379-76

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

THE USE OF PROBIT IN TOBIN'S LIMITED DEPENDENT VARIABLE MODEL WITH SPECIAL REFERENCE TO LABOUR SUPPLY STUDIES

UNIVERSITY OF WISCONSIN ~ MADISON

THE USE OF PROBIT IN TOBIN'S LIMITED DEPENDENT
VARIABLE MODEL WITH SPECIAL REFERENCE TO LABOUR
SUPPLY STUDIES

Lewis T. Evans
(Victoria University of Wellington)

November 1976

The author's interest in this subject was stimulated by a paper written by David Crawford and Steven Garber. This note has benefitted from the author's discussions with these persons and from comments made by Arthur S. Goldberger.

THE USE OF PROBIT IN TOBIN'S LIMITED DEPENDENT VARIABLE MODEL WITH SPECIAL REFERENCE TO LABOUR SUPPLY STUDIES*

Lewis T. Evans
(Victoria University of Wellington)

1. Introduction

The integration of individuals' decisions concerning labour force participation and hours of work is a problem frequently encountered in empirical analyses of labour supply. It is well known that Tobin's (1958) limited dependent variable model, Tobit, provides a solution to this difficulty. In addition, Probit can be used to estimate the coefficients of a function which expresses the probability of an individual's working. Crawford and Garber (1976) use Probit to analyse the labour force participation decision, and then Tobit to obtain estimates of the coefficients of the labour supply function. The purposes of this note are to point out the relationship between the Probit and Tobit coefficients and to evaluate some implications of this relationship.

2. Specification of the Model

In this model an unobserved variable, y*, is determined thus:

$$y^* = \underline{\chi}'\underline{\beta} + \varepsilon \qquad (1.1)$$

where $\varepsilon \approx N(0,\sigma^2)$ and independently of the (m x 1) vector $\underline{\chi}$. The observed

*The author's interest in this subject was stimulated by a paper written by David Crawford and Steven Garber. This note has benefitted from the author's discussions with these persons and from comments made by Arthur S. Goldberger.

y is given by

$$y = \begin{cases} 0 & \text{if } y^* \leq 0 \\ y^* & \text{if } y^* > 0. \end{cases} \tag{1.2}$$

Available data are assumed to consist of a random sample of T observations on y and $\underline{x}$.

Tobit

Tobin (1958) has demonstrated that the probability density of y conditional on $\underline{x}$ is

$$p(y) = \begin{cases} 0 & \text{for } y < 0 \\ F(-\underline{x}'\beta/\sigma) & \text{for } y = 0 \\ \frac{1}{\sigma} f\left\{(y-\underline{x}'\underline{\beta})/\sigma\right\} & \text{for } y > 0, \end{cases} \tag{2}$$

where f(·) and F(·) are the standard normal density and cumulative functions. If the observations are ordered so that for the first observation, T_1, y = 0, and for the remaining $T-T_1$ observations, y 0, the log likelihood function is

$$\ell_{T_0} = \sum_{t=1}^{T} \log F(-x_t'\beta/\sigma) - \left(\frac{T-T_1}{2}\right) \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=T_1+1}^{T} (y_t - \underline{x}_t'\beta)^2. \tag{3}$$

The maximum likelihood estimates of the vector of Tobit coefficients, $\underline{\beta}$, and thestandard error, σ, are chosen to maximize equation (3).

Probit

In order to specify the Probit model it is necessary to define a variable d from the observed y thus:

$$d = \begin{cases} 1 & \text{if } y > 0 \\ 0 & \text{if } y = 0. \end{cases}$$

The probability distribution of d conditional on $\underline{x}$ is given by Probit as

$$p(d) = \begin{cases} F(\underline{x}'\underline{\gamma}) & \text{for } d = 1 \\ 1-F(\underline{x}'\underline{\gamma}) = F(-\underline{x}'\underline{\gamma}) & \text{for } d = 0, \end{cases} \tag{4}$$

using the fact that $1-F(\cdot) = F(-\cdot)$. The maximum likelihood estimate of the vector of Probit coefficients, $\underline{\gamma}$, is chosen to maximize the log-likelihood function

$$\ell_{Pr} = \sum_{t=1}^{T} \log F(-\underline{x}_t'\underline{\ }) + \sum_{t=T_1+1}^{T} \log F(\underline{x}_t'\underline{\gamma}). \tag{5}$$

3. Relationship Between the Tobit and Probit Coefficients

It is apparent from the definition of d that $p(d=0) = p(y=0)$, and consequently

$$F(-\underline{x}'\underline{\gamma}) = F(-\underline{x}'\beta/\sigma), \tag{6.1}$$

from equations (4) and (2). In turn, equation (6.1) implies that

$$-\underline{x}'\underline{\gamma} = x'\beta/\sigma, \tag{6.2}$$

as $F(\cdot)$ is a 1 to 1 function. Since equation (6.2) must hold for all values of $\underline{x}$ it follows that

$$\underline{\gamma} = (1/\sigma)\underline{\beta}. \tag{7}$$

Equation (7) means that the Probit coefficients, γ_i $i=1, \ldots, m$, differ from the Tobit coefficients, β_i $i=1, \ldots, m$, by a scalar mutliple only, where the scalar is the standard error, σ.

4. Relative Efficiency of Tobit and Probit Estimates of γ.

Owing to the relationship expressed in equation (7) it is possible to obtain estimates of $\underline{\gamma}$ from Tobit as well as Probit. This section addresses the question: is the estimate of $\underline{\gamma}$ obtained via Tobit as efficient as that obtained from Probit?

From equations (3) and (7) the Tobit log-likelihood function can be written

$$\ell_{T_0} = \sum_{t=1}^{T_1} \log F(-\underline{x}_t'\underline{\gamma}) - \frac{(T-T_1)}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=T_1+1} (y_t/\sigma - \underline{x}_t'\underline{\gamma})^2, \qquad (8)$$

and used to obtain maximum likelihood estimates of and $\underline{\gamma}$. The relative efficiency of the Probit and Tobit estimates of $\underline{\gamma}$ can be ascertained from a comparison of the Probit and Tobit information matrices for $\underline{\gamma}$. Since differentiation is a linear operation, the difference between the hessians of the Probit and Tobit log-likelihood functions can be obtained by subtracting (8) from (5) and differentiating the remainder twice with respect to $\underline{\gamma}$. This procedure is implemented in the Appendix and it yields the matrix.

$$\frac{\partial^2 \ell_{Pr}}{\partial \underline{\gamma} \partial \underline{\gamma}'} - \frac{\partial^2 \ell_T}{\partial \underline{\gamma} \partial \underline{\gamma}'} = \sum_{t=T_1+1}^{T} \underline{x}_t \underline{x}_t' \left(1 - \underline{x}_t'\underline{\gamma} \frac{f(\underline{x}_t'\underline{\gamma})}{F(\underline{x}_t'\underline{\gamma})} - \frac{f(\underline{x}_t'\underline{\gamma})^2}{F(\underline{x}_t'\underline{\gamma})^2}\right). \qquad (9)$$

The t^{th} term in the summation of equation (9) consists of the product of a positive semi-definite matrix, $\underline{x}_t \underline{x}_t'$, and the variance

$$V(y_t/\sigma - \underline{x}_t'\underline{\gamma} \mid y_t/\sigma - \underline{x}_t'\underline{\gamma} > -\underline{x}_t'\underline{\gamma}) = 1 - \underline{x}_t'\underline{\gamma} \frac{f(\underline{x}_t'\underline{\gamma})}{F(\underline{x}_t'\underline{\gamma})} - \frac{f(\underline{x}_t'\underline{\gamma})^2}{F(\underline{x}_t'\underline{\gamma})^2}. \qquad (10)$$

Consequently, the right hand side of equation (9) is a positive semidefinite matrix and

$$\frac{\partial^2 \ell_{Pr}}{\partial \underline{\gamma} \partial \underline{\gamma}'} \geq \frac{\partial^2 \ell_{T0}}{\partial \underline{\gamma} \partial \underline{\gamma}'}. \tag{11}$$

It is apparent from equation (11)[1] that

$$-\left[\frac{\partial^2 \ell_{Pr}}{\partial \underline{\gamma} \partial \underline{\gamma}'}\right]^{-1} \geq -\left[\frac{\partial^2 \ell_{T0}}{\partial \underline{\gamma} \partial \underline{\gamma}'}\right]^{-1},$$

which means that the Probit information matrix for $\underline{\gamma}$ exceeds that of Tobit by a positive semidefinite matrix. It is concluded that Tobit yields a more efficient estimate of the vector of Probit coefficients, $\underline{\gamma}$, than does Probit.

5. Comment

An implication of the relationship between the Probit and Tobit coefficients expressed in equation (7), is that the Probit and Tobit models must include the same vector of explanatory variables. In fact, if different explanatory variables were included in each of these models, then either the Probit or Tobit model would be misspecified.

If the model is specified as in (1) then there is no reason to estimate coefficients by means of Probit. Hypotheses tests concerning $\underline{\gamma}$ can be formulated in terms of $\underline{\beta}$. Indeed, even if the sole purpose of analysis is to estimate the coefficients of a function which expresses the probability that $y > 0$, Tobit should be used. This is a consequence of the fact that Tobit yields a more efficient estimate of $\underline{\gamma}$ than does Probit--a result which is not surprising because Tobit employs the actual values of the y's, and hence uses more information than does Probit.

The preceding argument for the use of Tobit rather than Probit presupposes the availability of observations on $\underline{y}$ and $\underline{x}$. However, the data for some analyses may contain observations on d and $\underline{x}$ only. In this instance, the Probit estimate of $\underline{\gamma}$ can be used to test hypotheses concerning $\underline{\beta}$, even though an estimate of $\underline{\beta}$ cannot be obtained from Probit.[2] For example, a test of the hypothesis that a particular subset of $\underline{\beta}$ is zero can be formulated in terms of the corresponding subset of $\underline{\gamma}$ via Equation (7), and carried out with observations on d and $\underline{x}$ means of Probit.

In model (1) the two decisions--in the labour supply model, for example, these decisions are whether to work and how many hours to work if work occurs--are determined by essentially the same variables and parameters via Equation (7). Cragg (1971) has recognized this feature of model (1) and developed alternative models which permit the "whether to particupate" decision to be differentiated from the "how much if do participate" decision.

Appendix

The subtraction of (8) from (5) yields

$$\ell_{Pr} - \ell_{T_0} = \sum_{t=T_1+1}^{T} \log F(\underline{x}_t'\underline{\gamma}) + \left(\frac{T-T_1}{2}\right)\log \sigma^2 + \sum_{t=T\ +1}^{T} 1/2(y_t/\sigma - \underline{x}_t'\underline{\gamma})^2. \qquad \text{(A1)}$$

From (A1) the first derivative of $\ell_{Pr} - \ell_{To}$ with respect to $\underline{\gamma}$ is

$$\frac{\partial(\ell_{Pr} - \ell_{T_0})}{\partial\underline{\gamma}} = \sum_{t=T_1+1}^{T} \frac{f(\underline{x}_t'\underline{\gamma})}{F(\underline{x}_t'\underline{\gamma})}\underline{x}_t - \sum_{t=T_1+1}^{T} (y_t/\sigma - \underline{x}_t'\underline{\gamma})\underline{x}_t$$

$$= \sum_{t=T_1+1}^{T} \underline{x}_t\left(\frac{f(\underline{x}_t'\underline{\gamma})}{F(\underline{x}_t'\underline{\gamma})} - y_t/\sigma + \underline{x}_t'\underline{\gamma}\right). \qquad \text{(A2)}$$

Differentiating (A2) with respect to $\underline{\gamma}'$ yields

$$\frac{\partial^2(\ell_{Pr} - \ell_{T_0})}{\partial\underline{\gamma}\partial\underline{\gamma}} = \sum_{t=T\ +1}^{T} \underline{x}_t\left(\underline{x}_t' - \underline{x}_t'\frac{f(\underline{x}_t'\underline{\gamma})}{F(\underline{x}_t'\underline{\gamma})}\underline{x}_t' - \frac{f(\underline{x}_t'\underline{\gamma})^2}{F(\underline{x}_t'\underline{\gamma})^2}\underline{x}_t'\right)$$

$$= \sum_{t=T_1+1}^{T} \underline{x}_t\underline{x}_t'\left(1 - \underline{x}_t'\gamma\frac{f(\underline{x}_t'\underline{\gamma})}{F(\underline{x}_t'\underline{\gamma})} - \frac{f(\underline{x}_t'\underline{\gamma})^2}{F(\underline{x}_t'\underline{\gamma})^2}\right). \qquad \text{(A3)}$$

Using the fact that $\dfrac{\partial^2(\ell_{Pr} - \ell_{T_0})}{\partial\underline{\gamma}\partial\underline{\gamma}'} = \dfrac{\partial^2\ell_{Pr}}{\partial\underline{\gamma}\partial\underline{\gamma}'} - \dfrac{\partial^2\ell_{T_0}}{\partial\underline{\gamma}\partial\underline{\gamma}'}$, Equation (A3)

is presented in the text as Equation (9).

FOOTNOTES

[1]Equation (11) can be inferred directly from the truncated normal distribution material which is presented by Goldberger (1974).

[2]The parameters σ and $\underline{\beta}$ are not identified in the Probit model.

References

Cragg, John G. 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica, 39:829-44.

Crawford, David L., and Steven G. Garber. Forthcoming. The wage work response of female heads. In The Final Report of the Rural Income Maintenance Experiment, ed., D. Lee Rawden, vol. 3, chapter 4.

Goldberger, Arthur S. 1974. Limited dependent variables. Mimeographed. Course Notes for Economics 718. Madison: University of Wisconsin.

Tobin, James. 1958. Estimation of relationships for limited dependent variables. Econometrica, 26:24-36.