MULTIVARIATE REGRESSION AND SIMULTANEOUS EQUATIONS MODELS WITH SOME DEPENDENT VARIABLES TRUNCATED

by

Lung-Fei Lee*

Discussion Paper No. 76-79

(Revised June 1977)

*Research financed by NSF Grant SOC-76-04356 to the University of Florida. I would like to thank Professor G. S. Maddala for helpful comments. Thanks also to R. Litterman for correcting my English.

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

MULTIVATIATE REGRESSION AND SIMULTANEOUS EQUATION MODELS
WITH SOME DEPENDENT VARIABLES TRUNCATED

1. Introduction

Amemiya (2) in his interesting paper extended Tobin's model (11) to the Multivatiate Regression and Simultaneous Equations models. He considered the models where all the dependent variables are truncated normal. In this paper, we try to extend his models to the case where only some of the dependent variables are truncated. We consider the existence, identification and estimation problems in these models. The identification problem in this model is different from the usual simultaneous equation model. The rank conditions are only sufficient but not necessary. The estimation procedures are two step maximum likelihood procedures. Initial consistent estimates which utilize modified least squares and instrumental variables are found. To demonstrate the possible empirical applications, we point out some disequilibrium market models are in this framework.

The paper is organized as follows: In section 2, we review Amemiya's procedures and make some extensions. In section 3, we consider the estimation problems in multiple regression models with some dependent variables truncated. In section 4 and 5, we examine the general simultaneous equation model. The existence, identification and estimation problems are considered in detail. In section 6, we point out some empirical examples and in the last section, we draw our conclusions.

2. Review of Amemiya's Estimation Procedures and its Extension

The simultaneous equations system with truncated dependent variables considered by Amemiya is specified equivalently as follows,

$$\Gamma y_t \geq BX_t + u_t$$

$$y_t \geq 0 \text{ and}$$

$$(\Gamma y_t)_i = (BX_t + u_t)_i \quad \text{whenever } y_{it} > 0$$

where i stands for the i^{th} component of the relevant vector. In this model, X_t is a K-dimensional exogeneous variable, $(u_{1t}, u_{2t}, \ldots, u_{nt})' \sim N(0,\Omega)$ and is temporally independent. The sample observations for the dependent variable y_t take only the non-negative values. The vector y_t of endogeneous variables observed with all the components positive is an interior solution of the system. Any observed y_t with some components zero are corner solutions. Amemiya proves that for any u_t and X_t, the system has a unique solution if every principal minor of Γ is positive, i.e., Γ is a P-matrix, (Gale and Naikaido [4]).

Due to the difficulty of using corner solutions, Amemiya proposes an estimation procedure which utilizes only the interior solutions observations. Define y_t^* as follows,

$$y_t^* = y_t \text{ if } y_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{nt} \end{pmatrix} > 0 \ ^{(*)}$$

$$y_t^* = 0 \text{ otherwise.}$$

Denote $A = \Gamma^{-1}B$, $\nu_t = \Gamma^{-1}u_t$ and $\Sigma = \text{var}(\nu_t)$. With the first and second moments for the truncated multivariate normal distribution found by Tallis (3). Amemiya derives a simple relation for the system.

$$y_{it}^{*2} = \delta_i' z_{it} + \eta_{it}, \quad i = 1, \ldots, n, \quad t\epsilon S_1$$

where $z_{it} = \left[1, y_{it}(y_{1t}, y_{2t}, \ldots, y_{(i-1)t}, y_{(i+1)t}, \ldots, y_{nt}, X_t')\right]'$ and

$$\delta_i' = (1/\sigma^{ii})(1, -\sigma^{i1}, -\sigma^{i2}, \ldots, -\sigma^{i(i-1)}, -\sigma^{i(i+1)}, \ldots, -\sigma^{in}, \sigma^{i'}A).$$

S_1 is the set of t for which $y_t > 0$ and $E(\eta_{it}) = 0$, $\forall\ t\epsilon S_1$. The $\sigma^{i'}$ is the i^{th} row of Σ^{-1} and σ^{ij} is the ij^{th} element of Σ^{-1}. With this relationship, Amemiya suggests estimation of the parameters δ_i by an instrumental variables method;

1. regress y_{it} on X_t and certain higher powers of X_t, $t\epsilon S_1$,
2. replace y_{it} by the least squares predictor $\tilde{y}_{it}$ in z_{it} to get $\tilde{z}_{it}$.

(*) $y_t > 0$ means $y_{it} > 0$, $\forall\ i = 1, \ldots, n$.

3. use $\tilde{z}_{it}$ as the instrumental variables to have

$$\tilde{\delta}_i = (\Sigma_{t\varepsilon S_1} \tilde{z}_{it} z'_{it})^{-1} \Sigma_{t\varepsilon S_1} y^2_{it} \tilde{z}_{it}$$

Under certain assumptions, $\tilde{\delta}_i$ is a consistent estimator of δ_i. Since δ_i, $i = 1,\ldots,n$ can be estimated consistently, it follows that Σ and A can be estimated consistently from the relations between Σ, A and δ_i. However, the asympotic covariance of A will be very complicated and cumbersome. So Amemiya suggests that one can use these as initial estimates and apply the OLS to a reduced form equation with the bias in the error term adjusted to estimate α_i. The asymptotic variance-covariance matrix of the estimates can then be easily obtained as the least squares estimates in the regression equation with the heteroscedastic errors. Under the usual rank condition on A, Γ and B are identifiable from A and the consistent estimates of Γ and B can then be solved from A. The asymptotic distribution of the elements of Γ and B can also be derived from the asymptotic distributions of the elements of $\Gamma^{-1}B$ by means of Taylor expansion. However, when one considers the approach in more detail, the asymptotic distribution derived in this way is not the exact one. To get the exact distribution, the distributions of the instrumental variables which are extremely complicated are involved.

The structural estimates can be solved uniquely from the reduced form parameters only if the system is exactly identified. If any one of the structural equations is overidentified, the structural parameter can not be solved uniquely from the reduced form parameters. Nevertheless with the consistent estimates derived from the instrumental procedure, a two stage procedure instead of the indirect least squares can be proposed.

Without loss of generalization, it is enough to consider the first equation,

$$y_{1t} + \gamma_{12}y_{2t} + \ldots + \gamma_{1n}y_{nt} = \beta_1{}'X_t + u_{1t}, \quad \forall t\varepsilon S_1$$

Since $y_t{}^* = AX_t + \nu_t$, $\forall\ t\varepsilon S_1$, we have

$$y_{1t} = -\gamma_{12}y_{2t} - \ldots - \gamma_{1n}y_{nt} + \beta_1{}'X_t + u_{1t}$$

$$= -\gamma_{12}(\pi_2X_t) - \ldots - \gamma_{1n}(\pi_nX_t) + \beta_1{}'X_t + \nu_{1t}, \quad \forall t\varepsilon S_1$$

where π_i is the i^{th} row of A.

By theorem 1 in Amemiya, we have $P = \int_{a_1}^{\infty} \int_{a_2}^{\infty} \ldots \int_{a_n}^{\infty} f(\lambda)d\lambda$,

$$PE(\nu_{1t}) = \Sigma_{q=1}^{n} \sigma_{1q}f_q(a_q)F_{(q)}, \quad F_{(q)} = \prod_{S\neq q} \int_{a_S}^{\infty} f_{(q)}(\lambda)d\lambda, \quad \forall t\varepsilon S_1$$

where f_q is the marginal density of the q^{th} variable of $N(0,\Sigma)$; $f_{(q)}$ is the joint conditional density of the remaining n-1 variables given that the q^{th} variable of $N(0,\Sigma)$ is equal to a_q, a_q is the q^{th} element of $-AX_t$ and σ_{iq} is the $(i,q)^{th}$ element of Σ.

Thus we have

$$y_{1t} = -\gamma_{12}(\pi_2X_t) - \ldots -\gamma_{1n}(\pi_nX_t) + \beta'_1X_{1t} + \Sigma_{q=1}^{n} \frac{\sigma_{1q}f_q(a_q)F_{(q)}}{P} + \xi_{1t},$$

$$\forall t\varepsilon S_1$$

where $E(\xi_t | t\varepsilon S_1) = 0$.

After substituting the relevant estimates from $\hat{A}$ which are derived from $\tilde{\delta}_i$ into $(-\pi_iX_t)$, $\frac{1}{P} f_qF_{(q)}$, we can regress y_{1t} on $(-\tilde{\pi}_iX_t)$, $\frac{1}{\tilde{P}} f_q(\tilde{a}_q)\tilde{F}_{(q)}$ to get the least square estimates of γ_{1i}, β_1 and σ_{iq}.

All these estimates can be shown to be consistent. However, the asymptotic variance-covariance matrix is again complicated. With these consistent estimates, one can use them as initial estimates and a two step maximum likelihood procedure which has well-known asymptotic properties can be used. The two step maximum likelihood procedure is not extremely difficult to compute since it does not require repeated iterations. The two step maximum likelihood estimate Θ^* for the set of all parameters Θ in the model is

$$\overset{*}{\Theta} = \tilde{\Theta} + \left[\frac{\partial \ln L}{\partial \Theta} \frac{\partial \ln L}{\partial \Theta'}\right]\Bigg|_{\tilde{\Theta}}^{-1} \frac{\partial \ln L}{\partial \Theta}(\tilde{\Theta})$$

where $L(\Theta)$ is the likelihood function of the model and $\tilde{\Theta}$ is the initial consistent estimates. The estimates Θ^* are asymptotically normal and efficient and their asymptotic covariance matrix can be consistently estimated by

$$\left[\frac{\partial \ln L}{\partial \Theta} \frac{\partial \ln L}{\partial \Theta'}\right]^{-1}$$

The two step maximum likelihood procedure involves only first order derivatives which are not difficult to compute.

3. Multiple Regression Model with Some Dependent Variables Truncated.

In Amemiya's multiple regression system, all the dependent variables are truncated normal. In this section, we extend his model to the case when not all but only some of the dependent variables are truncated.

The system we consider is,

$$y_{1t} = \alpha_1 X_t + \nu_{1t}$$

$$y_{2t} = \alpha_2 X_t + \nu_{2t}$$

$$\vdots$$

$$y_{Gt} = \alpha_G X_t + \nu_{Gt}$$

$$y_{G+1t} = \alpha_{G+1} X_t + \nu_{G+1t}$$

$$\vdots$$

$$y_{nt} = \alpha_n X_t + \nu_{nt}$$

We can observe the endogeneous variables y_t if and only if $y_{G+1t} > 0, ---, y_{nt} > o$. Otherwise $y_t = 0$.

In this system, X_t is a K- dimensional exogeneous variables, $(\nu_{1t}, ---, \nu_{nt}) \sim N(0, \Sigma)$ and it is temporally independent. We would like to consider this system since it has the simultaneous structures and in fact, it is the reduced form of the simultaneous equations system that we will consider later on. In this system, the truncation of some dependent variables determines the truncation of all the other dependent variables.

Denote

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1^2 & \sigma_{12} \cdots & \sigma_{1G} & \sigma_{1G+1} \cdots\cdots \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 \cdots & \sigma_{2G} & \sigma_{2G+1} \cdots\cdots \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots \qquad\qquad \vdots \\ \sigma_{G1} & \sigma_{G2} \cdots & \sigma_G^2 & \sigma_{GG+1} \cdots\cdots \sigma_{Gn} \\ \sigma_{G+11} & \sigma_{G+12} \cdots & \sigma_{G+1G} & \\ \vdots & \vdots & \vdots & \Sigma_{G+1,n} \\ \sigma_{n1} & \sigma_{n2} \cdots & \sigma_{nG} & \end{array}\right]$$

The subsystem $y_{G+1,t}, \ldots, y_{nt}$ is Amemiya's system and hence the $\alpha_{G+1}, \ldots, \alpha_n$ and $\Sigma_{G+L,n}$ can be estimated by the instrumental procedures.

The likelihood function of this system is

$$\prod_{t\varepsilon S_1} f_{1,n}(y_{1t}-\alpha_1 X_t,\ldots,y_{nt}-\alpha_n X_t) \prod_{t\varepsilon S_2}\Big[1-\int_{-\alpha_{G+1}X_t}^{\infty}\cdots\int_{-\alpha_n X_t}^{\infty} f_{G+1,n}(\nu_{G+1},\ldots,\nu_n)$$

$$\cdot d\nu_{G+1}\cdots d\nu_n\Big]$$

where $S_1 = \{t|y_{G+1,t} > 0,\ldots,y_{nt} > 0\}$, $S_2 = \{1,\ldots,T\}-S_1$

and f_{ij} is the marginal density function of the multi-normal variables $(\nu_i, \nu_{i+1},\ldots,\nu_j)$. If initial consistent estimates can be found, two step maximum likelihood procedures can then be applied. Amemiya's instrumental variables procedure gives consistent estimates for the last $n-G$ equations. So it remains to find the procedure to estimate the remaining G equations.

Without loss of generalization, it is enough to consider the first equation. Consider the truncated mean of ν_{1t}; it is

$$E(\nu_{1t}|t\varepsilon S_1) = \int_{-\alpha_n X_t}^{\infty}\cdots\int_{-\alpha_{G+1}X_t}^{\infty}\int_{-\infty}^{\infty} \nu_{1t}\frac{f(\nu_{1t},\nu_{G+1t},\ldots,\nu_{nt})}{P_t}\,d\nu_{1t}d\nu_{G+1t}\cdots d\nu_{nt}$$

where

$$P_t = \int_{-\alpha_n X_t}^{\infty}\cdots\int_{-\alpha_{G+1}X_t}^{\infty} f(\nu_{G+1,t},\ldots,\nu_{nt})d\nu_{G+1t}\cdots d\nu_{nt}$$

Thus it follows,

$$E(\nu_{1t}|t\varepsilon S_1) = \int_{-\alpha_n X_t}^{\infty}\cdots\int_{-\alpha_{G+1}X_t}^{\infty}\Big(\int_{-\infty}\nu_{1t}f(\nu_{1t}|\nu_{G+1t},\ldots,\nu_{nt})d\nu_{1t}\Big)$$

$$\times\frac{g(\nu_{G+1\ t},\ldots,\nu_{nt})}{P_t}\,d\nu_{G+1,t},\ldots,d\nu_{nt}$$

$$= \left[\sigma_{1_{G+1}},\ldots,\sigma_{1_n}\right]\Sigma_{G+1,n}^{-1}\ \ X$$

$$\left[E(\nu_{G+1,t}|t\varepsilon S_1), \ldots, E(\nu_{n,t}|t\varepsilon S_1)\right]',$$

since $\left[\nu_{1t}|\nu_{G+1,t}, \ldots, \nu_{nt}\right]$

$$\sim N\left(\left[\sigma_{1G+1}, \ldots, \sigma_{1n}\right] \Sigma^{-1}_{G+1,n} \begin{bmatrix} \nu_{G+1,t} \\ \\ \nu_{n,t} \end{bmatrix}, \sigma_1{}^2 - \left[\sigma_{1G+1}, \ldots, \sigma_{1n}\right] \Sigma^{-1}_{G+1,n} \begin{bmatrix} \sigma_{1G+1} \\ \vdots \\ \sigma_{1n} \end{bmatrix}\right)$$

From Amemiya, we have as before

$$E(\nu_{kt}|t\varepsilon S_1) = \Sigma^n_{q=G+1}\, \sigma_{kq} f_q(\alpha_q X_t) F_{(q)}, \quad k = G+1, \ldots, n$$

where

$$F_{(q)} = \pi_{S\neq q} \int_{-\alpha_S X_t}^{\infty} f_{(q)}(\lambda)\, d\lambda.$$

The notations are similar to the notations defined in the previous sections except the relevant variable is $N(0, \Sigma_{G+1,n})$ only. That is, f_q is the marginal density function of the q^{th} random variable of $N(0,\Sigma_{G+1,n})$, $f_{(q)}$ is the joint conditional density function of the remaining n-G-1 variables given that the q^{th} variable of $N(0, \Sigma_{G+1,n})$ is equal to $-\alpha_q X_t$.

Thus after adjusting the truncated means, we have

$$y_{it} = \alpha_i X_t + \Sigma^n_{q=G+1}\, \sigma_{iq} f_q(\alpha_q' X_t) F_{(q)}/P_t + \xi_{it}, \quad \forall t\varepsilon S_1$$
$$\forall i = 1, \ldots, n$$

with $E(\xi_{it}|t\varepsilon S_1) = 0$

Using the instrumental variables method, we have the initial consistent estimates of $\alpha_{G+1}, \ldots, \alpha_n$ and $\Sigma_{G+1,n}$. After substituting these estimates into the expressions $f_{(q)}(\alpha_q' X_t) F_{(q)}/P_t$, $q = G+1, \ldots, n$, we can regress y_{it} on the X_t and $f_{(q)}(\alpha_q' X_t) F_{(q)}/P_t$ to have the estimates $\alpha_i, \sigma_{iG+1}, \ldots, \sigma_{in}$ for each equation i, $i = G+1, \ldots, n$.

With the above procedure, we have estimated all the unknown parameters except

$$\Sigma_G = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1G} \\ \vdots & & & \vdots \\ \sigma_{G1} & \sigma_{G2} & \cdots & \sigma_G^2 \end{bmatrix}$$

Since

$$E(\nu_{it}^2 | t\varepsilon S_1) = \int_{-\alpha_n X_t}^{\infty} \cdots \int_{-\alpha_{G+1} X_t}^{\infty} \left[\int_{-\infty}^{\infty} \nu_i^2 \frac{f(\nu_i, \nu_{G+1}, \ldots, \nu_n)}{P_t} d\nu_i \right] d\nu$$

$$= \int_{-\alpha_n X_t}^{\infty} \cdots \int_{-\alpha_{G+1} X_t}^{\infty} \left[\int_{-\infty}^{\infty} \nu_i^2 f(\nu_i | \nu_{G+1}, \ldots, \nu_n) d\nu_i \right] \frac{f(\nu_{G+1}, \ldots \nu_n)}{P_t} d\nu$$

$$= \int_{-\alpha_n X_t}^{\infty} \cdots \int_{-\alpha_{G+1} X_t}^{\infty} \left\{ \sigma_i^2 - \left[\sigma_{iG+1}, \ldots, \sigma_{in} \right] \Sigma_{G+1,n}^{-1} \begin{bmatrix} \sigma_{iG+1} \\ \vdots \\ \sigma_{in} \end{bmatrix} \right.$$

$$\left. + \left(\left[\sigma_{iG+1}, \ldots, \sigma_{in} \right] \Sigma_{G+1,n}^{-1} \begin{bmatrix} \nu_{G+1} \\ \vdots \\ \nu_n \end{bmatrix} \right)^2 \right\} \frac{f(\nu_{G+1}, \ldots, \nu_n)}{P_t} \, d\nu$$

where $d\nu$ denotes $d\nu_{G+1} \cdots \cdots d\nu_n$

$$= \sigma_i^2 - \left[\sigma_{iG+1}, \ldots, \sigma_{in} \right] \Sigma_{G+1,n}^{-1} \begin{bmatrix} \sigma_{iG+1} \\ \vdots \\ \sigma_{in} \end{bmatrix}$$

$$+ \left[\sigma_{iG+1}, \ldots, \sigma_{in} \right] \Sigma_{G+1,n}^{-1} \begin{bmatrix} E(\nu_{G+1}^2 | t\varepsilon S_1), \ldots, E(\nu_{G+1}\nu_n | t\varepsilon S_1) \\ \vdots \qquad \qquad \vdots \\ E(\nu_n \nu_{G+1} | t\varepsilon S_1), \ldots, E(\nu_n^2 | t\varepsilon S_1) \end{bmatrix}$$

$$\Sigma_{G+1,n}^{-1} \begin{bmatrix} \sigma_{iG+1} \\ \vdots \\ \sigma_{in} \end{bmatrix}, \qquad \forall\, i = 1, \ldots, G$$

Also,

$$E(\nu_{it}\nu_{jt} | t\varepsilon S_1) = \sigma_{ij} - \left[\sigma_{iG+1}, \ldots, \sigma_{in} \right] \Sigma_{G+1,n}^{-1} \begin{bmatrix} \sigma_{jG+1} \\ \vdots \\ \sigma_{jn} \end{bmatrix} .$$

$$+ \left[\sigma_{iG+1},\ldots,\sigma_{in}\right] \Sigma^{-1}_{G+1,n} \begin{bmatrix} E(\nu^2_{G+1}|t\varepsilon S_1), & \cdots & E(\nu_{G+1}\nu_n|t\varepsilon S_1) \\ \vdots & & \vdots \\ E(\nu_n\nu_{G+1}|t\varepsilon S_1), & \cdots & E(\nu_n{}^2|t\varepsilon S_1) \end{bmatrix}$$

$$\Sigma^{-1}_{G+1,n} \begin{bmatrix} \sigma_{jG+1} \\ \vdots \\ \sigma_{jn} \end{bmatrix} \qquad \forall\ i,j=1,\ldots,G$$

Denote,

$$\hat{V}_{it} = y_{it} - \hat{\alpha}_i X_t, \qquad \forall\ i = 1,\ldots,G$$

That is, $\hat{V}_{it}$ is the estimated residuals.

Thus we can estimate $\sigma_i{}^2$ and σ_{ij} consistently by using the estimated moments.

$$\hat{\sigma}_i{}^2 = \frac{1}{\#(S_1)} \sum_{t\varepsilon S_1} \hat{V}^2_{it} + \left[\hat{\sigma}_{iG+1},\ldots,\hat{\sigma}_{in}\right] \tilde{\Sigma}^{-1}_{G+1,n} \begin{bmatrix} \hat{\sigma}_{iG+1} \\ \vdots \\ \hat{\sigma}_{in} \end{bmatrix}$$

$$- \left[\hat{\sigma}_{iG+1},\ldots,\hat{\sigma}_{in}\right] \tilde{\Sigma}^{-1}_{G+1,n} \begin{bmatrix} \tilde{E}(\nu^2_{G+1}|t\varepsilon S_1), & \cdots & \tilde{E}(\nu_{G+1}\nu_n|t\varepsilon S_1) \\ \vdots & & \vdots \\ \tilde{E}(\nu_n\nu_{G+1}|t\varepsilon S_1), & \cdots & \tilde{E}(\nu_n{}^2|t\varepsilon S_1) \end{bmatrix}$$

$$\tilde{\Sigma}^{-1}_{G+1,n} \begin{bmatrix} \hat{\sigma}_{iG+1} \\ \vdots \\ \hat{\sigma}_{in} \end{bmatrix}$$

and

$$\hat{\sigma}_{ij} = \frac{1}{\#(S_1)} \sum_{t\varepsilon S_1} \hat{V}_{it}\hat{V}_{jt} + \left[\hat{\sigma}_{iG+1},\dots,\hat{\sigma}_{in}\right] \tilde{\Sigma}^{-1}_{G+1,n} \begin{bmatrix} \hat{\sigma}_{jG+1} \\ \vdots \\ \hat{\sigma}_{jn} \end{bmatrix}$$

$$-\left[\hat{\sigma}_{iG+1},\dots,\hat{\sigma}_{in}\right] \tilde{\Sigma}^{-1}_{G+1,n} \begin{bmatrix} \tilde{E}(\nu^2_{G+1}|t\varepsilon S_1), & \tilde{E}(\nu_{G+1}\nu_n|t\varepsilon S_1) \\ \tilde{E}(\nu_n\nu_{G+1}|t\varepsilon S_1), & \tilde{E}(\nu_n^2|t\varepsilon S_1 \end{bmatrix}.$$

$$\tilde{\Sigma}^{-1}_{G+1,n} \begin{bmatrix} \hat{\sigma}_{jG+1} \\ \vdots \\ \hat{\sigma}_{jn} \end{bmatrix} \qquad \forall\ i,j=1.,,,,.G$$

Thus all the unknown parameters Ω can be estimated consistently and the two step maximum likelihood procedure can be applied.

4. Simultaneous Equations with Some Dependent Variables Truncated Normal

In this section, we extend Amemiya's system to the case where only some of the dependent variables are truncated but not all of them.

The simultaneous system we will consider has the following specifications

$$\Gamma_{11}y_t^{(1)} + \Gamma_{12}y_t^{(2)} = B_1X_t + \mu_t^{(1)}$$

$$\Gamma_{21}y_t^{(1)} + \Gamma_{22}y_t^{(2)} \geq B_2X_t + \mu_t^{(2)}$$

$$y_t^{(2)} \geq 0$$

and

$(\Gamma_{21}y_t^{(1)} + \Gamma_{22}y_t^{(2)})i = (B_2X_t + \mu_t^{(2)})i$ whenever the i^{th} component of $y_t^{(2)}$ is greater than 0, where $y_t^{(1)} = (y_{1t}, y_{2t}, \ldots, y_{Gt})'$, $y_t^{(2)} = (y_{G+1t}, \ldots, y_{nt})'$ and $(\mu_{1t}, \mu_{2t}) \sim N(0, \Omega)$ which is serially independent.

As in Amemiya's original system, given any X_t, μ_{1t} and μ_{2t}, there may be no random value y_t satisfying the system. As an illustration, it is easy to show these cases happen. The following figures give an example

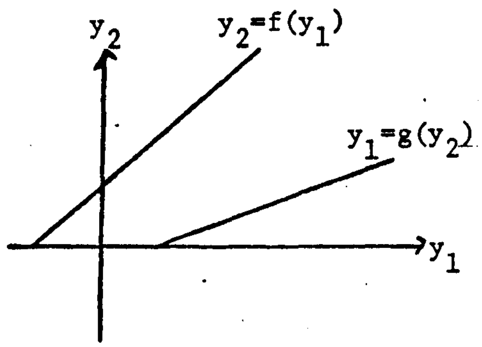

Figure 1: y_2 is truncated; there is no solution.

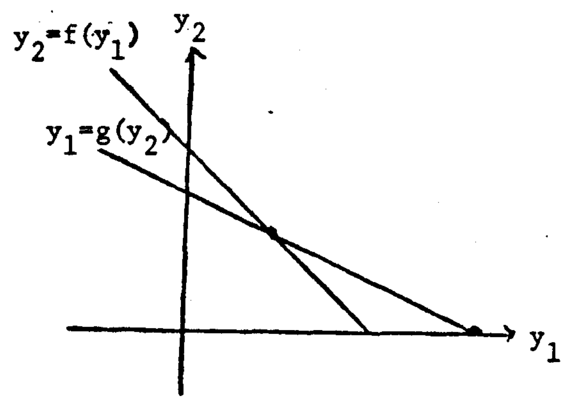

Figure 2: y_2 is truncated; there are multiple solutions.

However, under some general conditions, the system is well defined and there is unique solution for the system. Suppose Γ_{11} is non singular, then

$$y_t^{(1)} = - \Gamma_{11}^{-1}\Gamma_{12}y_t^{(2)} + \Gamma_{11}^{-1}B_1X_t + \Gamma_{11}^{-1}\mu_t^{(1)}$$

which implies in turn that

$$(\Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12})\, y_t^{(2)} \geqq -\Gamma_{21}\Gamma_{11}^{-1}B_1X_t + B_2X_t + \mu_t^{(2)} - \Gamma_{21}\Gamma_{11}^{-1}\mu_t^{(1)} \quad (*)$$

It follows from Amemiya's finding that if $\Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}$ is a P-matrix, there will be an unique solution $y_t^{(2)} \geqq 0$ satisfying (*). Hence the whole system will be well-defined if Γ_{11} is non singular and $\Gamma_{22}-\Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}$ is a P-matrix.

Now it remains to consider the estimation procedure for this model.

Define

$$y_t^{(1)*} = y_t^{(1)},\; y_t^{(2)*} = y_t^{(2)} \quad \text{if} \quad y_t^{(2)} > 0,$$

$y_t^{(1)*} = y_t^{(2)*} = 0$, otherwise and $y_t^* = (y_t^{(1)*}, y_t^{(2)*})$ is always observable. With these notations, we have the following reduced form defined on the nontruncated observations.

$$\begin{bmatrix} y_t^{(1)*} \\ y_t^{(2)*} \end{bmatrix} = \pi X_t + \nu_t \qquad \text{if } y_t^{(2)} > 0$$

where $\pi = \Gamma^{-1}B$, $\nu_t = \Gamma^{-1}\begin{bmatrix} \mu_t^{(1)} \\ \mu_t^{(2)} \end{bmatrix}$ and $\mathrm{var}(\nu_t) = \Sigma = \Gamma^{-1}\Omega\,\Gamma^{-1}$.

The reduced form parameters π and Σ can then be estimated consistently by the procedures in the last section. With the rank conditions satisfied by the simultaneous equations system, the Γ, B and Ω can be identified via the reduced form parameters. If the system is exactly identifiable, then the indirect least square is applicable. Otherwise, we may follow the procedure suggested in section 2. Thus the system can be estimated consistently. This estimation procedure depends on the usual rank identification conditions for the whole system. The sample observations of $y_t^{(1)}$ when $y_t^{(2)} \leq 0$ have not been used in the consistent estimation procedure even though they are utilized in the two step maximum likelihood procedure. These corner solutions are in fact very useful. In the following sections, we will show that the rank condition for the whole system is only a sufficient condition but not a necessary condition. Certain parts of the system can be identified under much weaker conditions. Also we will show how the corner observations can be utilized to give consistent estimates under those cases. To simplify the expression, we will consider a two equation model first and then the general models.

5.1 Two Equations Model with only one truncated variable.

In the previous sections, we consider the general simultaneous equations system with an arbitrary number of dependent variables truncated. In this section, we consider the special case of this model. The two equations system is much less complicated than the general system. Due to its relatively simpler structure, we can consider more detail about its structure, particularly the corner observations of the system.

More specifically, the two equations system has the following equivalent specification.

$$y_{1t} = \gamma_1 y_{2t} + \delta_1' X_t + \varepsilon_{1t}$$

$$y_{2t} = \gamma_2 y_{1t}^* + \delta_2' X_t + \varepsilon_{2t}$$

where

$$y_{1t}^* = \begin{cases} y_{1t} & \text{when } y_{1t} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

In this system, $(\varepsilon_{1t}, \varepsilon_{2t}) \sim N(0, \Omega)$; we can always observe y_{2t} and y_{1t}^* but not y_{1t}. In fact y_{1t} is an underlying index.

The P-matrix condition for the existence of the system is $1-\gamma_1\gamma_2 > 0$ in this case. In the general system, we have used only the interior solutions of the system and hence the rank conditions are required to give the identification. In the two equations model, it can be shown that the corner solutions give useful information about the system. The corner solutions can identify one of the equations.

For the two equations system above, there are two reduced forms corresponding to the regime where the truncation is effective or not. Denote $S_1 = \{t | y_{1t} > 0\}$ and $S_2 = \{t | y_{1t} \leq 0.\}$. The reduced form for the observations where the truncation is not effective is,

$$\left.\begin{aligned} y_{1t} &= \frac{1}{1-\gamma_1\gamma_2}(\delta_1' + \gamma_1\delta_2')X_t + \frac{1}{1-\gamma_1\gamma_2}(\varepsilon_{1t} + \gamma_1\varepsilon_{2t}) \\ y_{2t} &= \frac{1}{1-\gamma_1\gamma_2}(\delta_2' + \gamma_2\delta_1')X_t + \frac{1}{1-\gamma_1\gamma_2}(\gamma_2\varepsilon_{1t} + \varepsilon_{2t}) \end{aligned}\right\} \quad \forall t \varepsilon S_1$$

and it is,

$$\left.\begin{aligned} y_{1t} &= (\delta_1' + \gamma_1\delta_2')X_t + (\varepsilon_{1t} + \gamma_1\varepsilon_{2t}) \\ y_{2t} &= \delta_2' X_t + \varepsilon_{2t} \end{aligned}\right\} \quad \forall t \varepsilon S_2$$

where the truncation is effective.

In these forms, we can utilize the corner observations to estimate δ_2' consistently. It can be estimated as follows.

Consider the truncated mean $E(\varepsilon_{2t} | y_{1t} < 0)$. Under the condition that $1 - \gamma_1\gamma_2 > 0$ (which is the existence condition for the model), $y_{1t} < 0$ is equivalent to

$$(\delta_1' + \gamma_1\delta_2')X_t + (\varepsilon_{1t} + \gamma_1\varepsilon_{2t}) < 0.$$

Thus

$$E(\varepsilon_{2t} | y_{1t} < 0) = E(\varepsilon_{2t} | (\delta_1' + \gamma_1\delta_2')X_t + (\varepsilon_{1t} + \gamma_1\varepsilon_{2t}) < 0)$$

Define $\varepsilon_t = \frac{1}{\sigma^*}(\varepsilon_{1t} + \gamma_1\varepsilon_{2t})$ with $\sigma^{*2} = \mathrm{var}(\varepsilon_{1t} + \gamma_1\varepsilon_{2t})$.

$$\begin{aligned} E(\varepsilon_{2t} | y_{1t} < 0) &= E(\varepsilon_{2t} | (\frac{\delta_1' + \gamma_1\delta_2'}{\sigma^*}) X_t + \varepsilon_t < 0) \\ &= E(\varepsilon_{2t} | -(\frac{\delta_1' + \gamma_1\delta_2'}{\sigma^*}) X_t > \varepsilon_t) \\ &= -\sigma_{2\varepsilon}\, f(-\frac{\delta_1' + \gamma_1\delta_2'}{\sigma^*} X_t) / F(-\frac{\delta_1' + \gamma_1\delta_2'}{\sigma^*} X_t) \end{aligned}$$

where $\sigma_{2\varepsilon} = \text{cov}(\varepsilon_{2t}, \varepsilon_t)$; f, F are the standard normal density function and its distribution function respectively.

Hence

$$y_{2t} = \delta_2' X_t - \sigma_{2\varepsilon} f\left(-\frac{\delta_1' + \gamma_1 \delta_2'}{\sigma^*} X_t\right) / F\left(-\frac{\delta_1' + \gamma_1 \delta_2'}{\sigma^*} X_t\right) + \xi_t, \ \forall t \varepsilon S_2$$

where $E(\xi_t | t \varepsilon S_2) = 0$. The parameters $\frac{\delta_1' + \gamma_1 \delta_2'}{\sigma^*}$ can be estimated consistently by many procedures. One of the possible procedures is to define an indicator,

$$I_t = 1 \ \text{iff} \ y_{1t} > 0$$

$$I_t = 0 \ \text{iff} \ y_{1t} \leq 0,$$

and a probit model $I_t = F\left(-\frac{\delta_1' + \gamma_1 \delta_2'}{\sigma^*} X_t\right) + \eta_t$. The likelihood function of the probit model is strictly concave, so it can be estimated easily by numerical algorithms which guarantee convergence. With the parameter $\pi' = \frac{\delta_1' + \gamma_1 \delta_2'}{\sigma^*}$ estimated, say $\hat{\pi}'$, we can estimate δ_2' and $\sigma_{2\varepsilon}$ consistently by the ordinary least square to the equation,

$$y_{2t} = \delta_2' X_t - \sigma_{2\varepsilon} f(-\hat{\pi}' X_t) / F(-\hat{\pi}' X_t) + \tilde{\xi}_t$$

Another possibility is to use the Instrumental variable procedure by Amemiya to get the consistent estimate of π' to begin with.

Thus the parameters δ_2' can be estimated consistently and it must be identifiable. With the δ_2' identifiable, we will have more information about the identification of the equations.

With the observations where truncation is not effective, the parameters,

$$\pi_1' = \frac{1}{1 - \gamma_1 \gamma_2} (\delta_1' + \gamma_1 \delta_2')$$

$$\pi_2' = \frac{1}{1-\gamma_1\gamma_2}(\delta_2' + \gamma_2\delta_1')$$

can be estimated consistently by the instrumental variables procedure and hence they must be identifiable.

Rewrite these as,

$$\begin{pmatrix} 1 & -\gamma_1 \\ -\gamma_2 & 1 \end{pmatrix} \begin{pmatrix} \pi_1' \\ \pi_2' \end{pmatrix} = \begin{pmatrix} \delta_1' \\ \delta_2' \end{pmatrix}$$

which implies

$$(1, -\gamma_1) \begin{pmatrix} \pi_1' \\ \pi_2' \end{pmatrix} = \delta_1' \text{ and } (-\gamma_2, 1) \begin{pmatrix} \pi_1' \\ \pi_2' \end{pmatrix} = \delta_2'.$$

With the reduced form parameters π_1, π_2 and δ_2 identifiable, γ_2 can be identified under general conditions. γ_2 is identified if and only if π_1' has rank 1. If $\pi_1' = 0$, y_{1t} consists of disturbances only and is not explained by any observed factors. In empirical applications, this will not be the case in general and hence the y_{2t} equation is always identifiable. In particular, when $\gamma_1 = 0$, $\delta_1' \neq 0$, it is a recursive system and the second equation y_{2t} is identified. From this special case, it is clear that the identification is different from the usual simultaneous equation system without truncations.

On the other hand, it is not obvious that we can identify γ_1, δ_1 without imposing the rank condition on (π_1, π_2) or equivalently, without imposing the zero restrictions condition on the equation

$$y_{1t} = \gamma_1 y_{2t} + \delta_1' X_t + \varepsilon_{1t}.$$

In fact, it can be shown that this equation is unidentifiable without any further restriction. This can be shown easily by taking linear combinations of the two equations. For example, we can add the two equations and we have

$$y_{1t} = (\gamma_1 - 1)y_{2t} + \gamma_2 y_{1t}^* + (\delta_1' + \delta_2')X_t + (\varepsilon_{1t} + \varepsilon_{2t})$$

$$y_{2t} = \gamma_2 y_{1t}^* + \delta_2' X_t + \varepsilon_{2t}$$

Thus we now have two systems.

System 1, $y_{1t} = \gamma_1 y_{2t} + \delta_1 ' X_t + \varepsilon_{1t}$

$$y_{2t} = \gamma_2 y^*_{1t} + \delta_2 ' X_t + \varepsilon_{2t}$$

and

System 2, $y_{1t} = (\gamma_1 - 1) y_{2t} + \gamma_2 y_{1t}{}^* + (\delta_1 ' + \delta_2 ')X_t + (\varepsilon_{1t} + \varepsilon_{2t})$

$$y_{2t} = \gamma_2 y^*{}_{1t} + \delta_2 ' X_t + \varepsilon_{2t}$$

It is clear that those observations without truncation satisfy the above system. Those observations $(0, y_{2t})$ with truncation which satisfy the first system, equivalently satisfy the following equality and inequality;

$$0 \geq \gamma_1 y_{2t} + \delta_1 ' X_t + \varepsilon_{1t}$$

$$y_{2t} = \delta_2 ' X_t + \varepsilon_{2t}.$$

By addition, the observations also satisfy

$$0 \geq (\gamma_1 - 1) y_{2t} + (\delta_1 ' + \delta_2 ')X_t + (\varepsilon_{1t} + \varepsilon_{2t})$$

$$y_{2t} = \delta_2 ' X_t + \varepsilon_{2t}.$$

Thus those solutions which satisfy the first system satisfy the second system also. Similarly, observations which satisfy system 2 also satisfy system 1. Thus the two systems cannot be distinguished by the observations. Hence we can conclude that the equation

$$y_{1t} = \gamma_1 y_{2t} + \delta_1 ' X_t + \varepsilon_{1t}$$

is unidentifiable with no further restrictions. To achieve the identification of this equation, at least one exogeneous variable must not appear in this equation. With the parameters identifiable, we can discuss estimation procedures.

If the zero order restriction conditions are satisfied for both equations, we can estimate the coefficients by the procedure for the general system as pointed out in the last section.

However, if only the first equation satisfies the first order restrictions, we have to search other possible procedure to estimate the parameters. In that case, we have to utilize the corner solutions to get the estimates. The two equations system implies a switching simultaneous equations system,

$$\text{regime 1,}\quad y^*_{1t} = \gamma_1 y_{2t} + \delta_1' X_t + \varepsilon_{1t}$$

$$y_{2t} = \gamma_2 y^*_{1t} + \delta_2' X_t + \varepsilon_{2t} \qquad \text{iff} \quad \pi_1' X_t + \mu_{1t} > 0$$

$$\text{regime 2,}\quad y^*_{1t} = 0$$

$$y_{2t} = \delta_2' X_t + \varepsilon_{2t} \qquad \text{iff} \quad \pi_1' X_t + \mu_{1t} \leq 0.$$

where $\mu_{1t} = \frac{1}{1-\gamma_1\gamma_2}(\varepsilon_{1t} + \gamma_1\varepsilon_{2t})$.

Hence, the estimation procedure in our model is related to the estimation procedure for the simultaneous switching equations systems in (8). The reduced form of the simultaneous switching regression model is

$$y^*_{1t} = \pi_1' X_t + \mu_{1t}$$

$$y_{2t} = \pi_2' X_t + \mu_{2t} \qquad \text{iff} \quad \pi_1' X_t + \mu_{1t} > 0$$

and

$$y^*_{1t} = 0$$

$$y_{2t} = \delta_2'X_t + \varepsilon_{2t} \qquad \text{iff} \quad \pi_1'X_t + \mu_{1t} \leq 0$$

where $\mu_{2t} = \frac{1}{1-\gamma_1\gamma_2}(\gamma_2\varepsilon_{1t} + \varepsilon_{2t})$.

This implies

$$y^*_{1t} = \pi_1'X_tF(\psi_t) + \sigma_{u1}f(\psi_t) + \zeta_t, \quad \text{where} \quad \sigma^2_{u1} = \text{var}(\mu_{1t}); \quad \psi_t = \pi_1'X_t/\sigma_{u1} ;$$

ζ_t is an error term which has zero mean but different variance for different observation t. Substitute this expression into the second structural equation of the original system; it is

$$y_{2t} = \gamma_2(\pi_1'X_tF(\psi_t) + \sigma_{u1}f(\psi_t)) + \delta_2'X_t + (\gamma_2\zeta_t + \varepsilon_{2t}), \quad \forall t$$

The parameter π_1' and π_2' and σ_{u1} can be estimated consistently by using the observations in S_1, say by $\hat{\pi}_1'$, $\hat{\pi}_2'$ and $\hat{\sigma}_{u1}$. Denote

$$\hat{y}^*_{1t} = \hat{\pi}_1'X_tF(\hat{\psi}_t) + \hat{\sigma}_{u1}f(\hat{\psi}_t) , \quad \text{where} \quad \hat{\psi}_t = \hat{\pi}_1'X_t/\hat{\sigma}_{u1} .$$

$y_{2t} = \gamma_2\hat{y}^*_{1t} + \delta_2'X_{2t} + \xi_t$ where ξ_t has zero mean asymptotically.
Thus γ_2 and δ_2' can be estimated consistently by regression y_{2t} on $\hat{y}^*_{1t}$ and X_{2t}.

As for the estimation of the first equation, consider only those observations in S_1. With the zero order restriction conditions satisfied, the parameters are identifiable and hence it is estimable. Substituting

$y_{2t} = \pi_2'X_t + \mu_{2t}$, $\forall t \varepsilon S_1$ into the first equation, it is

$$y_{1t} = \gamma_1(\pi_2'X_t) + \delta_1'X_{1t} + (\varepsilon_{1t} + \gamma_2\mu_{2t})$$

$$= \gamma_1(\pi_2'X_t) + \delta_1'X_{1t} + \mu_{1t} \qquad \forall t \varepsilon S_1$$

With the zero order identification condition, $\pi_2'X_t$ will not be a linear combination of X_{1t} and hence γ_1 and δ_1 is estimable.

Adjusting the truncation mean of μ_{1t} we have

$$y_{1t} = \gamma_1(\pi_2'X_t) + \delta_1'X_{1t} + \sigma_{u1}\frac{f(\psi_t)}{F(\psi_t)} + \xi_{1t}, \qquad \forall t \varepsilon S_1$$

with $E(\xi_{1t}|t\varepsilon S_1) = 0$. With the reduced form estimates, $\hat{\pi}_1'$, $\hat{\pi}_2'$, $\hat{\sigma}_{u1}$ and $\hat{\psi}_t$

$$y_{1t} = \gamma_1(\hat{\pi}_2'X_t) + \delta_1'X_{1t} + \sigma_{u1}\frac{f(\hat{\psi}_t)}{F(\hat{\psi}_t)} + \xi_{1t}$$

and γ_1, δ_1 can be estimated by OLS. The likelihood function is

$$L(\gamma_1,\gamma_2,\delta_1,\delta_2,\sigma_1^2,\sigma_{12},\sigma_2^2|X, y_1, y_2)$$

$$= \prod_{t\varepsilon S_1} (1-\gamma_1\gamma_2)\ f(y_{1t}-\gamma_1 y_{2t}-\delta_1'X_t,\ y_{2t}-\gamma_2 y_{1t}-\delta_2'X_t)\ \times$$

$$\prod_{t\varepsilon S_2} \int_{-\infty}^{-\gamma_1 y_{2t}-\delta_1'X_t} f(\varepsilon_{1t}, y_{2t}-\delta_2'X_t)d\varepsilon_{1t}.$$

The two step maximum likelihood procedure can then be applied with all these consistent estimates.

5.2 Identification and Estimation of the general model

In this section, we would like to generalize the analysis for the two equations model to the general one. The general model as specified in section 4 is

$$\Gamma_{11}Y_t^{(1)} + \Gamma_{12}Y_t^{(2)} = B_1X_t + u_t^{(1)} \tag{5.2.1}$$

$$\Gamma_{21}Y_t^{(1)} + \Gamma_{22}Y_t^{(2)} \geqq B_2X_t + u_t^{(2)}, \quad Y_t^{(2)} \geqq 0 \tag{5.2.2}$$

where $Y_t^{(1)}$ is a G dimensional vector and $Y_t^{(2)}$ is a n-G dimensional vector; Γ_{11} and Γ_{22} are normalized to have unitary diagonal elements. As pointed out in the two equations model, equations $Y_t^{(2)}$ will be identified only if the usual rank conditions are satisfied for each equation in $Y_t^{(2)}$. But these are not necessary for the equations $Y_t^{(1)}$.

When $Y_t^{(2)} > 0$,

$$(\Gamma_{22}-\Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12})Y_t^{(2)} = (B_2-\Gamma_{21}\Gamma_{11}^{-1}B_1)X_t + u_t^{(2)} - \Gamma_{21}\Gamma_{11}^{-1}u_t^{(1)} \tag{5.2.3}$$

Assume the rank conditions are satisfied for each equation in $Y_t^{(2)}$. Based on subsamples $Y_t^{(2)} > 0$, parameters $\Gamma_{22}-\Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}$, $B_2-\Gamma_{21}\Gamma_{11}^{-1}B_1$ and the variance covariance matrix of $u_t^{(2)}-\Gamma_{21}\Gamma_{11}^{-1}u_t^{(1)}$ can be estimated consistently by the procedures mentioned in section 2.

When $Y_t^{(2)} = 0$, or equivalently

$(B_2-\Gamma_{21}\Gamma_{11}^{-1}B_1)X_t + u_t^{(2)} - \Gamma_{21}\Gamma_{11}^{-1}u_t^{(1)} \leqq 0$, we have

$$\Gamma_{11}Y_t^{(1)} = B_1X_t + u_t^{(1)} \tag{5.2.4}$$

B_1 and Γ_{11} will be identified when the usual rank condition holds for this subsystem (5.2.4). This follows from similar arguments in section 5.1 and section 3. It remains to consider the parameters Γ_{12}. Based on the subsamples $Y_t^{(1)}$ and $Y_t^{(2)}$ when $Y_t^{(2)} > 0$, the reduced form parameters π are always identifiable and we have

$$\begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \pi = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

which implies

$$\Gamma_{12}\pi_2 = B_1-\Gamma_{11}\pi_1 \quad \text{with} \quad \pi = \left[\begin{matrix} \pi_1 \\ \pi_2 \end{matrix} \right] .$$

Since π_2, π_1, B_1, Γ_{11} are all identified, Γ_{12} will be identifiable if and only if rank (π_2) = n-G. This condition is equivalent to requiring that any linear combinations of elements in $Y_t^{(2)}$ not consistent of disturbances only. This condition will always be satisfied in empirical applications. In summary, we conclude that the whole system will be identified when the usual rank conditions in (5.2.2) and (5.2.4) hold. In particular, when (5.2.1) consists of only one equation, (5.2.1) will always be identified.

Now let us consider the estimation of this general model. Equations in (5.2.2) can be estimated as in sections 2 and 3. To estimate (5.2.1), we can use a two stage least squares procedure. Without loss of generalization, let us consider the first equation of (5.2.1) which is

$$Y_{1t} = -\gamma_{12}Y_{2t}-\cdots-\gamma_{1G}Y_{Gt} - \gamma_{1G+1}Y_{G+1t}-\cdots-\gamma_{1n}Y_{nt} + \beta_1'X_t + u_{1t}$$

where

$Y_t^{(1)} = (Y_{1t},\ldots,Y_{Gt})'$, $Y_t^{(2)} = (Y_{G+1t},\ldots,Y_{nt})$. With the expected values $m_{it} = E(Y_{it})$

$$Y_{1t} = -\gamma_{12}m_{2t} - \cdots - -\gamma_{1n}m_{nt} + \beta'X_t + w_{1t}$$

where $E(w_{1t}|X_t) = 0$. It is easy to see that m_{it} can be estimated by the parameters in the several reduced forms and the parameters in (5.2.3). With these estimated $\hat{m}_{it}$ least squares can be applied to

$$Y_{1t} = -\gamma_{12}\,\hat{m}_{2t} - \cdots\cdots - \gamma_{1n}\,\hat{m}_{nt} + \beta'X_t + \tilde{w}_{1t} \quad .$$

Similarly, all equations in (5.2.1) can be consistently estimated. The variance matrix Ω can be estimated as in sections 2 and 3. Hence all the parameters can be consistently estimated. With the two step maximum likelihood procedure, efficient estimates and their asymptotic variance can then be derived.

6. Some Empirical Examples:

To illustrate the potential usefulness in economic application, we point out that a disequilibrium market model studied by Suit [9] and Goldfeld and Quandt [5] is in this framework. This disequilibrium market model is a model of the watermelon market. Following the simplified notations of Goldfeld and Quandt, the basic equations of the model can be written as

$$q_t = b_1 Z_{1t} + b_2 + u_{1t}$$

$$X_t = b_3 p_t + b_4 q_t + b_5 Z_{2t} + b_6 + u_{2t}$$

$$p_t = b_7 Z_{3t} + b_8 Y_t + b_9 + u_{3t}$$

$$Y_t = \min(q_t, X_t)$$

where Z_{it}, $i = 1,\ldots,3$ are vectors of predetermined variables; Z_{it}, q_t, p_t and Y_t are observable but not X_t. The first equation describes the determination of the crop. The second equation is the harvest equation which states that the intended amount of harvest is a function of current price, the size of the crop itself and other factors. The third equation is a standard demand equation. The last equation says that the actual harvest is the minimum quantity of intended harvest and the size of the crop. Under certain circumstances, it may not be worthwhile to harvest the entire crop. On the other hand, it may be possible that the intended harvest exceeds the crop and in this case the actual harvest will equal the crop. The structure of this model is quite general and it can be applied to many other agricultural product markets. For a more detailed description of this model, one can refer to the original papers.

Let $Y_{1t} = q_t$, $Y_{2t} = p_t$, $Y_{3t} = q_t - X_t$ and $Y_{3t}^* = \begin{cases} Y_{3t} & \text{if } Y_{3t} > 0 \\ 0 & \text{if } Y_{3t} \leq 0 \end{cases}$

The model can be rewritten as

$$Y_{1t} = b_1 Z_{1t} + b_2 + u_{1t}$$

$$Y_{2t} = b_7 Z_{3t} - b_8 Y_{3t}^* + b_8 Y_{1t} + b_9 + u_{3t} \ ,$$

$$Y_{3t}^* \geq -b_3 Y_{2t} + (1-b_4) Y_{1t} - b_5 Z_{2t} - b_6 - u_{2t} \ , \quad Y_{3t}^* \geq 0 \ .$$

Thus this is a three equations model with only one truncated endogeneous variables. Goldfeld and Quandt estimated this model by the maximum likelihood procedure. As suggested above, one can simplify the iteration procedure to a two step maximum likelihood procedure with initial consistent estimates derived from instrumental procedures and modified least squares. Goldfeld and Quandt pointed out that $b_3 > 0$ and $b_8 < 0$ would give existence of the model. From our analysis above, the existence condition is

$$1 - b_3 b_8 > 0 \ .$$

Hence the conditions mentioned in Goldfeld and Quandt are relatively stronger.

The other examples are the models of the income maintenance experiment studied by Hausman and Wise [6,7]. In their models, only the truncated samples are observed, i.e., only the interior solutions are observed in the terminologies of our models. It is easy to see that our estimation procedure; the one in which only the interior solutions are utilized, can be used rather than the iteration procedures they employed.

7. Conclusions and Remarks:

In this paper, we consider some multivariate and simultaneous equation models with truncated dependent variables. These models extend Tobin's model [11] and Amemiya's model [2] to the general cases where only some of the dependent variables are truncated normal. The existence condition, identification and estimation problem are investigated. The existence condition is achieved under a certain P-matrix [4] condition on the coefficients of the system. The identification problems are different from the usual simultaneous equation system or Amemiya's model where all the endogeneous variables are truncated. We have shown that the corner solutions will identify some coefficients in the system. In particular, when there is only one dependent variable which is not truncated, this equation will always be identifiable. The usual rank condition for simultaneous equation models is only a sufficient condition but not a necessary one. In the simultaneous equation model with all dependent variables truncated, Amemiya suggested some instrumental variable procedures which are computationally simple and an indirect least squares method. Unfortunately, his suggestions on the asymptotic tests derived are not correct. Instead of indirect least square, we extend it to the two stage least squares method which is more convenient for handling the overidentified cases. The estimates are consistent but as in Amemiya's procedures, the exact asymptotic variances have complicated expressions. To get more tractable asymptotic properties, we suggest using a two step maximum likelihood procedure which requires only first order derivates and does not require repeated iterations. To extend this two step maximum likelihood procedure to our models, we suggest some modified least squares methods which are based partly on Amemiya's instrumental variables procedures to get consistent estimates. These consistent procedures also utilize the samples from the corner solutions. To demonstrate that our models are useful in economics, we point out that a disequilibrium market model on agricultural products considered by Suit [9] and Goldfeld and Quandt [5] is an example of our models. Other examples are the income experiment models of Hausman and Wise [6,7].

The estimation procedures proposed are not iterative methods so that they are computationally tractible. The only complications involved are a fixed number of numerical evaluations of multivariate normal probabilities. The computations of numerical multivariate normal probablilties were expected to be unattractive for dimensions greater than 2. But a recent investigation by Dutt [3], suggests that efficient computational formulas for computing normal probabilities of dimension up to 6. Monte Carlo studies by Warner [12] indicated that for the bivariate case, contrary to expectations, computations are not expensive. So we can expect that our proposed two step maximum likelihood procedures would be useful in practice.

REFERENCES

(1) Amemiya, T., "Regression Analysis When the Dependent Variable is Truncated Normal," Econometrica, 41, (1973), 997-1016.

(2) Amemiya, T., "Multivariate Regression and Simultaneous Equation Models when the Dependent Variables are Truncated Normal", Econometrica, 42, (1974), 999-1012.

(3) Dutt, J. E., "Numerical Aspects of Multivariate Normal Probabilities in Econometric Models", Annals of Economic and Social Measurement 5, (1976), 547-561.

(4) Gale, G. and H. Naikaido, "The Jacobian Matrix and the Global Univalence of Mappings", Mathematische Annalen, 159 (1965), 81-93.

(5) Goldfeld, S. M. and Richard Quandt, "Estimation in a Disequilibrium Model and the Value of Information", Journal of Econometrics 3 (1975), 325-348.

(6) Hausman, J. A. and D. A. Wise, "Social Experimentation, Truncated Distributions, and Efficient Estimation", forthcoming in Econometrica, 1977.

(7) Hausman, J. A. and D. A. Wise, "The Evaluation of Results from Truncated Samples: The New Jersey Income Maintenance Experiment", Annals of Economic and Social Measurement, Vol. 5, No. 4, 1976, 421-445.

(8) Lee, L. F., "Limited Information Estimation of Some Switching Regression Models," Discussion Paper 75-20, Department of Economics, University of Rochester, 1975.

(9) Suits, D., "An Econometric Model of the Watermelon Market", Journal of Farm Economics, 37, (1955), 237-251.

(10) Tallis, G. M., "The Moment Generating Function of the Truncated Multi-Normal Distributions," Journal of the Royal Statistical Society (Series B), 23 (1961), 223-229.

(11) Tobin, J., "Estimation of Relationships for Limited Dependent Variables", Econometrica, 26, (1958), 24-36.