Female Labor Supply

- The main difference between female and male labor supply models is that the estimation of the former requires to make serious attempts to incorporate the participation information (boundary solution) contained in the labor supply decision rule (expression (5)).
- It might seem almost unbelievable now, but until the mid 1970s the standard practice in empirical research was to estimate labor supply functions over samples of working women only.
- This was done either directly by confining the analysys exclusively to labor force participants or indirectly by imputing wage to nonworking women from a wage equation estimated over a sample of workers.
- The first step towards seriously taking into account this censoring problem was to recognize that labor force participation involves a comparison between market wage offers, W_m, and the value of home time at zero hours of work or reservation wage, W_r.

- This means that a participation rate of, say, 55%, indicates that 55% of the women have market wages that exceed their reservation wages. Ben-Porath and Lewis in the early 1970s were the first ones to clearly state this issue.

- Gronau (1974) developed more formally the interpretation of the participation equation suggested by Ben-Porath and Lewis. The probability of participation was seen as depending upon whether offered market wages exceed reservations wages, where

$$W_m = \alpha X + u_1 \tag{1}$$

$$W_r = \beta Y + u_2, \tag{2}$$

then the probability of participation can be written as

$$\frac{\alpha X - \beta Y}{\sigma_p} > \frac{u_2 - u_1}{\sigma_p} = \frac{u_p}{\sigma_p} \tag{3}$$

X and Y are the systematic determinants and u_1 and u_2 are normally distributed errors in the market and reservation wages equations, respectively.

- Notice that the participation error u_p is distributed as a standardized normal, probit estimation of this probability was appropriate.

- Gronau in related work raised the issue of censoring or selectivity bias in the wage generating function for women. The difficulties estimating labor supply and wage equations are closely linked because identical sample censoring issues arise. The expected wage for workers may be written as

$$E(W_m) = \alpha X + E\left(u_1 | \frac{\alpha X - \beta Y}{\sigma_p} > \frac{u_2 - u_1}{\sigma_p}\right) \tag{4}$$

If one estimates a wage equation using samples of working women, biases result because the same sets of variables that determine wages enter in as a criterion for sample eligibility. The estimated wage function confounds the true behavioral wage function with the rules of sample inclusion.

- The idea is that when making comparisons between groups of people who differ in their labor force participation rates, some of the observed differences may have little to do with behavioral disimilarities but merely reflect sample censoring.

- Heckman (1974, Ecta) extended Gronau's work in many dimensions. It relaxed the lack of correlation between the errors in the reservations wages and market wages equations. Second, he integrated in one framework decisions regarding wages, hours, and participation. The approach allows one to estimate a common set of parameters which underly the function determining the probability that a woman works, her hours of work, her observed wage rate, and her shadow wage.

- The shadow wage is defined as the marginal value placed on a woman's leisure evaluated at each unit of leisure, and therefore hours of work.

- We can write the model as follows:

$$W_s = b_1 H + \beta Y + u_2 \tag{5}$$

$$W_m = \alpha X + u_1. \tag{6}$$

  A woman works if $W_m > W_{s|h=0}$, that means her market wage exceeds her shadow wage at zero hours of work. For working women, shadow and market wages are equated at the margin, so that one can solve for the hours equation

$$H = 1/b_1(W_m - \beta Y) - \frac{u_2}{u_1} \; if \; \alpha X - \beta Y > u_2 - u_1$$

$$H = 0 \; if \; \alpha X - \beta Y < u_2 - u_1$$

$$W_m = \alpha X + u_1 \; if \; \alpha X - \beta Y > u_2 - u_1$$

$$W_m \; unobserved \; if \; \alpha X - \beta Y < u_2 - u_1.$$

Given these equations it is possible to write the likelihood function for the observed sample distribution of hours and wages for K workers and $N-K$ nonworkers. Then estimate the model by Maximum Likelihood. The model is methodologically very advanced, but computationally complicated, so it has been less used than simpler methods.

- Heckman in 1975 circulated a five-page draft on sample censoring that was the seed of his celebrated 1979 Ecta paper. Building upon the work of Amemiya, he presented the problem of censoring bias as another variant of an specification error.

  Consider a random sample of I observations. The equations for individual i may be written as

$$Y_{1i} = X_{1i}\beta_1 + U_{1i} \tag{7}$$

$$Y_{2i} = X_{2i}\beta_2 + U_{2i}, \tag{8}$$

  where X is a vector of exogenous regressors. The unconditional expectation of the error terms are zero, and there is potential correlation between the two error terms.

- One implicit, and non-trivial assumption of this setup is that if all data were available, unbiased estimators of the parameters of each equation could be achieved by least squares, and all parameters would be identified.

- Suppose one seeks to estimate (7), but data are missing on Y_{1i} for certain observations. The crucial issue is why are some observations missing?

  Then we can write the regression function for the subsample of available data as

$$E(Y_{1i}|X_{1i}, sampleselectionrule) = X_{1i}\beta_1 + E(U_{1i}|sampleselectionrule)$$

If the conditional expectation of U_{1i} is zero, the selected sample regression function is the same as the population regression function. Then we can use least squares using the subsample of available data to estimate the population regression function. The only cost here is efficiency.

- But in the more general case consider a selection rule where we only observe values of Y_{1i} when $Y_{2i} \geq 0$. Then we could define a dummy variable to be equal to one under that condition.
- We can write

$$E(U_{1i}|sampleselectionrule) = E(U_{1i}|Y_{2i} \geq 0) = E(U_{1i}|U_{2i} \geq -X_{2i}\beta_2) \tag{9}$$

If U_{1i} and U_{2i} are independent, then the selection rule is independent of the behavioral function being estimated, the conditional mean of U_{1i} is zero.

- But in general, the conditional mean of the U_{1i} disturbance does not vanish. Then the selected sample regression function may be written as

$$E(Y_{1i}|X_{1i}, Y_{2i} > 0) = X_{1i}\beta_1 + E(U_{1i}|U_{2i} \geq -X_{2i}\beta_2). \qquad (10)$$

The selected sample regression function depends on X_{1i} and X_{2i}. Regression estimators of (7) computed on the selected sample omit the final term above. Thus the problem of sample selection bias, initially viewed as a missing dependent variable problem, may be formulated as an ordinary omitted explanatory variable problem.

- Using known results Heckman shows that

$$E(U_{1i}|U_{2i} > -X_{2i}\beta_2) = \frac{\sigma_{12}}{(\sigma_{22})^{1/2}}\lambda_i$$

where

$$\lambda_i = \frac{\phi(Z_i)}{1 - \Phi(Z_i)} \qquad (11)$$

and

$$Z_i = -\frac{X_{2i}\beta_2}{(\sigma_{22})^{1/2}}$$

This means that

$$E(Y_{1i}|X_{1i}, Y_{2i} \geq 0) = X_{1i}\beta_1 + \frac{\sigma_{12}}{(\sigma_{22})^{1/2}}\lambda(Z_i). \qquad (12)$$

- Therefore, the bias in estimating wage and hours over samples of working women results from the omission of the variable λ. If one knew or could estimate that variable, then it would be possible to obtain consistent wage and hours equations with samples of workers.

- Heckman suggests to estimate λ from a first-stage probit on the probability of participation. This formulation clarifies the biases in estimating relationships using only the sample of workers. The bias depends on the sign of λ in the main equation, as well as the sign of the coefficients in the auxiliary regression.

- The simplicity of this technique has made it one of the most popular methods in applied econometrics.

- Notice that λ is the Inverse Mills' ratio, and is the ratio of the ordinate of a standard normal to the tail area of the distribution

  $$\lambda_i = \frac{\phi(Z_i)}{1 - \Phi(Z_i)} \tag{13}$$

  Notice that the denominator is the probability that a population observation with characteristics X_{2i} is selected into the sample.

- One of the advantages of this methodology is that if all population observations have an equal chance of being sampled, λ will be essentially zero, and the least squares estimator has optimal properties.

- Also, if the disturbances affecting sample selection are independent of the disturbances affecting the behavioral functions of interest, then we can omit the Inverse Mills' ratio from the regression.

- Using well-known theorems involving truncated normal distributions, he proves that the expected wage and hours for a sample of workers can be written as

$$E(W_m) = \alpha X + \sigma_{1p}\lambda \tag{14}$$

$$E(h) = \beta_0(\alpha X - \beta Y) + \sigma_p \lambda, \tag{15}$$

where λ is the ratio of the height of the density to the right tail area. σ_{1p} is the correlation between the errors in the participation and wage equations, and σ_p is the standard deviation of the participation error.

- Notice that we will only observe market wages for certain women because their productivity in the market exceeds their productivity in the home, but this does not mean that the more market productive women are the ones observed working.

The Heckman and MaCurdy Model

Adopting the Heckman and MaCurdy (1980, 1982) utility function (which closely resembles MaCurdy (1981)),

$$U_i(C(t),L(t)) = \gamma_{C_{i,t}} \cdot (\frac{C^{\beta}(t)-1}{\beta}) + \gamma_{L_{i,t}} \cdot (\frac{L^{\alpha}(t)-1}{\alpha}), \tag{16}$$

where $\alpha, \beta < 1$ for concavity. The λ-constant labor supply/leisure demand function of women i is

$$L_i(t) = T - H_i(t) = \begin{cases} \left\{\frac{1}{\gamma_{L_{i,t}}} \cdot (\frac{1+\rho}{1+r})^t \cdot \lambda_i(0) \cdot W_i(t)\right\}^{\frac{1}{\alpha-1}} & \text{if women } i \text{ works;} \\ T & \text{otherwise,} \end{cases} \tag{17}$$

Notice that as a result of the different setup, the intertemporal elasticity of substitution between hours worked (denoted by δ above) is now $-\frac{1}{\alpha-1} > 0$. Taking natural logarithms of this function, and approximating $\ln\frac{1+\rho}{1+r}$ by $\rho - r$ as above, the equations become

$$\ln L_i(t) = \ln(T - H_i(t)) = \begin{cases} \frac{1}{\alpha-1}\left\{\ln\lambda_i(0) + (\rho - r)\cdot t - \ln\gamma_{L_{i,t}} + \ln W_i(t)\right\} & \text{if women } i \text{ works;} \\ \ln T & \text{otherwise,} \end{cases} \tag{18}$$

Suppose that observed and unobserved heterogeneity and the individual's wage profile are specified in the following way

$$\ln\gamma_{L_{i,t}} = X_i(t)\phi + \varepsilon_{1,i}(t), \tag{19}$$

$$\ln W_i(t) = Z_i(t)\kappa + \varepsilon_{2,i}(t),$$

where the errors are assumed to be structured as follows

$$\varepsilon_{1,i}(t) = \eta_{1,i} + v_{1,i}(t), \tag{20}$$

$$\varepsilon_{2,i}(t) = \eta_{2,i} + v_{2,i}(t).$$

The labor supply equation may now be written as

$$\ln(T - H_i(t)) = \begin{cases} F_i + \frac{\rho-r}{\alpha-1}\cdot t - Z_i(t)\frac{\phi}{\alpha-1} - X_i(t)\frac{\kappa}{\alpha-1} + \frac{1}{\alpha-1}(v_{2,i}(t) - v_{1,i}(t)) \\ \quad \text{if } \frac{1}{\alpha-1}(v_{2,i}(t) - v_{1,i}(t)) \leq -F_i - \frac{\rho-r}{\alpha-1}\cdot t + Z_i(t)\frac{\phi}{\alpha-1} + X_i(t)\frac{\kappa}{\alpha-1} + \ln T, \\ \ln T \;\; \text{if } \frac{1}{\alpha-1}(v_{2,i,t} - v_{1,i,t}) > -F_i - \frac{\rho-r}{\alpha-1}\cdot t + Z_i(t)\frac{\phi}{\alpha-1} + X_i(t)\frac{\kappa}{\alpha-1} + \ln T, \end{cases} \tag{21}$$

where $F_i = \frac{1}{\alpha-1}(\ln\lambda_i(0) - \eta_{1,i} + \eta_{2,i})$. The constant individual-specific taste parameters $\eta_{1,i}$, and $\eta_{2,i}$ are allowed to be arbitrarily correlated. The taste components $u_1 = \frac{1}{\alpha-1}(v_{2,i}(t) - v_{1,i}(t))$, and $u_2 = v_{2,i}(t)$ are assumed to be mean zero bivariate normally distributed with no serial correlation, i.e.

$$E(u_{1,i}(t)) = E(u_{2,i}(t)) = 0, E(u_{n,i}(t)u_{m,i}(t')) = \sigma^2_{nm}, \ \ n,m = 1,2, t = t',$$

$$E(u_{n,i}(t)u_{m,i}(t')) = 0, \ \ n,m = 1,2, t \neq t'.$$

In this form, the labor supply model combined with the wage equation yields a two equation Panel Tobit system that can be estimated using Maximum Likelihood. F_i is treated as a fixed effect in the estimation. Notice that a second fixed effect, η_2, arises from the wage equation (19). As in MaCurdy (1981), in a second step Heckman and MaCurdy (1980, 1982) regress the predicted fixed effect on the determinants of the marginal utility of wealth (wealth, future fertility, etc.). Women who never supply labor to the labor market have to be excluded from the estimation sample since the fixed effects cannot be identified. We will abstract from this additional complication in the construction of the likelihood function below.[1] Given the assumptions on the error the joint density of u_1 and u_2 is the following bivariate normal density

$$f_{u_1,u_2}(u_1,u_2) = \frac{1}{2\pi\sigma_{11}\sigma_{22}(1-\xi^2)^{\frac{1}{2}}} \ e^{-\frac{1}{2(1-\xi^2)}\left\{(\frac{u_1}{\sigma_{11}})^2 - 2\xi\frac{u_1}{\sigma_{11}}\frac{u_2}{\sigma_{22}} + (\frac{u_2}{\sigma_{22}})^2\right\}},$$

where $\xi = \frac{\sigma_{12}}{\sigma_{11}\sigma_{22}}$ is the correlation coefficient. We can always write the joint density as $f_{u_1,u_2}(u_1,u_2) = f_{u_2|u_1}(u_2|u_1) \cdot f_{u_1}(u_1)$, where $f_{u_1}(u_1)$ is the marginal density of $f_{u_1,u_2}(u_1,u_2)$ with respect to u_1. Using the previous result, we need the density of $f_{u_1}(u_1)$ and the conditional density $f_{u_2|u_1}(u_2|u_1)$ to form the bivariate normal Tobit likelihood function. It can be shown that the marginal density is

$$f_{u_1}(u_1) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_{11}} \ e^{-\frac{1}{2}\left\{\frac{u^2_{1,i}(t)}{\sigma^2_{11}}\right\}}, \tag{22}$$

and the conditional is

$$f_{u_2|u_1}(u_2|u_1) = \frac{1}{(2\pi)^{\frac{1}{2}}(\sigma^2_{22} - \frac{\sigma^4_{12}}{\sigma^2_{11}})^{\frac{1}{2}}} \ e^{-\frac{1}{2}\left\{\frac{(u_{2,i}(t) - \frac{\sigma^2_{12}}{\sigma^2_{11}}u_{1,i}(t))^2}{\sigma^2_{22} - \frac{\sigma^4_{12}}{\sigma^2_{11}}}\right\}}. \tag{23}$$

[1] Heckman and MaCurdy 1980 (see appendix there) also present results that control for this type of selection and find little evidence that it matters using a sample of white women from the PSID 1968-75.

The log-likelihood for a sample of N women who are observed for $K+1$ periods takes the following form (not essential constants are omitted)

$$\ln L = \sum_{i=1}^{N}\sum_{t=0}^{K}(1-d_i(t))\ln\Phi\left(\frac{R_i(t)-\ln T}{\sigma_{11}}\right) \tag{24}$$

$$+\sum_{i=1}^{N}\sum_{t=0}^{K} d_i(t)\left\{-\frac{1}{2}\ln\sigma_{11}^2 - \frac{1}{2\sigma_{11}^2(\ln(T-H_i(t))-R_i(t))^2}\right\}$$

$$+\sum_{i=1}^{N}\sum_{t=0}^{K} d_i(t)\left\{-\frac{1}{2}\ln\left(\sigma_{22}^2-\frac{\sigma_{12}^4}{\sigma_{11}^2}\right)-\frac{1}{2}\frac{\left\{\ln W_i(t)-X_i(t)\kappa-\eta_{2,i}-\frac{\sigma_{12}^2}{\sigma_{11}^2}(\ln(T-H_i(t))-R_i(t))\right\}^2}{\sigma_{22}^2-\frac{\sigma_{12}^4}{\sigma_{11}^2}}\right\},$$

where $R_i(t)=F_i+\frac{\rho-r}{\alpha-1}\cdot t - Z_i(t)\frac{\phi}{\alpha-1}-X_i(t)\frac{\kappa}{\alpha-1}$, and $d_i(t)$ equals 1 if the individual participates in the labor force in period t and equals 0 otherwise. The first RHS expression represents the likelihood contribution of the censored observations from the participation decision, the second part is the labor supply/leisure demand equation using the marginal distribution, expression (22), conditional on participation, and the third part is the likelihood contribution from the wage equation using the conditional density, expression (23), also conditional on participation. Note that the estimation of the fixed effects using this likelihood function requires that participation occurs at least once for each sample member. Heckman and MaCurdy also present maximum likelihood estimates that correct for women who never participate.

For a detailed discussion of the results see Heckman and MaCurdy (1980, 1982). Most notably, in their revised version they report that the intertemporal elasticity of leisure is $\frac{1}{\alpha-1}=-0.406$ based on an eight year panel sample of 672 married women, age 30-65 in 1968, from the 1968-1975 PSID. Individual characteristics are shown to have the expected effects: labor supply is inversely related to wealth and young children affect the value of time at home ('leisure time').

References

- G. Becker and G. Ghez, ”The Allocation of Time and Goods Over the Life-Cycle,” Columbia University Press, 1975.
- J. Heckman, ”Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption Over the Life Cycle,” American Economic Review (March 1974).
- S. Frederick, G. Loewenstein, and T. O'Donoghue, ”Time Discounting and Time Preference: A Critical Review,” Journal of Economic Literature (June 2002).
- T. MaCurdy, ”An Empirical Model of Labor Supply in a Life-Cycle Setting,” Journal of Political Economy, (December 1981).
- J. Heckman and T. MaCurdy, ”A Life Cycle Model of Female Labor Supply,” The Review of Economic Studies (Jan. 1980).
- J. Heckman and T. MaCurdy, ”Corrigendum on A Life Cycle Model of Female Labor Supply,” The Review of Economic Studies (May 1982).
- J. Altonji, ”Intertemporal Substitution in Labor Supply: Evidence from Micro data,” The Journal of Political Economy 94(3), Part 2: Hoover Institution Labor Conference. (Jun., 1986).
- D. Card, ”Intertemporal Labor Supply: An Assessment,” in C. Sims, ed., Advances in Econometrics Sixth World Congress, vol. II, Cambridge University Press, 1994.
- F. Laisney, W. Pohlmeier, and M. Staat, ”Estimation of Labour Supply Functions Using Panel Data: A Survey,” in Matyas and Sevestre, eds., The Econometrics of Panel Data, 1996.