# 12 Constraints on the Parameters in Simultaneous Tobit and Probit Models

Peter Schmidt

## 12.1 Introduction

This chapter considers simultaneous tobit and probit models. When the observed (truncated or dichotomous) endogenous variables appear as right-hand side (r.h.s.) variables in some or all equations, the parameters of the model must satisfy certain constraints. It turns out that the constraints for the tobit and probit cases are similar in form. However, in the tobit case the constraints are inequalities, while in the probit case they are equalities. In the probit case these constraints rule out some models that would seem apparently reasonable, at least at first glance, and impose a degree of recursivity on the allowable models.

## 12.2 Simultaneous Tobit Models

Throughout this section a superscript* will represent an unobservable variable; the same variable without the asterisk will represent its observable (truncated) counterpart. In a tobit context the relationship between unobservables and observables is of course

$$y = \begin{cases} y^* & \text{if } y^* > 0, \\ 0 & \text{if } y^* \leq 0, \end{cases}$$

for any variable $y$ subject to a tobit truncation.

The models to be considered here have the observed endogenous variables as explanatory variables. For example, a simple two-equation model with one truncation could be written as follows:

$$y_1^* = \gamma_1 y_2 + \boldsymbol{\beta}_1' \mathbf{X} + \varepsilon_1, \tag{12.1}$$

$$y_2 = \gamma_2 y_1 + \boldsymbol{\beta}_2' \mathbf{X} + \varepsilon_2, \tag{12.2}$$

$$y_1 = \begin{cases} y_1^* & \text{if } y_1^* > 0, \\ 0 & \text{if } y_1^* \leq 0. \end{cases} \tag{12.3}$$

(Here $\mathbf{X}$ is a vector of exogenous variables, and $\varepsilon_1$ and $\varepsilon_2$ are disturbances.) Note it is the observed $y_1$ (not the unobserved $y_1^*$) that appears on the r.h.s. of (12.2). Models of this general type have been considered by Amemiya (1974), Lee (1976), and Schmidt and Sickles (1978).

On the other hand, models in which the truncated versions of endogenous variables never appear as r.h.s. variables have been considered by Nelson and Olson (1978) and Amemiya (1979). None of the constraints of the type investigated here arise in such models.

The distinction between these two types of models is basically the distinction as to whether $y_1$ or $y_1^*$ should appear on the r.h.s. of (12.2). In my opinion this should depend on whether in a particular application it is $y_1$ or $y_1^*$ that has a meaningful economic interpretation.

Such issues aside, we now return to system (12.1) through (12.3). Let us add a subscript $t$ indicating observation. Now for $t$ such that $y_{1t}^* > 0$, $y_{1t}^* = y_{1t}$ and the reduced form for $y_{1t}^*$ (or $y_{1t}$) is

$$y_{1t}^* = \frac{1}{1 - \gamma_1\gamma_2}[(\boldsymbol{\beta}_1 + \gamma_1\boldsymbol{\beta}_2)'\mathbf{X}_t + (\varepsilon_{1t} + \gamma_1\varepsilon_{2t})]. \tag{12.4}$$

On the other hand, for $t$ such that $y_{1t}^* \leq 0$, $y_{1t} = 0$, and the reduced form for $y_{1t}^*$ is

$$y_{1t}^* = [(\boldsymbol{\beta}_1 + \gamma_1\boldsymbol{\beta}_2)'\mathbf{X}_t + (\varepsilon_{1t} + \gamma_1\varepsilon_{2t})]. \tag{12.5}$$

From (12.4) and (12.5) it follows that we must have $1 - \gamma_1\gamma_2 > 0$. This condition ensures that for any $\mathbf{X}_t$, $\varepsilon_{1t}$, $\varepsilon_{2t}$, the model produces one and only one $y_{1t}$. This condition is necessary for the internal consistency (perhaps "unique solvability" would be a better phrase) of the model.

To see that this condition is indeed necessary, assume the opposite—let $1 - \gamma_1\gamma_2 < 0$. Then if the r.h.s. of (12.5) is positive, the r.h.s. of (12.4) is negative, which implies two different $y_{1t}^*$ but no solution for $y_{1t}$. On the other hand, if the r.h.s. of (12.5) is negative, the r.h.s. of (12.4) is positive, which implies two different $y_{1t}^*$ and two solutions for $y_{1t}$. Neither difficulty arises if $1 - \gamma_1\gamma_2 > 0$. Also neither difficulty would arise if $y_1^*$ appeared in place of $y_1$ on the r.h.s. of (12.5). Then (12.4) would be the reduced form for every $t$.

## 12.3 All Endogenous Variables Truncated

As noted previously, the earliest models of this type are those of Amemiya (1974), who considers models in which each of the endogenous variables is subject to a tobit truncation. We can write his model as follows:

$$\mathbf{Y}^* = \mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B} + \boldsymbol{\varepsilon}. \tag{12.6}$$

Here $\mathbf{Y}^*$, $\mathbf{Y}$, and $\boldsymbol{\varepsilon}$ are of dimension $1 \times G$; $\mathbf{X}$ is $1 \times K$; $\boldsymbol{\Gamma}$ is $G \times G$; and $\mathbf{B}$ is $K \times G$. ($G$ is therefore the number of equations.) The relationship between $\mathbf{y}^*$ and $\mathbf{y}$ is

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0, \\ 0 & \text{if } y_i^* \leq 0, \end{cases} \tag{12.7}$$

$i = 1, \ldots, G$, where $y_i$ is the $i$th column of $\mathbf{Y}$, and $y_i^*$ is the $i$th column of $\mathbf{Y}^*$.

Amemiya (1974, p. 1006) shows (by reference to a theorem in linear programming) that the conditions for internal consistency of this model are the following:

CONDITION 12.1: All principal minors of $(\mathbf{I} - \boldsymbol{\Gamma})$ must be positive.

## 12.4 Some Endogenous Variables Truncated

We now consider the case in which some, but not necessarily all, endogenous variables are subject to a tobit truncation. This case has been considered by Lee (1976) and Sickles and Schmidt (1978).

The model is the same as (12.6) above, except that only the first $S$ endogenous variables are truncated. Namely, we retain equation (12.6) but replace (12.7) with the following:

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0, \\ 0 & \text{if } y_i^* \leq 0, \end{cases} \tag{12.8}$$

$i = 1, \ldots, S$, and

$$y_i = y_i^*, \tag{12.9}$$

$i = S + 1, \ldots, G$. Alternatively we can write

$$(\mathbf{Y}_1^*, \mathbf{Y}_2) = (\mathbf{Y}_1, \mathbf{Y}_2)\begin{bmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{bmatrix} + \mathbf{X}(\mathbf{B}_1, \mathbf{B}_2) + (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2). \tag{12.10}$$

Here $\mathbf{Y}_1^*$ represents the first $S$ untruncated variables, which are related to $\mathbf{Y}_1$ as in (12.8).

As shown by Lee (1976), $\mathbf{Y}_2$ can be substituted out of this expression, which yields

$$\begin{aligned} \mathbf{Y}_1^* = {} & \mathbf{Y}_1[\boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{12}(\mathbf{I} - \boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}] + \mathbf{X}[\mathbf{B}_1 + \mathbf{B}_2(\mathbf{I} - \boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}] \\ & + [\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2(\mathbf{I} - \boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}]. \end{aligned} \tag{12.11}$$

But this is a system form with all endogenous variables truncated, so we can simply invoke the results of the previous section. That is, the condition for internal consistency of the model given by (12.6), (12.8), and (12.9) is that all principal minors of $\mathbf{I} - [\mathbf{\Gamma}_{11} + \mathbf{\Gamma}_{12}(\mathbf{I} - \mathbf{\Gamma}_{22})^{-1}\mathbf{\Gamma}_{21}]$ be positive. This can in turn be rewritten somewhat. Define

$$\mathbf{A} \equiv \mathbf{I} - \mathbf{\Gamma}, \tag{12.12}$$

and partition it as $\mathbf{\Gamma}$ is partitioned in (12.10). (Thus $\mathbf{A}_{11} = \mathbf{I} - \mathbf{\Gamma}_{11}$, $\mathbf{A}_{22} = \mathbf{I} - \mathbf{\Gamma}_{22}$, $\mathbf{A}_{12} = -\mathbf{\Gamma}_{12}$, $\mathbf{A}_{21} = -\mathbf{\Gamma}_{21}$.) Then the condition just given can be expressed in the following way:

CONDITION 12.2: All principal minors of $(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})$ must be positive.

As noted, condition 12.2 and its derivation are due to Lee (1976). An equivalent condition, given as condition 12.3, is derived by Sickles and Schmidt (1978).

CONDITION 12.3: All principal minors of $A$ that involve at least the last $G-S$ rows and columns must have the same sign.

The Sickles-Schmidt derivation need not be repeated here. However, it is easy to show that conditions 12.2 and 12.3 are indeed equivalent. To see this, we use the fact that

$$|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}| = \frac{|\mathbf{A}|}{|\mathbf{A}_{22}|}.$$

Now let $\mathbf{A}^*$ be formed from $\mathbf{A}$ by dropping corresponding rows and columns not involved in $\mathbf{A}_{22}$ (none of the last $G-S$ rows and columns is dropped). We therefore have

$$\mathbf{A}^* = \begin{bmatrix} \mathbf{A}_{11}^* & \mathbf{A}_{12}^* \\ \mathbf{A}_{21}^* & \mathbf{A}_{22} \end{bmatrix}$$

and

$$|\mathbf{A}_{11}^* - \mathbf{A}_{12}^*\mathbf{A}_{22}^{-1}\mathbf{A}_{21}^*| = \frac{|\mathbf{A}^*|}{|\mathbf{A}_{22}|}. \tag{12.13}$$

The l.h.s. is just a principal minor of $(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})$, while $|\mathbf{A}^*|$ is a principal minor of $\mathbf{A}$ involving at least $\mathbf{A}_{22}$. Thus the condition that all principal minors of $(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})$ be positive is equivalent to the condition that all principal minors of $\mathbf{A}$ involving at least $\mathbf{A}_{22}$ have the same sign.

## 12.5 Both $Y$ and $Y^*$ as Explanatory Variables

In section 12.2 we discussed the distinction between models in which the truncated variables appeared as r.h.s. variables and models in which only untruncated variables appeared as r.h.s. variables. We can of course allow both possibilities, although no one appears to have proposed such models so far. Suppose therefore that analogously to (12.6) we write

$$\mathbf{Y}^* = \mathbf{Y}\boldsymbol{\Gamma} + \mathbf{Y}^*\boldsymbol{\Delta} + \mathbf{XB} + \boldsymbol{\varepsilon}. \tag{12.14}$$

Clearly we can solve for $\mathbf{Y}^*$ to obtain

$$\mathbf{Y}^* = \mathbf{Y}\boldsymbol{\Gamma}(\mathbf{I} - \boldsymbol{\Delta})^{-1} + \mathbf{XB}(\mathbf{I} - \boldsymbol{\Delta})^{-1} + \boldsymbol{\varepsilon}(\mathbf{I} - \boldsymbol{\Delta})^{-1}. \tag{12.15}$$

But this is exactly of the form of (12.6), and the ensuing discussion still holds, with $\boldsymbol{\Gamma}$ replaced by $\boldsymbol{\Gamma}(\mathbf{I} - \boldsymbol{\Delta})^{-1}$.

What this points out is that the presence of truncated endogenous variables as r.h.s. variables is responsible for the constraints on the parameters. Whether untruncated endogenous variables appear as r.h.s. variables does not matter.

## 12.6 Simultaneous Probit Models

Simultaneous probit models appear to be somewhat less well worked out than simultaneous tobit models. Basically there is only the work of Heckman (1978) to refer to.

The notation of this section will be similar to that of the last section. For example, $y^*$ will represent an unobservable variable, with observable counterpart

$$y = \begin{cases} 1 & \text{if } y^* > 0, \\ 0 & \text{if } y^* \leq 0. \end{cases}$$

Constraints on the parameters are implied whenever such dichotomous variables appear as r.h.s. variables. These constraints are very similar in form to those of the corresponding tobit model, except that they are equalities rather than inequalities.

As a simple example to illustrate the nature of the problem, consider the two-equation model

$$y_1^* = \gamma_1 y_2 + \boldsymbol{\beta}_1'\mathbf{X} + \varepsilon_1, \tag{12.16}$$

$$y_2 = \gamma_2 y_1 + \boldsymbol{\beta}_2'\mathbf{X} + \varepsilon_2, \tag{12.17}$$

$$y_1 = \begin{cases} 1 & \text{if } y_1^* > 0, \\ 0 & \text{if } y_1^* \leq 0. \end{cases} \tag{12.18}$$

(This is identical to system (12.1) through (12.3), except for the rule relating $y_1$ to $y_1^*$.) The solution for $y_1^*$ is

$$y_1^* = (\gamma_1\gamma_2)y_1 + (\boldsymbol{\beta}_1 + \gamma_1\boldsymbol{\beta}_2)'\mathbf{X} + (\varepsilon_1 + \gamma_1\varepsilon_2). \tag{12.19}$$

From this it is easy to see that we must require $\gamma_1\gamma_2 = 0$. As in the tobit case the constraint ensures a unique outcome $y_1$ for any value of $\mathbf{X}$, $\varepsilon_1$ and $\varepsilon_2$. We can see this by noting that

$$\begin{aligned} y_1 &= 0 \quad \text{if } (\varepsilon_1 + \gamma_1\varepsilon_2) \leq -(\boldsymbol{\beta}_1 + \gamma_1\boldsymbol{\beta}_2)'\mathbf{X}, \\ y_1 &= 1 \quad \text{if } (\varepsilon_1 + \gamma_1\varepsilon_2) > -(\boldsymbol{\beta}_1 + \gamma_1\boldsymbol{\beta}_2)'\mathbf{X} - \gamma_1\gamma_2. \end{aligned}$$

If and only if $\gamma_1\gamma_2 = 0$, one and only one of these outcome must occur. This conclusion has previously been pointed out, for the two-equation case just considered, by Maddala and Lee (1976).

Two points are worth stressing here, in anticipation of what will come. First, the condition $\gamma_1\gamma_2 = 0$ is similar to the tobit condition $\gamma_1\gamma_2 < 1$. Second, the condition $\gamma_1\gamma_2 = 0$ imposes recursivity.

## 12.7 All Endogenous Variables Truncated

We consider the model

$$\mathbf{Y}^* = \mathbf{Y}\boldsymbol{\Gamma} + \mathbf{X}\mathbf{B} + \boldsymbol{\varepsilon}, \tag{12.20}$$

$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0, \\ 0 & \text{if } y_i^* \leq 0, \end{cases} \tag{12.21}$$

$i = 1, 2, \ldots, G$. This is very simular to the tobit system (12.6) to (12.7), see section 12.3 for more detail on the notation.

The condition for the internal consistency of this model is essentially that it be recursive. This can be expressed more precisely by the following condition:

**CONDITION 12.4:** There must be no nonzero product (chain) of the form $\Gamma_{i_1 i_2} \Gamma_{i_2 i_3} \ldots \Gamma_{i_{r-1} i_r} \Gamma_{i_r i_1}$, where for any $r \leq G$, $\{i_1, \ldots, i_r\}$ is any set of nonrepeated integers chosen from $\{1, 2, \ldots, G\}$.

To see why this condition is necessary, suppose that it does not hold. Thus we have a nonzero chain of the form

$$\Gamma_{ij}\Gamma_{jk} \cdots \Gamma_{pq}\Gamma_{qr}\Gamma_{ri}, \tag{12.22}$$

where $\{i, j, k, \ldots, p, q, r\}$ is a set of integers as before. Thus we have the situation:

$$\begin{aligned}
y_i^* &= \Gamma_{ri} y_r + \mathbf{X}\boldsymbol{\beta}_i + \varepsilon_i, \\
y_r^* &= \Gamma_{qr} y_q + \mathbf{X}\boldsymbol{\beta}_q + \varepsilon_q, \\
y_q^* &= \Gamma_{pq} y_p + \mathbf{X}\boldsymbol{\beta}_p + \varepsilon_p, \\
&\vdots \\
y_k^* &= \Gamma_{jk} y_j + \mathbf{X}\boldsymbol{\beta}_k + \varepsilon_k, \\
y_j^* &= \Gamma_{ij} y_i + \mathbf{X}\boldsymbol{\beta}_j + \varepsilon_j,
\end{aligned}$$

where for simplicity we have suppressed other endogenous r.h.s. variables. Now note that $y_r$ depends on $y_r^*$, which depends on $y_q$, which depends on $y_q^*$, which depends on $y_p, \ldots$, which depends on $y_k^*$, which depends on $y_j$, which depends on $y_j^*$, which depends on $y_i$. Since $y_i^*$ depends on $y_r$, we now have $y_i^*$ depending on $y_i$. Explicitly we could write

$$y_i^* = \Gamma_{ri} f(y_i, \mathbf{X}, \varepsilon_j, \ldots, \varepsilon_r, \text{parameters}) + \mathbf{X}\boldsymbol{\beta}_i + \varepsilon_i,$$

where $f$ shows that $y_r$ depends on $y_i$ (among other things). We then note that

$$\begin{aligned}
y_i = 1 \quad &\text{if } \varepsilon_i > -\mathbf{X}\boldsymbol{\beta}_i - \Gamma_{ri} f(1, \mathbf{X}, \varepsilon_j, \ldots, \varepsilon_r, \text{parameters}), \\
y_i = 0 \quad &\text{if } \varepsilon_i \leq -\mathbf{X}\boldsymbol{\beta}_i - \Gamma_{ri} f(0, \mathbf{X}, \varepsilon_j, \ldots, \varepsilon_r, \text{parameters}).
\end{aligned}$$

If the chain in (12.22) is nonzero, we are not guaranteed that one and only one of these outcomes will occur.

Condition 12.4 is just a condition of recursivity. As such it is equivalent to the following:

**CONDITION 12.5:** It must be possible to reorder variables (and equations) in such a way that all elements of $\Gamma$ on or below the diagonal equal zero.

It is also possible to express this condition in other equivalent ways, which more closely resemble the statements for the tobit case. For example, condition 12.4 (and hence condition 12.5) can be shown to be equivalent to the following condition:

CONDITION 12.6: All principal minors of $\Gamma$ must equal zero.

It is clear that condition 12.4 implies condition 12.6. To show their equivalence, it is therefore necessary to show that condition 12.6 implies condition 12.4. This can be done by induction. Consider the case $G = 2$. The principal minors of $\Gamma$ are $\Gamma_{11}$, $\Gamma_{22}$, and $\Gamma_{11}\Gamma_{22} - \Gamma_{12}\Gamma_{21}$. The requirement that these equal zero clearly implies that the chains $\Gamma_{11}$, $\Gamma_{22}$, $\Gamma_{12}\Gamma_{21}$, and $\Gamma_{21}\Gamma_{12}$ all equal zero. Hence condition 12.6 implies condition 12.4 for $G = 2$. Now assume that this implication holds for arbitrary value $G - 1$. This immediately implies that all chains of length $G - 1$ or less equal zero. What remains to be shown then is that $|\Gamma| = 0$ and all chains of length $G - 1$ equal to zero (together) imply that all chains of length $G$ must equal zero. This we show by contradiction. Suppose there is a nonzero chain of length $G$. Since the ordering of variables is arbitrary, we may as well suppose that this is the chain

$$\Gamma_{12}\Gamma_{23}\Gamma_{34}\ldots\Gamma_{G1} \neq 0. \tag{12.23}$$

Consider the implications of this. Since all chains of lengths of length 2 are zero, we must have

$$\Gamma_{21} = \Gamma_{32} = \Gamma_{43} = \ldots = \Gamma_{G,G-1} = \Gamma_{1G} = 0. \tag{12.24}$$

Similarly, since all chains of length 3 are zero,

$$\Gamma_{31} = \Gamma_{42} = \ldots \Gamma_{G,G-2} = \Gamma_{1,G-1} = \Gamma_{2G} = 0. \tag{12.25}$$

Since all chains of length 4 are zero,

$$\Gamma_{41} = \Gamma_{52} = \ldots = \Gamma_{G,G-3} = \Gamma_{1,G-2} = \Gamma_{2,G-1} = \Gamma_{3G} = 0, \tag{12.26}$$

and so forth. The final implication is that since all chains of length $G - 1$ equal zero,

$$\Gamma_{G-1,1} = \Gamma_{G2} = \Gamma_{13} = \ldots = \Gamma_{G-2,G} = 0. \tag{12.27}$$

Now equations (12.24) through (12.27) each imply that $G$ off-diagonal elements of $\Gamma$ must equal zero. Furthermore the coefficients appearing in these equations are all distinct. Hence (12.24) through (12.27) together set to zero all $G(G - 2)$ off-diagonal elements of $\Gamma$ that do not appear in (12.23). Given that the diagonal elements of $\Gamma$ also equal zero, this means that the expression in (12.23) equals $|\Gamma|$. But this contradicts that fact that $|\Gamma| = 0$. This completes the proof that condition 12.6 implies condition 12.4, and hence that they are equivalent.

Furthermore by the same line of proof we can show that the following requirement is equivalent to conditions 12.4 through 12.6:

CONDITION 12.7: All principal minors of $(\mathbf{I}-\boldsymbol{\Gamma})$ must equal one.

The similarity between this condition and condition 12.1 for the tobit model (that all principal minors must be positive) is evident.

## 12.8 Some Endogenous Variables Truncated

We now consider the case in which some endogenous variables, but not necessarily all of them, are of the probit type. Suppose that the first $S$ variables are of the probit type. Then the model under consideration is as given in equation (12.20), but with (12.21) replaced by

$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0, \\ 0 & \text{if } y_i^* \leq 0, \end{cases} \tag{12.28}$$

$i = 1, \ldots, S$ and

$$y_i = y_i^*, \tag{12.29}$$

$i = S+1, \ldots, G$. Alternatively as in section 12.4 we can write

$$(\mathbf{Y}_1^*, \mathbf{Y}_2) = (\mathbf{Y}_1, \mathbf{Y}_2)\begin{bmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{bmatrix} + \mathbf{X}(\mathbf{B}_1, \mathbf{B}_2) + (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2). \tag{12.30}$$

Substituting $\mathbf{Y}_2$ out of (12.30), we obtain

$$\begin{aligned} \mathbf{Y}_1^* = {} & \mathbf{Y}_1[\boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{12}(\mathbf{I}-\boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}] + \mathbf{X}[\mathbf{B}_1 + \mathbf{B}_2(\mathbf{I}-\boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}] \\ & + [\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2(\mathbf{I}-\boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}], \end{aligned} \tag{12.31}$$

which is identical to (12.11) except for the relationship of $\mathbf{Y}_1$ to $\mathbf{Y}_1^*$. But this is now a system in which all endogenous variables are of the probit type, so the results of the last section apply. In particular we need to require that all principal minors of $[\boldsymbol{\Gamma}_{11} + \boldsymbol{\Gamma}_{12}(\mathbf{I}-\boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}]$ be zero, or that all principal minors of $[\mathbf{I}-\boldsymbol{\Gamma}_{11}-\boldsymbol{\Gamma}_{12}(\mathbf{I}-\boldsymbol{\Gamma}_{22})^{-1}\boldsymbol{\Gamma}_{21}]$ equal one, and so on. This can be rewritten, as in section 12.4, by defining $\mathbf{A} = \mathbf{I}-\boldsymbol{\Gamma}$ (as in equation 12.12), and partitioning it conformably to $\boldsymbol{\Gamma}$. In this way we obtain the following:

CONDITION 12.8: All principal minors of $(\mathbf{A}_{11} - \mathbf{A}_{12}\,\mathbf{A}_{22}^{-1}\,\mathbf{A}_{21})$ must equal one.

An equivalent condition is the following:

CONDITION 12.9: All principal minors of $\mathbf{A}$ which involve at least the last $G-S$ rows columns must be equal.

The proof of the equivalence of conditions 12.8 and 12.9 follows exactly the lines of the discussion following condition 12.3 in section 12.4, with a few obvious changes. As a result there is no point in presenting it again. However, the analogy between the tobit and probit cases is again worth noting.

## 12.9 Both $Y$ and $Y^*$ as Explanatory Variables

As in the tobit case (section 12.5) we can consider models with both $\mathbf{Y}$ and $\mathbf{Y}^*$ as r.h.s. variables. (Heckman 1978 has previously considered some such models.) The model is as given in (12.14), but with the relationship between $\mathbf{Y}$ and $\mathbf{Y}^*$ given by (12.28) and (12.29). We can solve for $\mathbf{Y}^*$ in terms of $\mathbf{Y}, \mathbf{X}$, and $\boldsymbol{\varepsilon}$ to obtain

$$\mathbf{Y}^* = \mathbf{Y}\boldsymbol{\Gamma}(\mathbf{I} - \boldsymbol{\Delta})^{-1} + \mathbf{X}\mathbf{B}(\mathbf{I} - \boldsymbol{\Delta})^{-1} + \boldsymbol{\varepsilon}(\mathbf{I} - \boldsymbol{\Delta})^{-1}, \tag{12.32}$$

which is the same as (12.15), of course. But this is now of the form discussed in section 12.4 and therefore that discussion now applies.

As an example, consider the model of Heckman (1978) "translated" to the present notation:

$$\begin{aligned} y_1^* &= \gamma_{11} y_1 + \gamma_{21} y_2 + \mathbf{X}\boldsymbol{\beta}_1 + \varepsilon_1, \\ y_2 &= \gamma_{12} y_1 + \delta y_1^* + \mathbf{X}\boldsymbol{\beta}_2 + \varepsilon_2, \\ y_1 &= \begin{cases} 1 & \text{if } y_1^* > 0, \\ 0 & \text{if } y_1^* \leq 0. \end{cases} \end{aligned} \tag{12.33}$$

In matrix form

$$(y_1^*, y_2) = (y_1, y_2)\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & 0 \end{bmatrix} + (y_1^*, y_2)\begin{bmatrix} 0 & \delta \\ 0 & 0 \end{bmatrix} + \mathbf{X}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) + (\varepsilon_1, \varepsilon_2).$$

It is then easy to calculate that

$$\mathbf{I} - \boldsymbol{\Gamma}(\mathbf{I} - \boldsymbol{\Delta})^{-1} = \begin{bmatrix} 1 - \gamma_{11} & -\delta\gamma_{11} - \gamma_{12} \\ -\gamma_{21} & 1 - \delta\gamma_{21} \end{bmatrix}.$$

By condition 12.9 we must have all principal minors of this matrix involving at least the last row and column equal:

$$1 - \delta\gamma_{21} = (1 - \gamma_{11})(1 - \delta\gamma_{21}) + \gamma_{21}(-\delta\gamma_{11} - \gamma_{12}),$$

or

$$\gamma_{11} + \gamma_{12}\gamma_{21} = 0. \tag{12.34}$$

As Heckman shows, with this restriction (12.33) becomes

$$\begin{aligned} y_1^* &= \gamma_{21}y_2^l + \mathbf{X}\boldsymbol{\beta}_1 + \varepsilon_1, \\ y_2^l &= \delta y_1^* + \mathbf{X}\boldsymbol{\beta}_2 + \varepsilon_2, \end{aligned} \tag{12.35}$$

where $y_1$ is related to $y_1^*$ as before, and $y_2 = y_2^l + \gamma_{12}y_1$.[1] This is a model of unusual structure: $\mathbf{X}$ and $\varepsilon$ determine $(y_1^*, y_2^l)$; $y_1^*$ then determines $y_1$; $(y_2^l, y_1)$ then determines $y_2$. One might describe this structure as "quasi-recursive," or some such phrase. Clearly, however, inclusion of both $y_1$ and $y_1^*$ as r.h.s. variables in (12.33) has allowed strict recursivity to be escaped.

Similar comments apply to similar versions of the model. Consider briefly the model like (12.33) but with two probit variables:

$$\begin{aligned} y_1^* &= \gamma_{11}y_1 + \gamma_{21}y_2 + \delta_1 y_2^* + \mathbf{X}\boldsymbol{\beta}_1 + \varepsilon_1, \\ y_2^* &= \gamma_{12}y_1 + \gamma_{22}y_2 + \delta_2 y_1^* + \mathbf{X}\boldsymbol{\beta}_2 + \varepsilon_2. \end{aligned} \tag{12.36}$$

We can then derive the constraints:

$$\begin{aligned} \gamma_{11} + \delta_1\gamma_{12} &= 0, \\ \gamma_{22} + \delta_2\gamma_{21} &= 0, \\ \gamma_{12}\gamma_{21} &= 0. \end{aligned} \tag{12.37}$$

There are a number of ways that these constraints might be satisfied. Clearly in all of them either $\gamma_{12}$ or $\gamma_{21}$ (or both) must equal zero. We may as well consider

$$\gamma_{21} = 0,$$

since due to the symmetry of the notation this is no different than considering $\gamma_{12} = 0$. From the second line of (12.37), this implies $\gamma_{22} = 0$. With $\gamma_{21} = \gamma_{22} = 0$, we have then eliminated $y_2$ as a r.h.s. variable in (12.36):

1. Note that $y_2^l$ is the value of $y_2$ when $y_1 = 0$. For example, if $y_1$ were a unionism dummy, and $y_2$ were a wage rate, then $y_2$ would be the nonunion wage rate. The first equation of (12.35) then states that the probability of unionism depends on the nonunion wage.

$$y_1^* = \gamma_{11}y_1 + \delta_1 y_2^* + \mathbf{X}\boldsymbol{\beta}_1 + \varepsilon_1,$$
$$y_2^* = \gamma_{21}y_1 + \delta_2 y_1^* + \mathbf{X}\boldsymbol{\beta}_2 + \varepsilon_2$$

(still subject to $\gamma_{11} + \delta_1\gamma_{12} = 0$). But this is essentially identical to (12.33). (It differs only in the replacement of $y_2$ in equation 12.34 by $y_2^*$ here.) We can, imposing $\gamma_{11} + \delta_1\gamma_{12} = 0$, write

$$\begin{aligned} y_1^* &= \delta_1 y_2^l + \mathbf{X}\boldsymbol{\beta}_1 + \varepsilon_1, \\ y_2^l &= \delta_2 y_1^* + \mathbf{X}\boldsymbol{\beta}_2 + \varepsilon_2, \end{aligned} \tag{12.38}$$

where now $y_2^* = y_2^l + \gamma_{12}y_1$. The structure of the model is again quasi-recursive: $(\mathbf{X}, \varepsilon) \rightarrow (y_1^*, y_2^l)$; $y_1^* \rightarrow y_1$; $(y_2^l, y_1) \rightarrow y_2^*$; $y_2^* \rightarrow y_2$.

Similar special cases are easily worked out as well. They all have in common some element of recursivity, which limits the type of structure the model can have. More work is needed to fully understand what sorts of structures are internally consistent.

## 12.10 Conclusions

This chapter has derived the constraints that arise in simultaneous tobit and probit models. The constraints are necessary so that a given set of exogenous variables and disturbances yield a unique solution for the endogenous variables. They are very similar for the two cases, with the same principal minors required to be positive in the tobit case, but required to equal one in the probit case.

However, these constraints are far more interesting in the probit case, since the equality constraints essentially remove certain variables from certain equations. The result is to make simultaneous equations models recursive, in one way or another. The implications of this appear not yet to be fully understood.

It should be noted that these constraints arise when the (truncated) observable variables ($y$'s) appear as r.h.s. variables, either in place of or in addition to the underlying unobservable variables ($y^*$'s). It is therefore reasonable to ask whether we might do better to avoid these problems by not using $y$'s as r.h.s. variables. At least in the probit case, my own feeling is that this is not reasonable, since it is often natural to think of the effect of some discrete variable, quite apart from the effect of an underlying index of that variable. Indeed the fact that we observe legislatures passing laws, schools granting diplomas, and so forth, suggests that the effects of a discrete $y$ are not always thought to be captured by its underlying $y^*$.

Another point worth making is that the necessity of the constraints discussed here may raise questions of whether we ought to even consider simultaneous tobit and probit models at all. It is by no means clear that the best way to model truncated and/or qualitative variables is to embed them (essentially by analogy) into the usual simultaneous equations model. At this point it is not clear what the alternatives would be. However, the difficulties one runs into, especially in the probit case, do indicate that the analogy to the usual simultaneous equations model is not completely straightforward.

Finally, it should be noted that the simultaneous tobit and probit models considered here can be viewed as special cases of a general nonlinear model with unobserved variables. In both cases we have appearing as a r.h.s. variable a function of an unobservable endogenous variable. Multimarket disequilibrium models (e.g., where observable quantity is the minimum of supply and demand) also fit this general category and necessitate similar constraints. A general treatment of this problem is an interesting topic for future work. For some recent work on such a general treatment see Gourieroux, Laffont, and Monfort (1978).

## References

Amemiya, T. 1974. Multivariate Regression and Simultaneous Equation Models when the Dependent Variables are Truncated Normal. *Econometrica*. 42; 999–1012.

Amemiya, T. 1979. The Estimation of a Simultaneous Equation Tobit Model. *International Economic Review* 20: 169–182.

Gourieroux, C., J. J. Laffont, and A. Monfort. 1978. Coherency Conditions in Simultaneous Linear Equation Models with Endogenous Switching Regimes. Ecole Polytechnique Working Paper A190-0478.

Heckman. J. J. 1978. Dummy Endogneous Variables in a Simultaneous Equation System. *Econometrica*. 46; 931–959.

Lee, L. F. 1976. Multivariate Regression and Simultaneous Equations Models with Some Dependent Variables Truncated. Unpublished manuscript.

Maddala, G. S., and L. F. Lee. 1976. Recursive Models with Qualitative Endogenous Variables. *Annals of Economic and Social Measurement*. 5; 525–545.

Nelson, F. S., and L. Olson. 1978. Specification and Estimation of a Simultaneous Equation Model with Limited Dependent Variables. *International Economic Review*. 19: 695–710.

Sickles, R. C., and P. Schmidt. 1978. Simultaneous Equation Models with Truncated Dependent Variables: A Simultaneous Tobit Model. *Journal of Economics and Business*. 31: 11–21.