CONSISTENT ESTIMATION OF A MULTIVARIATE DOUBLY TRUNCATED OR CENSORED TOBIT MODEL

by

Lung-Fei Lee

Discussion Paper No. 81 - 153, August 1981

Abstract

A multivariate regression model with some of the dependent variables doubly truncated normal is specified. A recursive relation for the moments of the truncated multivariate normal distribution is derived. Based on the recursive relations, some computationally simple and consistent instrumental variables estimation procedures are proposed. The asymptotic distribution of the estimators is analyzed.

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

CONSISTENT ESTIMATION OF A MULTIVARIATE DOUBLY TRUNCATED OR CENSORED NORMAL TOBIT MODEL

Lung-Fei Lee*

1. Introduction

Amemiya [1974] extended Tobin's model [1958] to the Multivariate Regression and Simultaneous Equation models when all the dependent variables are singly censored. Subsequent works of Sickles and Schmidt [1978] and Lee [1979] extended the models to the case where only some of the dependent variables are singly censored. Applications of those models in economics can be found in Sickles, Schmidt and Witte [1979], Goldfeld and Quandt [1975] (see also Lee [1976]) and Waldman [1981]. As the likelihood functions for the models are quite complicated, maximum likelihood methods are not attractive. Fortunately, computationally simple and consistent estimators for those models are available as in Amemiya [1974], Sickles and Schmidt [1978] and Lee [1979]. These consistent estimators are instrumental variables estimators which are derived from the relationships between the first two moments of the singly truncated multi-normal distribution. Subsequent generalization of the Tobit model is in Rosett and Nelson [1975]. They extend the single equation Tobit model to allow double censoring and have studied the maximum likelihood estimation.

In this article, we generalize the above models to the multivariate regression models with both singly and doubly truncated or censored dependent variables. We will derive some recursive formulae for the moments of the truncated normal distribution and suggest a consistent instrumental variable estimator for this general model. The derivations of the recursive formulae are very simple and are readily applicable to many truncated or censored non-normal multivariate distributions.

2. The Model

The multivariate regression model with N multivariate normally distributed disturbances is

$$Y_t^* = \alpha x_t + \varepsilon_t \qquad t = 1,\ldots,T \tag{2.1}$$

where x_t is a k x 1 exogenous variable vector, Y_t^* is a N x 1 vector of underlying unobservable dependent variable. α is a N x k matrix of unknown coefficients and $\varepsilon_t \sim$ i.i.d. $N(0, \Sigma)$ where Σ is a positive definite matrix. The Y_t^*, $t=1,\ldots,T$, are not completely observed since there is censoring. The observed vectors of the dependent variables Y_t, $t=1,\ldots,T_1$ where $T_1 < T$ are censored samples. The sample Y_t is observed and $Y_t = Y_t^*$ if and only if it satisfies the conditions (i) and (ii):

(i) $0 < Y^*_{G+j,t} < k_{G+j}, \qquad j = 1,\ldots,J;$

(ii) $0 < Y^*_{G+J+\ell,t} < \infty, \qquad \ell = 1,\ldots,N-G-J$

where $Y_t^{*\prime} = (Y^*_{1t},\ldots, Y^*_{Gt}\ Y^*_{G+1t},\ldots,Y^*_{G+Jt}, Y^*_{G+J+1t},\ldots,Y^*_{Nt})$, $0 \le G \le N$ and $0 \le J \le N$. If any conditions in (i) or (ii) are not satisfied, Y_t^* is unobserved. The upper truncation points k_{G+j} and the lower truncation point 0 are known constants in the model.[1/] To justify the asymptotic properties of the estimator, we assume the following regularity conditions: (1) the parameter space H is compact and constains the true parameter vector, vec(α, Σ), as an interior,[2/] (2) the exogenous variables vectors x_t have bounded range, and the empirical distribution functions of x defined by the counting measure converges to a limiting distribution as T goes to infinity, and (3) $\lim_{T\to\infty} \frac{1}{T} \Sigma_{t=1}^{T} x_t x_t'$ is positive definite.

This general model contains the singly truncated multivariate regression models in Amemiya [1974] and Lee [1979] as special cases. When $G = 0$ and $J = 0$, all the dependent variables in Y_t are singly truncated and it is the model in Amemiya [1974]. When $J = 0$ and $1 \le G < N$, it corresponds to the model in Lee [1979].

The above model is specified for the censored sampling design in which all the T observations on x_t are available and we know that $T - T_1$ samples on Y_t^* of the T random samples are censored out. There is another sampling situation, namely the truncated sampling design, in which the T random samples are drawn from the populations in which all the conditions (i) and (ii) are always satisfied. For the truncated samples, all the observations x_t and Y_t are available, $T_1 = T$ but the disturbances ε_t in

$$Y_t = \alpha x_t + \varepsilon_t \qquad t = 1,\ldots,T$$

will be distributed as truncated normal with density function

$$f(\varepsilon|x) = (2\pi)^{-N/2}|\Sigma|^{-1/2} \exp\{-\tfrac{1}{2}\,\varepsilon'\Sigma^{-1}\varepsilon\}/P(x),$$

defined on $-\alpha_i x < \varepsilon_i < k_i - \alpha_i x$, $i = G+1,\ldots,G+J$ and $-\alpha_j x < \varepsilon_j < \infty$ for $j = G+J+1,\ldots,N$, where

$$P(x) = \int_{-\infty}^{\infty}\!\cdots\!\int_{-\infty}^{\infty} \int_{-\alpha_{G+1}x}^{k_{G+1}-\alpha_{G+1}x} \cdots \int_{-\alpha_{G+J}x}^{k_{G+J}-\alpha_{G+J}x} \int_{-\alpha_{G+J+1}x}^{\infty} \cdots$$

$$\int_{-\alpha_N x}^{\infty} (2\pi)^{-N/2}|\Sigma|^{-1/2} \exp\{-\tfrac{1}{2}\,\varepsilon'\Sigma^{-1}\varepsilon\}d\varepsilon$$

The censored samples and the truncated samples cases are apparently different in their likelihood functions. However, for the instrumental variable estimation procedures that we will be considered, they are applicable to both cases.[3/]

3. Differential Equations and the Truncated Moments

In Amemiya [1974] and Lee [1979] for the singly truncated multivariate normal distributions, explicit expressions for the first two moments around zero are derived and the relationship between these two moments are then developed. The arguments involved are complicated and will involve tedious generalizations. In this section, we adopt other arguments which are rather simple, allow us to derive a very simple relationship between moments of any orders for our general model, and are easily extended to a variety of multivariate distributions. Our analysis will proceed for the censored sampling case, but the final results are applicable to both censored and truncated samples.

Consider a N-multivariate normal random variable u with density function

$$f(u) = (2\pi)^{-\frac{N}{2}}|\Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(u-\mu)'\Sigma^{-1}(u-\mu)\} \tag{3.1}$$

It follows immediately from (3.1) that the density function is characterized by the differential equations,

$$\frac{d}{du}(\ell n f(u)) = -\Sigma^{-1}(u-\mu). \tag{3.2}$$

Let σ^i be the ith column of Σ^{-1} and u_i be the ith component of the vector u, i=1,...,N. It follows that

$$\frac{\partial}{\partial u_i} f(u) = -\sigma^{i'}(u-\mu)f(u) \tag{3.3}$$

Premultiplying each side by the ℓth power u_i^ℓ of u_i, we have

$$u_i^\ell \frac{\partial f(u)}{\partial u_i} = - \sigma^{i'}(u - \mu)\, u_i^\ell\, f(u) \tag{3.4}$$

and hence

$$\int_{a_1}^{b_1} \cdots \int_{a_N}^{b_N} u_i^\ell \frac{\partial f(u)}{\partial u_i}\, du_N \ldots du_1 = - \sigma^{i'}(E^*(u_i^\ell u) - \mu E^*(u_i^\ell)) \tag{3.5}$$

where $E^*(u_i^\ell\, u) \equiv \int_{a_1}^{b_1} \cdots \int_{a_N}^{b_N} u_i^\ell\, u\, f(u) du_N \ldots du_1$, etc., denote the incomplete moments. Since, by the integration by parts,

$$\int_{a_i}^{b_i} u_i^\ell \frac{\partial f(u)}{\partial u_i}\, du_i = u_i^\ell\, f(u)\Big|_{u_i=a_i}^{b_i} - \ell \int_{a_i}^{b_i} u_i^{\ell-1}\, f(u) du_i\ ,$$

we have

$$\int_{a_1}^{b_1} \cdots \int_{a_N}^{b_N} u_i^\ell \frac{\partial f(u)}{\partial u_i}\, du_N \ldots du_1 = b_i^\ell\, F_i(b_i) - a_i^\ell\, F_i(a_i) - \ell E^*(u_i^{\ell-1}) \tag{3.6}$$

where $F_i(c) \equiv \int_{a_1}^{b_1} \cdots \int_{a_{i-1}}^{b_{i-1}} \int_{a_{i+1}}^{b_{i+1}} \cdots \int_{a_N}^{b_N} f(u_1,\ldots,u_{i-1},c,u_{i+1},\ldots,u_N) du_1 \ldots$

$du_{i-1} du_{i+1} \ldots du_N$. It follows from the expressions in (3.5) and (3.6) that

$$b_i^\ell\, F_i(b_i) - a_i^\ell\, F_i(a_i) - \ell E^*(u_i^{\ell-1}) = -\sigma^{i'} E^*(u_i^\ell u) + \sigma^{i'} \mu\, E^*(u_i^\ell) \tag{3.7}$$

for $\ell = 0,1,2,\ldots,\infty$ and $i=1,\ldots,N$. These simple recursive formulae provide the basis for the relationships of the moments of the truncated distribution.

Let the probability, $\int_{a_i}^{b_i} F_i(c) dc$, be denoted by F and $E(u_i^\ell)$ and $E(u_i^\ell u)$ be the moments of u_i^ℓ and $u_i^\ell u$ respectively of the truncated distribution with density function $\frac{f(u)}{F}$ defined on $\mathop{X}\limits_{i=1}^{N} [a_i,b_i]$. It follows immediately, from the definitions, that $E^*(u_i^0) = F$, $E(u_i^0) = 1$, $E(u_i^\ell) = E^*(u_i^\ell)/F$ and $E(u_i^\ell u) = E^*(u_i^\ell u)/F$. From (3.7), we have the following recursive formulae for the moments of the truncated normal distribution,

$$\sigma^{i'} E(u_i^\ell u) = \sigma^{i'} \mu\, E(u_i^\ell) + \ell E(u_i^{\ell-1}) + \frac{a_i^\ell F_i(a_i)}{F} - \frac{b_i^\ell F_i(b_i)}{F} \qquad (3.8)$$

$$\ell = 0,1,2,\ldots$$

For the model of (2.1), $\mu = \alpha x$ and (3.8) can be divided into three categories:

(1) $i \varepsilon \{1,\ldots,G\}$. For these equations, $a_i = -\infty$ and $b_i = \infty$, and the terms $a_i^\ell\, F_i(a_i)$ and $b_i^\ell\, F_i(b_i)$ vanish for all $\ell = 0,1,2,\ldots$. Thus (3.8) becomes

$$\sigma^{i'} E(Y) = \sigma^{i'} \alpha x \qquad \ell = 0 \qquad (3.9)$$

and

$$\sigma^{i'} E(Y_i^\ell Y) = \sigma^{i'} \alpha x\, E(Y_i^\ell) + \ell E(Y_i^{\ell-1}) \qquad \ell = 1,2,\ldots, \qquad (3.10)$$

(2) $i \varepsilon \{G+J+1,\ldots,N\}$. In this case, $a_i = 0$ and $b_i = \infty$. Here (3.8) is

$$\sigma^{i'} E(Y) = \sigma^{i'} \alpha x + \frac{F_i(0)}{F} \qquad \ell = 0 \qquad (3.11)$$

and

$$\sigma^{i'} E(Y_i^\ell Y) = \sigma^{i'} \alpha x\, E(Y_i^\ell) + \ell E(Y_i^{\ell-1}) \qquad \ell = 1,2,\ldots, \qquad (3.12)$$

(3) $i \varepsilon \{G+1,\ldots,G+J\}$. In this case, $a_i = 0$ and $b_i = k_i$, and correspondingly (3.8) is

$$\sigma^{i'}E(Y) = \sigma^{i'}\alpha x + \frac{F_i(0)}{F} - \frac{F_i(k_i)}{F} \qquad \ell = 0 \tag{3.13}$$

and

$$\sigma^{i'}E(Y_i^{\ell}Y) = \sigma^{i'}\alpha x\, E(Y_i^{\ell}) + \ell E(Y_i^{\ell-1}) - k_i^{\ell}\,\frac{F_i(k_i)}{F}\,, \quad \ell = 1,2,\ldots. \tag{3.14}$$

In contrast to models with single truncations, the terms $k_i^{\ell}\,\frac{F_i(k_i)}{F}$ in (3.14) do not vanish. However, the equations (3.14) do obviously imply the following recursive relationship:

$$\sigma^{i'}E(Y_i^{\ell}Y) = k_i\, \sigma^{i'}E(Y_i^{\ell-1}Y) + \sigma^{i'}\alpha x(E(Y_i^{\ell}) - k_iE(Y_i^{\ell-1})) + \ell E(Y_i^{\ell-1}) - k_i(\ell-1)E(Y_i^{\ell-2}) \qquad \ell = 2,3,\ldots, \tag{3.15}$$

which will be used to derive instrumental variable estimators.

4. Consistent Instrumental Variables Estimator

With the relations (3.10), (3.12) and (3.15), the instrumental variables estimation procedure first appearing in Amemiya [1974] can be extended to our general model. For the equations i, $i \varepsilon \{1,\ldots,G\} \cup \{G+J+1,\ldots,N\}$, the relations in (3.10) and (3.12) imply that the conditional means of the disturbances $\eta_{it}^{(\ell)}$ in the following equations are zero:

$$\sigma^{i'}Y_{it}^{\ell}Y_t = \sigma^{i'}\alpha x Y_{it}^{\ell} + \ell Y_{it}^{\ell-1} + \eta_{it}^{(\ell)} \qquad \ell = 1,2,\ldots,\quad t = 1,\ldots,T_1 \tag{4.1}$$

where $E(\eta_{it}^{(\ell)}) = 0$. For the equations i, $i \varepsilon \{G+1,\ldots,G+J\}$, we have

$$\sigma^{i'}(Y_{it}^{\ell}Y_t - k_iY_{it}^{\ell-1}Y_t) = \sigma^{i'}\alpha x(Y_{it}^{\ell} - k_iY_{it}^{\ell-1}) + (\ell Y_{it}^{\ell-1} - k_i(\ell-1)Y_{it}^{\ell-2}) + \eta_{it}^{(\ell)} \qquad \ell = 2,3,\ldots;\quad t = 1,\ldots,T_1 \tag{4.2}$$

where $E(\eta_{it}^{(\ell)}) = 0$. Thus, double truncation amounts to the appearance of additional terms in $\ell-1$ and $\ell-2$ powers of Y_{it}.

Let $\sigma^{i'} = (\sigma^{i1},\ldots,\sigma^{iN})$. For the convenience of estimation, the equations (4.1) and (4.2) can be rearranged as follows:

$$Y_{it}^{\ell+1} = \frac{1}{\sigma^{ii}}\,\ell Y_{it}^{\ell-1} + \frac{1}{\sigma^{ii}}\,\sigma^{i'}\alpha x Y_{it}^{\ell} - \frac{1}{\sigma^{ii}}\sum_{\substack{j=1\\ j\neq i}}^{N} \sigma^{ij} Y_{it}^{\ell} Y_{jt} + \frac{1}{\sigma^{ii}}\,\eta_{it}^{(\ell)} \qquad (4.1)'$$

$$\ell = 1,2,\ldots,;\quad i = 1,\ldots,G,G+J+1,\ldots,N$$

$$(Y_{it} - k_i)Y_{it}^{\ell} = \frac{1}{\sigma^{ii}}\,(\ell Y_{it} - k_i(\ell-1))Y_{it}^{\ell-2} + \frac{1}{\sigma^{ii}}\,\sigma^{i'}\alpha x(Y_{it}-k_i)Y_{it}^{\ell-1}$$

$$- \frac{1}{\sigma^{ii}}\sum_{\substack{j=1\\ j\neq i}}^{N} \sigma^{ij}(Y_{it} - k_i)Y_{it}^{\ell-1}Y_{jt} + \frac{1}{\sigma^{ii}}\,\eta_{it}^{(\ell)} \qquad (4.2)'$$

$$\ell = 2,3,\ldots,;\quad i = G+1,\ldots,G+J$$

To derive an instrumental variables estimator, consider separately each ℓ for the equations i, $i\varepsilon\{1,\ldots,G\} \cup \{G+J+1,\ldots,N\}$ and the corresponding $\ell+1$ for the remaining equations with $\ell \geq 1$.[4/] For example, for $\ell=1$, we are considering the estimation of α and Σ from the following equations:

$$Y_{it}^{2} = \frac{1}{\sigma^{ii}} + \frac{1}{\sigma^{ii}}\,\sigma^{i'}\alpha\, xY_{it} - \frac{1}{\sigma^{ii}}\sum_{\substack{j=1\\ j\neq i}}^{N} \sigma^{ij} Y_{it}Y_{jt} + \frac{1}{\sigma^{ii}}\,\eta_{it}^{(1)} \qquad (4.3)$$

$$i = 1,\ldots,G,G+J+1,\ldots,N$$

and

$$(Y_{it}-k_i)Y_{it}^{2} = \frac{1}{\sigma^{ii}}\,(2Y_{it}-k_i) + \frac{1}{\sigma^{ii}}\,\sigma^{i'}\alpha\, x(Y_{it}-k_i)Y_{it}$$

$$- \frac{1}{\sigma^{ii}}\sum_{j\neq i}^{N} \sigma^{ij}(Y_{it}-k_i)Y_{it}Y_{jt} + \frac{1}{\sigma^{ii}}\,\eta_{it}^{(2)} \qquad (4.4)$$

$$i = G+1,\ldots,G+J$$

To simplify the notations, let us write

$$\delta_i = (\frac{1}{\sigma^{ii}})(-\sigma^{i1},\ldots,-\sigma^{i(i-1)},1,\ -\sigma^{i(i+1)},\ldots,-\sigma^{iN},\ \sigma^{i'}\alpha)'$$

For each ℓ, $\ell \geq 1$, define

$$z_{it}^{(\ell)} \begin{cases} = (Y_{it}^{\ell}(Y_{1t},\ldots,Y_{i-1t}),\ \ell Y_{it}^{\ell-1},\ Y_{it}^{\ell}(Y_{i+1t},\ldots,Y_{Nt},x'))' & \text{for } i=1,\ldots,G,G+J+1,\ldots,N \\ = ((Y_{it}-k_i)Y_{it}^{\ell}(Y_{1t},\ldots,Y_{i-1t}),((\ell+1)Y_{it}-k_i\ell)Y_{it}^{\ell-1}, & \\ \quad (Y_{it}-k_i)Y_{it}^{\ell}(Y_{i+1t},\ldots,Y_{Nt},x'))' & \text{for } i=G+1,\ldots,G+J \end{cases}$$

and

$$q_{it}^{(\ell)} \begin{cases} = Y_{it}^{\ell+1} & \text{for } i=1,\ldots,G,G+J+1,\ldots,N \\ = Y_{it}^{\ell+1}(Y_{it}-k_i) & \text{for } i=G+1,\ldots,G+J. \end{cases}$$

Then we have

$$q_{it}^{(\ell)} = \delta_i' z_{it}^{(\ell)} + \xi_{it}^{(\ell)} \qquad i=1,\ldots,N\ ;\ t=1,\ldots,T_1 \qquad (4.5)$$

where $\xi_{it}^{(\ell)}$ are the corresponding disturbances. As suggested in Amemiya [1974], the instrumental variables $\hat{z}_{it}^{(\ell)}$ can be constructed by regressing each Y_{it} $t=1,\ldots,T_1$ on x_t and certain higher powers of the elements of x_t and substituting the least squares predictor $\hat{Y}_{it}$ for Y_{it} in $z_{it}^{(\ell)}$. The instrumental variable estimator of δ_i is then

$$\hat{\delta}_i^{(\ell)} = (\Sigma_{t=1}^{T_1} \hat{z}_{it}^{(\ell)} z_{it}^{(\ell)'})^{-1} \Sigma_{t=1}^{T_1} \hat{z}_{it}^{(\ell)} q_{it}^{(\ell)} \qquad (4.6)$$

for each i. Under the previous assumptions and the assumption that $\underset{T\to\infty}{\text{plim}} \frac{1}{T} \Sigma_{t=1}^{T} \hat{z}_{it}^{(\ell)} z_{it}^{(\ell)'}$ is non singular for the censored sampling case or $\underset{T\to\infty}{\text{plim}} \frac{1}{T} \Sigma_{t=1}^{T} \hat{z}_{it}^{(\ell)} z_{it}^{(\ell)'}$ for the truncated sampling case, the consistency and asymptotic normality follow as in Amemiya [1973]. The disturbances $\xi_{it}^{(\ell)}$ are obviously heteroscedastic and the explicit expressions will depend on the high moments and the joint multivariate normal probability. The analytical asymptotic covariance matrix is thus quite complicated to be computed as the computation of the multivariate normal probability involves numerical multiple integrals. This difficulty, however, can be overcome if we use the sample moments as an alternative. It is easy to show by lemma 2 in Amemiya ([1973], p. 1002) that, under the specified regularity conditions, the asymptotic covariance of $\hat{\delta}_i^{(\ell)}$ can be estimated by

$$(\Sigma_{t=1}^{T} \hat{z}_{it}^{(\ell)} z_{it}^{(\ell)'})^{-1} \Sigma_{t=1}^{T} (\hat{\xi}_{it}^{(\ell)})^2 \hat{z}_{it}^{(\ell)} \hat{z}_{it}^{(\ell)'} (\Sigma_{t=1}^{T} z_{it}^{(\ell)'} \hat{z}_{it}^{(\ell)})^{-1} \quad (4.7)$$

where $\hat{\xi}_{it}^{(\ell)} = q_{it}^{(\ell)} - \hat{\delta}_i^{(\ell)} z_{it}^{(\ell)}$.[5/] Similarly, the asymptotic covariance of $\hat{\delta}_i^{(\ell)}$ and $\hat{\delta}_j^{(\ell)}$ can be estimated by

$$(\Sigma_{t=1}^{T} \hat{z}_{it}^{(\ell)} z_{it}^{(\ell)'})^{-1} \Sigma_{t=1}^{T} \hat{\xi}_{it}^{(\ell)} \hat{\xi}_{jt}^{(\ell)} (\Sigma_{t=1}^{T} z_{jt}^{(\ell)'} \hat{z}_{jt}^{(\ell)})^{-1} \quad (4.8)$$

Since there is a one to one correspondence between the parameter vector $\delta = (\delta_1', \ldots, \delta_N')'$ and the parameter vector vec (α, Σ) where the symmetry in Σ are not imposed. Thus one can derive a consistent estimator $(\hat{\alpha}^{(\ell)}, \hat{\Sigma}^{(\ell)})$ for (α, Σ). The asymptotic covariance of $(\hat{\alpha}^{(\ell)}, \hat{\Sigma}^{(\ell)})$ can then be derived from (4.7) and the transformation between δ and vec (α, Σ) as follows.

Let δ_{ij} be the jth component of δ_i, $j=1,\ldots,N$ and $\delta_{i*} = \alpha'\sigma^i$. It follows that

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\delta_{11}} & -\frac{\delta_{21}}{\delta_{22}} & \cdots & -\frac{\delta_{N1}}{\delta_{NN}} \\ -\frac{\delta_{12}}{\delta_{11}} & \frac{1}{\delta_{22}} & \cdots & -\frac{\delta_{N2}}{\delta_{NN}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\delta_{1N}}{\delta_{11}} & -\frac{\delta_{2N}}{\delta_{22}} & \cdots & \frac{1}{\delta_{NN}} \end{bmatrix}$$

Substituting the estimates $\hat{\delta}_{ij}^{(\ell)}$ for δ_{ij}, we have a consistent estimate $(\widehat{\Sigma^{-1}})^{(\ell)}$ of the matrix Σ^{-1}. To simplify notations, we will suppress the superscript (ℓ) in this paragraph. The inverse of $\widehat{\Sigma^{-1}}$ provides an estimate $\hat{\Sigma}$ of Σ. Let $\delta_* = [\delta_{1*}, \ldots, \delta_{N*}]$. Since $\delta_* = \alpha'\Sigma^{-1}$, a consistent estimate of α is $\hat{\alpha} = (\widehat{\Sigma^{-1}})' \hat{\delta}_*'$. To derive the asymptotic covariance matrices of $\hat{\Sigma}$ and $\hat{\alpha}$, it is convenient to derive the asymptotic covariance matrix of $\widehat{\Sigma^{-1}}$ first. Let $\underline{\delta}_i = (\delta_{i1}, \ldots, \delta_{iN})'$ be the sub-vector of δ_i consisting of the first N components. Furthermore, let $\underline{\delta} = (\underline{\delta}_1', \underline{\delta}_2', \ldots, \underline{\delta}_N')'$. By the Taylor expansion,

$$\mathrm{vec}(\widehat{\Sigma^{-1}}) - \mathrm{vec}(\Sigma^{-1}) \overset{D}{=} \frac{\partial\, \mathrm{vec}(\Sigma^{-1})}{\partial\, \underline{\delta}'} (\hat{\underline{\delta}} - \underline{\delta}) \qquad (4.9)$$

where $\overset{D}{=}$ means that the expressions on both sides multiplied by $\sqrt{T_1}$ have the same asymptotic distribution. With some tedious algebras, it is straightforward to show that $\frac{\partial\, \mathrm{vec}(\Sigma^{-1})}{\partial\, \underline{\delta}'}$ is a block diagonal matrix;

$$\frac{\partial\ \mathrm{vec}(\Sigma^{-1})}{\partial \underline{\delta}'} = \begin{bmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & \ddots \quad A_N \end{bmatrix} = A, \text{ say},$$

where

$$A_i = -\frac{1}{\delta_{ii}} I_N + \frac{1}{\delta_{ii}^2} (e_i' \otimes \delta_i - e_i' \otimes e_i), \quad i=1,\ldots,N;$$

I_N is the $N \times N$ identity matrix and $e_i = (0,\ldots,0,1,0,\ldots,0)$ is the ith unit vector of the N dimensional Euclidean space. Since

$$\mathrm{vec}(\hat{\Sigma}) - \mathrm{vec}(\Sigma) \stackrel{D}{=} \frac{\partial\ \mathrm{vec}(\Sigma)}{\partial\ \mathrm{vec}(\Sigma^{-1})'} (\mathrm{vec}(\widehat{\Sigma^{-1}}) - \mathrm{vec}(\Sigma^{-1})) \text{ and}$$

$$\frac{\partial\ \mathrm{vec}(\Sigma)}{\partial\ \mathrm{vec}(\Sigma^{-1})} = -(\Sigma' \otimes \Sigma), \text{ it follows}$$

$$\mathrm{vec}(\hat{\Sigma}) - \mathrm{vec}\ (\Sigma) \stackrel{D}{=} -(\Sigma' \otimes \Sigma)A(\hat{\underline{\delta}} - \underline{\delta}) \tag{4.10}$$

Let $\Omega_{\hat{\underline{\delta}}}$ be the asymptotic covariance matrix of $\hat{\underline{\delta}}$ as appropriately constructed from (4.7) and (4.8). The asymptotic covariance matrix of $\mathrm{vec}(\hat{\Sigma})$ is $(\Sigma' \otimes \Sigma)A\Omega_{\hat{\underline{\delta}}}A'(\Sigma \otimes \Sigma')$. Since $\hat{\alpha} = (\widehat{\Sigma^{-1}})'\hat{\delta}_*'$, the asymptotic covariance matrix of $\hat{\alpha}$ can be derived as follows. By the Taylor expansion,

$$\mathrm{vec}(\hat{\alpha}') - \mathrm{vec}(\alpha') \stackrel{D}{=} \frac{\partial\ \mathrm{vec}(\alpha')}{\partial \delta'} (\hat{\delta} - \delta) \tag{4.11}$$

where $\delta = (\delta_1', \delta_2', \ldots, \delta_N')'$. Since $\mathrm{vec}(\alpha') = (\Sigma^{-1'} \otimes I_N)\mathrm{vec}(\delta_*) = (I_N \otimes \delta_*)\mathrm{vec}(\Sigma^{-1})$, it follows $\frac{\partial\ \mathrm{vec}(\alpha')}{\partial \delta'} = (I_N \otimes \delta_*) \frac{\partial\ \mathrm{vec}(\Sigma^{-1})}{\partial \delta'} + (\Sigma^{-1'} \otimes I_N) \frac{\partial\ \mathrm{vec}(\delta_*)}{\partial \delta'}$. Furthermore, we can show that

$$\frac{\partial\ \mathrm{vec}(\delta_*)}{\partial\delta'} = I_N \otimes [\underline{0}, I_k]$$

and

$$\frac{\partial\ \mathrm{vec}(\Sigma^{-1})}{\partial\delta'} = \begin{bmatrix} (A_1,\underline{0}') & & & 0 \\ & (A_2,\underline{0}') & & \\ & & \ddots & \\ 0 & & & (A_N,\underline{0}') \end{bmatrix} = B, \quad \text{say;}$$

where $\underline{0}$ denotes the zero matrix of dimension $k \times N$.

Hence

$$\mathrm{vec}(\hat{\alpha}') - \mathrm{vec}(\alpha') \overset{D}{=} [(I_N \otimes \delta_*)B + (\Sigma^{-1'} \otimes I_N)(I_N \otimes [\underline{0}, I_k])](\hat{\delta} - \delta) \quad (4.12)$$

Let $\Omega_{\hat{\delta}}$ be the asymptotic covariance matrix of $\hat{\delta}$. The asymptotic covariance matrix of $\mathrm{vec}(\hat{\alpha}')$ is

$$[(I_N \otimes \delta_*)B + (\Sigma^{-1'} \otimes I_N)(I_N \otimes [\underline{0}, I_k])]\Omega_{\hat{\delta}}[B'(I_N \otimes \delta_*') + (I_N \otimes \begin{bmatrix} \underline{0}' \\ I_k \end{bmatrix})(\Sigma^{-1} \otimes I_N)].$$

The estimation described above provides a consistent estimate $(\hat{\alpha}^{(\ell)}, \hat{\Sigma}^{(\ell)})$ of (α, Σ) for each ℓ, $\ell \geq 1$. In practice, it is sufficient to use several ℓ of low orders for the estimation purpose without loss of much information. The instrumental variables estimation in Amemiya [1974] corresponds to the case that uses only $\ell = 1$. Suppose we have used the equations in (4.5) for $\ell = 1,\ldots,m$ and derived the corresponding estimates $\hat{\alpha}^{(\ell)}$ and $\mathrm{vec}(\hat{\Sigma}^{(\ell)})$ for each ℓ, it is possible to pool the estimators so as to derive a more efficient one. A relatively simple pooling procedure is to use the generalized least square (mixed estimation) procedure. The set of estimates $\hat{\theta}^{(\ell)} = (\hat{\alpha}^{(\ell)\prime}, \text{vec}(\hat{\Sigma}^{(\ell)\prime}))'$ can be rearranged as

$$\begin{bmatrix} \hat{\theta}^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \hat{\theta}^{(m)} \end{bmatrix} = \begin{bmatrix} I \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ I \end{bmatrix} \theta + \begin{bmatrix} \xi_1 \\ \cdot \\ \cdot \\ \cdot \\ \xi_m \end{bmatrix} \tag{4.13}$$

where $\theta = (\alpha', \text{vec}(\Sigma)')'$ and $\xi_i = \hat{\theta}^{(i)} - \theta$. The asymptotic covariance matrices of the ξ's can be derived in a similar fashion as described in previous paragraphs. The generalized least squares procedure is then applied to (4.13).

Finally, it should be noted that since the above estimation procedure does not impose positive-definiteness property on Σ, the derived estimates of Σ need not be positive definite. This difficulty will not create problems if the main interest is to estimate the vector of coefficients $\text{vec}(\alpha)$. Some discussions on this problem are in Amemiya [1974].

5. Remarks

The approaches described in this article are ready to be generalized to other models where the distributions are characterized by differential equations in the form $\frac{d}{du} \ell n\, f(u) = Q_1(u)/Q_2(u)$ where $Q_1(u)$ and $Q_2(u)$ are finite order polynomials. Apparently, this includes the family of multivariate distributions of van Uven which generalizes the univariate Pearson family of distributions to the multivariate case. The multivariate normal distribution is an important member of the van Uven family. Many other popular multivariate distributions also belong to this family, see, e.g. Elderton and Johnson [1969].

Finally, it is useful to point out some related works in the literature. For the univariate Pearson family of distributions, Pearson (see the citations in Elderton and Johnson [1969], p. 36 and Johnson and Kotz [1970], pp. 9-15) derived the recursive relations of the complete moments and suggested the use of the method of moments to estimate the parameters. Subsequently, Cohen [1951, 1953] derived the recursive relations for the truncated Pearson distributions. The main concern in those articles is to estimate the unknown parameters and the distribution. Amemiya [1973, 1974] is apparently the first one to suggest the use of instrumental variables method to estimate the truncated normal regression model. In a recent article of Hartley and Swanson [1980], for the doubly truncated univariate normal regression model of Rossett and Nelson [1975], they notice the recursive formulae and suggest an instrumental variables estimation. In Lee [1981], based on recursive formulae for the truncated Pearson family of distributions, non-linear two stage least squares estimations are developed for the Tobit models with Pearson family of distributions. The potential usefulness of the recursive formulae of moments in other fields such as decision sciences can be found in Winkler et al [1972].

Footnotes

(*) An earlier preliminary version of this article was presented at the CEME Conference on Qualitative Decision Theory and Discrete Data Analysis at Harvard University on April 2-4, 1981. I appreciate receiving valuable comments from Jeff Dubin, Dale Poirier and Jim Adams. Financial support from the National Science Foundation under grants SES-8006481 to the University of Minnesota are gratefully acknowledged. Any errors are of my own. This article was partly completed during my visit to the Center for Econometrics and Decision Sciences, University of Florida, Gainesville, 1980-1981.

1. More general specification of the model can allow the truncation points to be varying across samples. Our estimation procedure is applicable to this situation with slight modifications.

2. $\text{vec}(\alpha,\Sigma)$ denotes the operator 'vec' applied to the block matrix $[\alpha,\Sigma]$, i.e., arrange the columns of the matrix $[\alpha,\Sigma]$ into a large column vector.

3. This is so, since the estimation method will utilize the same subset of sample information in both models.

4. Theoretically, one can set up a system of equations from (4.1)' and (4.2)' and estimate the parameters by the generalized three stage least squares procedure. The difficulty in this procedure is the computational complication in the evaluation of the covariance matrix of the system. One can easily show that the covariance matrix involves the evaluation of the multivariate normal probabilities F. Unless the probability F is of lower order, the computation is rather complicated.

5. This sample covariance matrix is valid for both censoring and truncated sampling designs, even though the explicit analytical expressions for the covariance matrices of the limit distributions of $\sqrt{T}\,(\hat{\delta}_i^{(\ell)} - \delta_i)$ for the different sampling designs are different.

References

Amemiya, T. [1973], "Regression Analysis When the Dependent Variable is Truncated Normal," Econometrica 41, pp. 997-1016.

Amemiya, T. [1974], "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal," Econometrica 42, pp. 999-1012.

Cohen, A. C., Jr., [1951], "Estimation of Parameters in Truncated Pearson Frequency Distributions," Annals of Mathematical Statistics 22, pp. 256-265.

Cohen, A. C., Jr., [1953], "Estimating Parameters in Truncated Pearson Frequency Distributions Without Resort to Higher Moments," Biometrika 40, pp. 50-57.

Elderton, W. P. and N. J. Johnson [1969], Systems of Frequency Curves, Cambridge University Press.

Goldfeld, S. M. and R. Quandt [1975], "Estimation in a Disequilibrium Model and the Value of Information," Journal of Econometrics 3, pp. 325-348.

Hartley, M. J. and E. V. Swanson [1980], "Maximum Likelihood Estimation of the Truncated and Censored Normal Regression Models," Discussion Paper, Development Economics Department, World Bank.

Johnson, N. L. and S. Kotz [1970], Continuous Univariate Distributions - 1, Houghton Mifflin Co., Boston.

Lee, L. F. [1976], "Multivariate Regression and Simultaneous Equation Models with Some Dependent Variables Truncated," Discussion Paper No. 76-79, Department of Economics, University of Minnesota, Minneapolis, MN.

Lee, L. F. [1979], "On the First and Second Moments of the Truncated Multi-Normal Distribution and A Simple Estimator," Economics Letters 3, pp. 165-169.

Lee, L. F. [1981], "Estimation of Some Non-Normal Limited Dependent Variables Models," Manuscript, University of Florida.

Rosett, R. N. and F. D. Nelson [1975], "Estimation of the Two Limit Probit Regression Model," Econometrica 43:1, pp. 141-146.

Sickles, R. C. and P. Schmidt [1978], "Simultaneous Equations Models with Truncated Dependent Variables: A Simultaneous Tobit Model," Journal of Economics and Business 31, pp. 11-21.

Sickles, R. C., P. Schmidt and A. D. Witte [1979], "An Application of the Simultaneous Tobit Model: A Study of the Determinants of Criminal Recidivism," Journal of Economics and Business 31, pp. 166-171.

Tobin, J. [1958], "Estimation of Relationships for Limited Dependent Variables," Econometrica 26, pp. 24-26.

Waldman, D. M. [1981], "An Economic Interpretation of Parameter Constraints in a Simultaneous - Equations Model with Limited Dependent Variables," forthcoming in International Economic Review.

Winkler, R. L., G. M. Roodman and R. R. Britney [1972], "The Determination of Partial Moments," Management Science 19, No. 3, pp. 290-296.