The Econometric Analysis of Labour Supply

The Participation and Hours of Work Decisions

Miles Corak
Department of Economics
The Graduate Center, City University of New York
MilesCorak.com @MilesCorak

Lecture 4
Labor Economics I
ECON 87100

Introduction

The econometric analysis of labour supply recognizes the peculiar nature of the endogenous variables that interest us

1. the participation decision is a binary variable

$$P_i = 0 \text{ iff } H_i = 0$$

$$P_i = 1 \text{ iff } H_i > 0$$

2. the hours of work variable is a censored variable

$$H_i = max(h_i^*, 0)$$

Introduction

Two objectives for this lecture

1. Examine the econometrics of binary and limited dependent variable models
2. Use this background to jointly model labour force participation and hours of work decisions

- James Heckman (1974), "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, Vol. 42 (July), 679-94.
  - an illustration of our first "identification strategy", the simultaneous equations model

Univariate Binary Choice Models

We rely on the intuition from our study of Ben-Porath (1973), using the relationship between the wage rate and the reservation wage as defining the labour force participation decision, and thinking probabilistically to model it. What we actually observe:

a sample of $i = 1 \ldots N$ individuals that is randomly selected and independently drawn from the populatioin, where for each i we observe

1. a dependent variable

$$P_i = 1 \text{ if } i \text{ in labour force}$$
$$P_i = 0 \text{ if } i \text{ not in labour force}$$

2. independent variables

$$X_i \equiv \text{a row vector of observed characteristics}$$

- preference related and human capital characteristics

Univariate Binary Choice Models

1. Estimation using Least Squares

specify a linear model and estimate the β with least squares ignoring the fact that the dependent variable only takes two values, 0 and 1.

$$Y_i = X_i\beta + u_i$$

1. the error cannot be assumed to be normally distributed

Y_i	u_i
1	$1 - X_i\beta$
0	$-X_i\beta$

- implying u_i is binomial and discrete, not normal and continuous

Univariate Binary Choice Models

1. Estimation using Least Squares

specify a linear model and estimate the β with least squares ignoring the fact that the dependent variable only takes two values, 0 and 1.

$$Y_i = X_i\beta + u_i$$

1. the error cannot be assumed to be normally distributed

- least squares estimates of β are still unbiased and consistent, and asymptotically $\hat{\beta}$ is distributed normally
- statistical inference is going to be a problem

Univariate Binary Choice Models

1. Estimation using Least Squares

specify a linear model and estimate the β with least squares ignoring the fact that the dependent variable only takes two values, 0 and 1.

$$Y_i = X_i\beta + u_i$$

1. the error cannot be assumed to be normally distributed
2. u_i does not have a constant variance

$$var(u_i) = E(u_i^2)$$

- heteroscedasticity could be remedied using GLS with appropriately devised weights

$$\frac{1}{\sigma_i} = \frac{1}{\sqrt{X_i\hat{\beta}(1-X_i\hat{\beta})}}$$

Univariate Binary Choice Models

1. Estimation using Least Squares

specify a linear model and estimate the β with least squares ignoring the fact that the dependent variable only takes two values, 0 and 1.

$$Y_i = X_i\beta + u_i$$

1. the error cannot be assumed to be normally distributed
2. u_i does not have a constant variance
3. the predicted value is not bounded on the 0,1 interval
   - $X_i\beta$ has a probabilistic interpretation so it must lie on the closed interval $0 \leq X_i\beta \geq 1$, there is nothing to guarantee that $\hat{\beta}$ for all i will respect this restriction
   - predicted values of the model may violate the closed unit interval restriction, and there is no way to address this problem with least squares

Univariate Binary Choice Models

2. The General Set up

Consider a linear model where the dependent variable is a "latent" (unobserved) variable

$$y_i^* = X_i\beta + u_i$$

- y_i^* is assmued to be continuous
- think of it as a type of index, or propensity
- monotonically positively related to the dependent variable of interest, for our purposes it is the difference between the market wage and the reservation wage

u_i is *iid* and for now we make no distributional assumptions

- it has a representation as a probability density function $g(u_i)$ and an associated cummulative distribution function $G(Z_i) \equiv \int_{-\infty}^{Z_i} g(u)du$, where Z_i is any real number

Univariate Binary Choice Models

3. Observable outcomes

Y_i represents the observable outcome

$$Y_i = 1 \text{ iff } y_i^* > 0$$

$$Y_i = 0 \text{ iff } y_i^* \leq 0$$

4. Interpretation

if u_i is the regular error term, it has mean 0, so we assume $E(u_i|X_i) = 0 \; \forall \; i = 1 \ldots N$ implying $E(y_i^*|X_i) = X_i\beta$

the Y_i's represent the realizations of a binomial process

Y_i	Probability
1	$Pr(y_i = 1) = Pr(y_i^* > 0)$
0	$Pr(y_i = 0) = Pr(y_i^* \leq 0)$

Univariate Binary Choice Models

5. Aside on the properties of CDFs

$$G(-\infty) = 0$$
$$G(+\infty) = 1$$

- $G(Z_i) \equiv \int_{-\infty}^{Z_i} g(u)du = Pr(u \leq Z_i)$
- $Pr(u > Z_i) = \int_{Z_i}^{\infty} g(u)du$
- we can express this in terms of $G(Z_i)$ as:

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty} g(u)du - \int_{-\infty}^{Z_i} g(u)du \\
&= G(+\infty) - G(Z_i) \\
&= 1 - G(Z_i)
\end{aligned}
$$

Univariate Binary Choice Models

5. Aside on the properties of CDFs

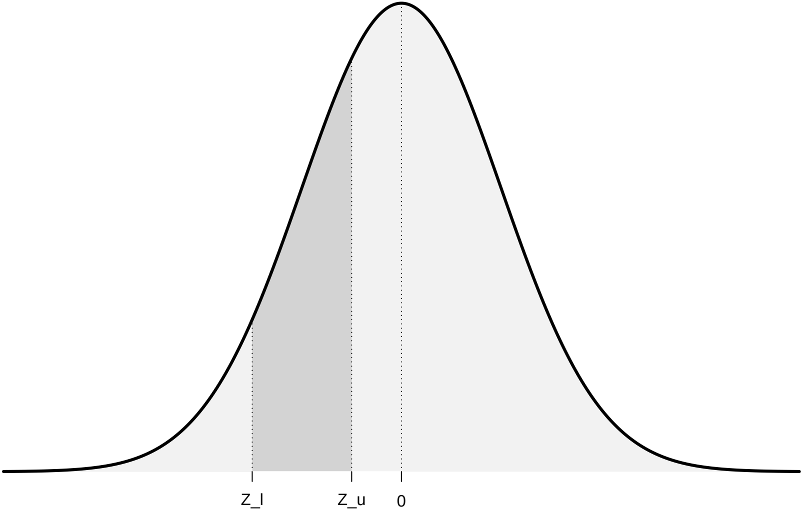
$$g(x) = \frac{dG(x)}{dx}$$

- $Pr(Z_l \leq u \leq Z_u) = \int_{Z_l}^{Z_u} g(u)du$
- as a way of representing all possible probabilities

Univariate Binary Choice Models

$$Pr(Z_l \leq u \leq Z_u) = \int_{Z_l}^{Z_u} g(u)du$$

Representing any probility as an area under the Probability Distribution Function



Univariate Binary Choice Models

5. Aside on the properties of CDFs

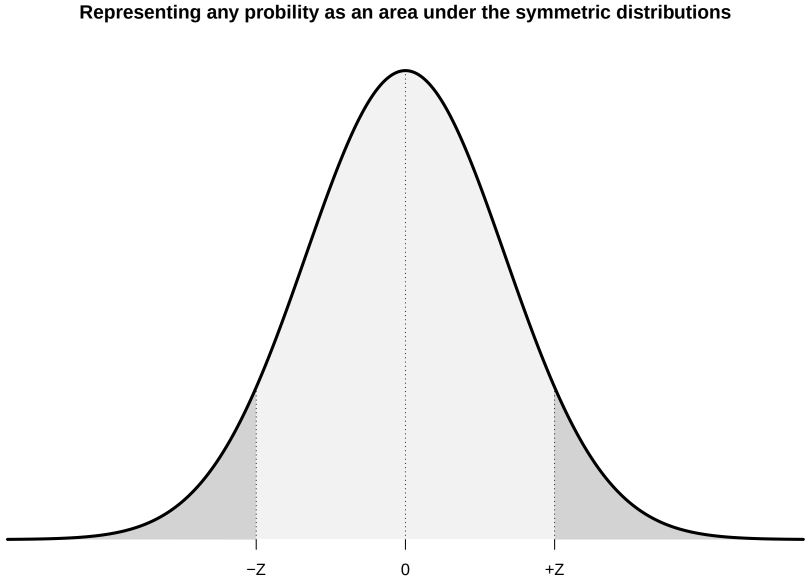
- The form of the PDF and the CDF has not been restricted
- We now restrict our attention to those functions that are "symmetric"
  - this implies that all measures of location—mode, median, mean—are the same
  - consider any value of u, say Z
  - symmetry means $G(-Z) = 1 - G(+Z)$

$$G(-Z) = 1 - G(+Z)$$
$$G(+Z) = 1 - G(-Z)$$

- So that we are speaking of a particular class of distributions for the random variable, the symmetric distributions

Univariate Binary Choice Models

For the class of symmetric distributions $G(-Z) = 1 - G(+Z)$



Univariate Binary Choice Models

6. Characterizing the probability of the outcome

We want to characterize in analytical terms:

$$Pr(Y_i = 1) = Pr(y_i^* > 0)$$
$$Pr(Y_i = 0) = Pr(y_i^* \leq 0)$$

Univariate Binary Choice Models

6. Characterizing the probability of the outcome

$$
\begin{aligned}
Pr(Y_i = 1) &= Pr(y_i^* > 0) \\
&= Pr(X_i\beta + u_i > 0) \\
&= Pr(u_i > -X_i\beta) \\
&= 1 - Pr(u_i \leq -X_i\beta) \\
&= 1 - G(-X_i\beta) \\
&= G(X_i\beta)
\end{aligned}
$$

$$
\begin{aligned}
Pr(Y_i = 0) &= Pr(y_i^* \leq 0) \\
&= 1 - G(X_i\beta)
\end{aligned}
$$

Univariate Binary Choice Models

7. The sample likelihood

for every observation in the sample we can write the contribution of individual i to the sample likelihood:

$$L_i(\beta, ...) = [G(X_i\beta)]^{Y_i}[1 - G(X_i\beta)]^{1-Y_i}$$

$$Y_i = 1 \implies L_i = G(X_i\beta)$$
$$Y_i = 0 \implies L_i = 1 - G(X_i\beta)$$

Univariate Binary Choice Models

7. The sample likelihood function

the sample likelihood function for the full set of N randomly and independently selected observations is:

$$L(\beta) = \prod_{i=1}^{N} L_i(\beta)$$

$$= \prod_{i=1}^{N} [G(X_i\beta)]^{Y_i}[1 - G(X_i\beta)]^{1-Y_i}$$

$$= \prod_{Y_i=1} [G(X_i\beta)] \prod_{Y_i=0} [1 - G(X_i\beta)]$$

there are two sets of observations, so we can partition the sample

Univariate Binary Choice Models

7. The sample log likelihood function

for computational purposes the log likelihood function is often used

$$
\begin{aligned}
l(\beta) &= lnL(\beta) \\
&= \sum_{i=1}^{N} \left\{ Y_i \ln[G(X_i\beta)] + (1 - Y_i) \ln[1 - G(X_i\beta] \right\} \\
&= \sum_{Y_i=1} \ln[G(X_i\beta)] + \sum_{Y_i=0} \ln[1 - G(X_i\beta)]
\end{aligned}
$$

we have developed a general framework and have, except for invoking symmetry, put no restrictions on the actual form of the distribution funciton

- implementation requires us to make an assumption as to the functional form

Univariate Binary Choice Models

8. Different distributions imply different models

    1. The univariate probit model

        - assume the univariate normal distribution, with mean μ and standard deviation $\sigma > 0$

$$g(x) = \frac{1}{(2\pi)^{1/2}} exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$$

    2. The univariate logit model

        - assume the univariate logistic distribution, which has a simple analytical expression for the probability density function

$$g(x) = \frac{exp(x)}{[1+exp(x)]^2}$$

Censored Regression Model

There are two types of observations in the data

- **non limit observations** in which the outcome is observed

$$y_i = y_i^*, y_i^* > 0$$

- **limit observations** in which the outcome is not observed

$$y_i = 0, y_i^* \leq 0$$

We can't apply standard methods to the estimation of $y_i^* = X_i\beta + u_i$

- if we use least squares $\hat{\beta}$ will be biased and inconsistent whether we drop the limit observations and use only the non limit observations, or if we set the limit observations to zero

Censored Regression Model

1. The non-limit observations

Since $y_i = max(y_i^*, 0)$ we observe y_i when $y_i^* > 0$.

- This is calling for a conditional density function: $\phi(y_i | y_i > 0)$, the distribution of y_i given that it is observed.
- we need $\phi(y_i | y_i > 0) \times Pr(y_i > 0)$
  - $\phi(y_i | y_i > 0)$ is the traditional normal density function
  - $Pr(y_i > 0)$ is a probability that may be represented analytically

Censored Regression Model

1. The non-limit observations

representing $Pr(y_i > 0)$ analytically

$$
\begin{aligned}
Pr(y_i > 0) &= Pr(y_i^* > 0) \\
&= Pr(X_i\beta + u_i > 0) \\
&= Pr\Big(\frac{u_i}{\sigma} > \frac{-X_i\beta}{\sigma}\Big) \\
&= 1 - Pr\Big(\frac{u_i}{\sigma} \leq \frac{-X_i\beta}{\sigma}\Big) \\
&= 1 - F\Big(\frac{-X_i\beta}{\sigma}\Big) \\
&= F\Big(\frac{X_i\beta}{\sigma}\Big)
\end{aligned}
$$

Censored Regression Model

1. The non-limit observations

representing $\phi(y_i|y_i > 0)$ analytically

$$\phi(y_i|y_i > 0) = \frac{\phi(y_i)}{Pr(y_i > 0)}$$
$$= \frac{\phi(y_i)}{F\left(\frac{X_i\beta}{\sigma}\right)}$$

Censored Regression Model

1. The non-limit observations

representing $\phi(y_i | y_i > 0)$ analytically

the contribution to the sample likelihood function of each non-limit observation can be derived

- using the change of variable theorem to standardize $\phi(y_i)$ to have a mean of zero and standard deviation of one

$$\phi(y_i) = \frac{1}{\sigma} f(\frac{y_i - X_i\beta}{\sigma})$$

Censored Regression Model

1. The non-limit observations

representing $\phi(y_i|y_i > 0)$ analytically

the contribution to the sample likelihood function of each non-limit observation can be derived

$$\phi(y_i|y_i > 0) \times Pr(y_i > 0) = \frac{1}{\sigma}\frac{f(\frac{y_i - X_i\beta}{\sigma})}{F(\frac{y_i - X_i\beta}{\sigma})} \times F\Big(\frac{y_i - X_i\beta}{\sigma}\Big)$$
$$= \frac{1}{\sigma} f\Big(\frac{y_i - X_i\beta}{\sigma}\Big)$$

this is the contribution to the sample likelihood of each non-limit observation

Censored Regression Model

2. The limit observations

representing $y_i^* \leq 0$ analytically, where $F(.)$ is the standardized univariate normal distribution

$$\begin{aligned} Pr(y_i^* \leq 0) &= Pr(X_i\beta + u_i \leq 0) \\ &= Pr(u_i \leq -X_i\beta) \\ &= Pr\Big(\frac{u_i}{\sigma} \leq \frac{-X_i\beta}{\sigma}\Big) \\ &= F\Big(\frac{-X_i\beta}{\sigma}\Big) \\ &= 1 - F\Big(\frac{X_i\beta}{\sigma}\Big) \end{aligned}$$

Censored Regression Model

3. The lilelihood function

contribution of the ith observation to the sample likelihood is

$$\left[\frac{1}{\sigma} f\left(\frac{y_i - X_i\beta}{\sigma}\right)\right]^{Y_i}\left[1 - F\left(\frac{X_i\beta}{\sigma}\right)\right]^{1-Y_i}$$

Censored Regression Model

3. The lilelihood function

sample likelihood function for N independently and randomly selected observations

$$
\begin{aligned}
L(\beta,\sigma) &= \prod_{i=1}^{N}\left[\frac{1}{\sigma}f\left(\frac{y_i - X_i\beta}{\sigma}\right)\right]^{Y_i}\left[1 - F\left(\frac{X_i\beta}{\sigma}\right)\right]^{1-Y_i} \\
l(\beta,\sigma) &= \ln L(\beta,\sigma) \\
&= \sum_{i=1}^{N} Y_i\left[\ln\left(\frac{1}{\sigma}\right) + \ln f\left(\frac{y_i - Xi\beta}{\sigma}\right)\right] + (1 - Y_i)\ln\left[1 - F\left(\frac{X_i\beta}{\sigma}\right)\right]
\end{aligned}
$$

this is the sample log likelihood function for a one-limit Tobit model, in which β and σ are separately identifiable (unlike the univariate binary model)

Heckman's SEM Tobit model of labour supply

1. Introduction

Our study of this paper is meant as an illustration:

1. of the use of theory to guide empirical research
2. an example of an identification strategy to estimate causal parameters

The paper is a response to the then prevailing practice of:

1. using a common set of variables to explain wage rates, hours of work, and the decision to work (among married women)
2. the tendency to throw out information and work with a selected sample that will imply biased results
   - censored samples
   - truncated samples

Heckman's SEM Tobit model of labour supply

1. Introduction

A wage function and a reservation wage function form the structural equations of a system that may be used to derive a common set of parameters determining the probability a women works, and her:

- hours of work
- observed wage rate
- reservation wage

The empirical approach is an extension of Tobit to Simultaneous Equations Models, using the entire sample of observations.

Heckman's SEM Tobit model of labour supply

2. Shadow prices and market wages

Constrained utiility maximization permits the derivation of a shadow price function

- for leisure demand we refer to this as the reservation wage function
- reservation wage functions are defined at corners, and two corner solutions are possible
  - all time is spent in market work when the wage rate exceeds the reservation wage
  - no time is spent in market work when hte wage rate is less than the reservation wage

Heckman's SEM Tobit model of labour supply

2. Shadow prices and market wages

The reservation wage function

$$W^* = g(h, W_m, P, A, Z)$$

W^* marginal value of a women's time, $W^* = P \times MRS_{XL}$

h hours of work, $\delta W^*/\delta h > 0$

W_m husband's wage rate

P vector of goods prices

A asset income, $\delta W^*/\delta A > 0$ if leisure is normal

Z other exogneous contraints

Heckman's SEM Tobit model of labour supply

2. Shadow prices and market wages

The reservation wage function

$$W^* = g(h, W_m, P, A, Z)$$

The market wage function

$$W = B(E, S)$$

informed by human capital theory

S number of years of schooling, $\delta W / \delta S > 0$

E number of years of labour market experience, $\delta W / \delta E > 0$

Heckman's SEM Tobit model of labour supply

2. Shadow prices and market wages

The reservation wage function

$$W^* = g(h, W_m, P, A, Z)$$

The market wage function

$$W = B(E, S)$$

If working hours can adjust, then an interior solution implies that $W = W^*$ as an equilibrium condition. If the individual is a non-participant, then $W^* \geq W$.

Heckman's SEM Tobit model of labour supply

3. Specification of functional form

$$
\begin{aligned}
l(W_i^*) &= \beta_0 + \beta_1 h_i + \beta_2 (W_m)_i + \beta_3 P_i + \beta_4 A_i + \beta_5 Z_i + \epsilon_i \\
l(W_i) &= b_0 + b_1 S_i + b_2 E_i + u_i
\end{aligned}
$$

ϵ_i and u_i are jointly normally distributed with mean zero and possibly correlated. The disturbances are uncorrelated with the regressors. Observed hours of work will depend upon the disturbances.

Heckman's SEM Tobit model of labour supply

4. Reduced forms for working women

$W_i^* = W_i$ is an equilibrium condition that permits the derivation of reduced form equations for observed wages and hours

$$h_i = \frac{1}{\beta_1}[b_0 - \beta_0 + b_1 S_i + b_2 E_i - \beta(W_m)_i - \beta_3 P_i - \beta_4 A_i - \beta_5 Z_i] + \frac{u_i - \epsilon_i}{\beta_1}$$

$$l(W_i) = b_0 + b_1 S_i + b_2 E_i + u_i$$

The observations to estimate these two equations are available only if the women are working. The distributions of the disturbances are conditional upon W_i^*, and are hence conditional distributions.

Heckman's SEM Tobit model of labour supply

5. Joint distribution of observed hours and wages for the i^{th} working women

$$j(h_i, l(W_i)|W_i^* < W_i)_{h=0} = \frac{n(h_i, l(W_i))}{pr([W_i > W_i^*]_{h=0})}$$

j(.) is the conditional distribution

n(.) is an unconditional distriubtion, a multivariate normal density

pr(.) is the probability that a women works, a univariate cummulative normal distribution function

Heckman's SEM Tobit model of labour supply

6. The likelihood function

maximization of this function yields consistent asymptotically unbiased, and efficient parameter estimates that are asymptotically normally distributed

$$
\begin{aligned}
L = & \prod_{i=1}^{K} j(h_i, l(W_i) | W_i > W^*_{i,h=0}) pr([W_i > W_i^*]_{h=0}) \\
& \times \prod_{i=K+1}^{T} pr([W_i < W_i^*]_{h=0}) \\
= & \prod_{i=1}^{K} n(h_i, l(W_i)) \prod_{i=K+1}^{T} pr([W_i > W_i^*]_{h=0})
\end{aligned}
$$

T refers to the sample size, K of whom work and $T - K$ do not.

Heckman's SEM Tobit model of labour supply

7. Summing up

An empirical approach that is no longer defendable:

1. exclusion restrictions are arbitrarily imposed
2. explicit functional form and distributional assumptions are required

The challenge in developing an identification strategy involves explicitly understanding the sources of exogenous variation in the data.