

The Tobit Model

Quantitative Microeconomics

R. Mora

Department of Economics
Universidad Carlos III de Madrid

Outline

- 1 Motivation: The Female Labour Supply
- 2 The Married Women Labor Supply and the Tobit Model
- 3 ML Estimation for the Tobit Model

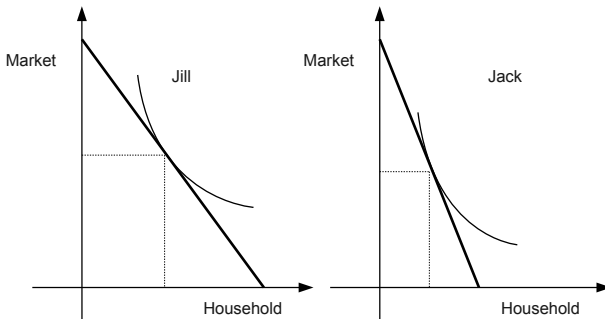
The House Allocation of Time in Real Life

%	Leisure	Personal	Household	Market
Married Women	13	47	30	10
Married Men	14	46	16	24
Unmarried Women	14	48	21	17
Unmarried Men	15	45	21	19

Source: University of Michigan, US data

Time Allocation among Singles

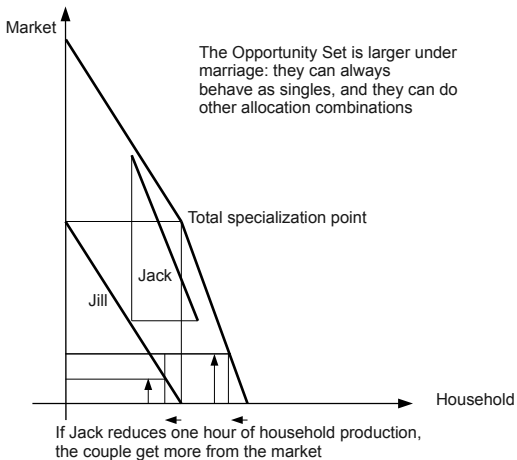
Jack is less productive In the Household



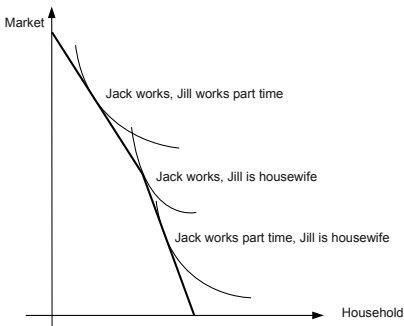
Economic Theories on Marriage

- Principal-Agent Problem: taking care of the kids and the car by people you trust
- Scale Economies: the marginal kid hardly increases the costs
- Risk-Sharing: pooling resources reduces risks
- none of these economic channels explain why married women work mostly in the household

Specialization: Opportunity Frontier of Married Couple



Alternative Marriage Agreements



Whether Jack works full time and Jill works part time depends on abilities and preferences

Jack Works Full Time, Jill Works Part Time

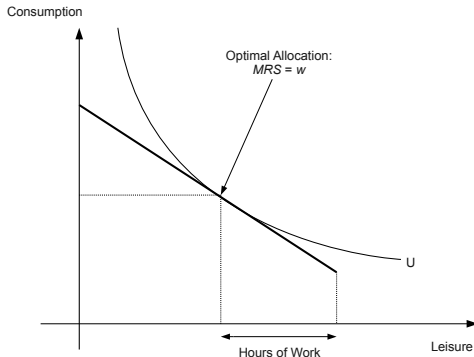
- market activities are more productive than household activities
⇒ both work a few hours in the market
- women hourly wages tend to be lower than men's ⇒ if they are equally productive in the household, then women have a comparative advantage in working in the household
- we will revisit the crucial issue of why women earn lower wages when we look at selection issues
- now we are going to try to understand the determinants of female labor force participation given wages

The Married Women Labor Supply

It can be decomposed into two stages

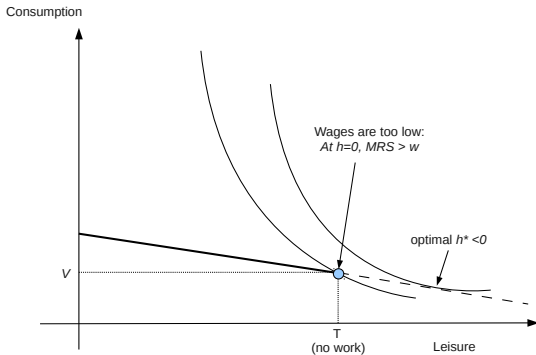
- first decision: participation decision: $w > w_R$
 - second decision: if $w > w_R$, how many working hours?
-
- the first decision is like a Probit model because the participation decision is binary
 - the second decision is like a linear regression model because the amount of time worked can be considered continuous
 - both decisions are strongly linked: factors that make a married woman more likely to participate make her work more hours

The Optimal Allocation of Leisure: Internal Solution



internal solution: $h = h^* (MRS = w) > 0$

The Optimal Allocation of Leisure: Corner Solution



corner solution: $h = 0$ if $h^* (MRS = w) \leq 0$

The Tobit Model

Example: Married Women Labor Supply

- optimality condition ($MRS = w$): $h^* = \beta x + \varepsilon, \varepsilon \sim N(0, \sigma^2)$
- participation condition ($MRS < w$): $h^* > 0$
- If $h^* > 0$, then actual hours of work: $h = h^*$
- If $h^* \leq 0$, then actual hours of work: $h = 0$

$$h = \max\{0, \beta x + \varepsilon\}, \varepsilon \sim N(0, \sigma^2)$$

Labor Supply Controls

Which controls should be in vector x ?

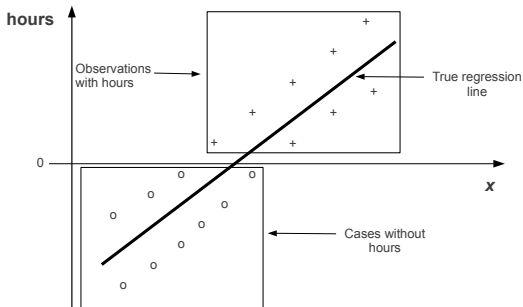
- Personal: Non-labor income, spouse's income, number of kids, human capital,...
- Economic conditions: market wages, unemployment rates,...
- Strictly speaking, for the labor supply we require the wage offers. This creates two problems:
 - We do not have information on wage offers for those who are not working.
 - A worker's wage offer is likely related to unobservable characteristics which arguably affect simultaneously the worker's labor supply: Wages and hours worked are simultaneously determined for each worker.

Observable Data

- the econometrician observes whether the married woman participates in the labor market or not
- if the married woman participates, then the econometrician observes the hours of work
- if the married woman does not participate, the econometrician does not observe the optimal number of hours that the married woman would choose to work (in this case, it would be a negative number)

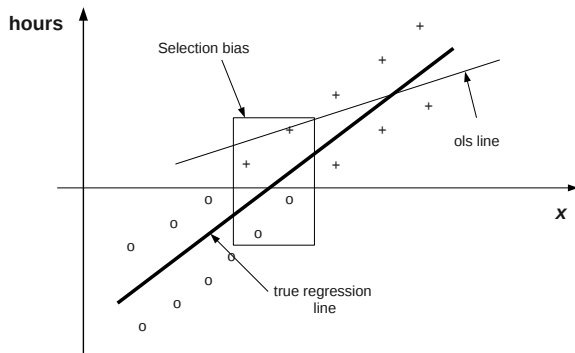
Using Only the Married Women Who are Working

Can we estimate β by OLS using only the data from the married women who choose to work?



Selection Bias

The OLS sample is not iid: we only observe (h_i, x_i) if $h_i > 0$



ML Estimation (1/2)

If we estimate by Maximum Likelihood, we use the full sample: including women who choose to work with information of the hours they work and also women who choose not to work

Density of a woman who works $h_i > 0$ hours

$$\begin{aligned} f(h_i | x_i) &= f(\beta_0 x_i + \varepsilon_i | x_i) \\ &= \left(\frac{1}{\sigma_0}\right) \phi\left(\frac{\varepsilon_i}{\sigma_0}\right) \end{aligned}$$

Probability that a woman does not work ($h_i = 0$)

$$\begin{aligned} \Pr(h_i = 0 | x_i) &= \Pr(\beta_0 x_i + \varepsilon_i \leq 0 | x_i) \\ &= 1 - \Phi\left(\frac{\beta_0 x_i}{\sigma_0}\right) \end{aligned}$$

ML Estimation (2/2)

- Writing both cases simultaneously:

$$f(h_i | x_i) = \left[\left(\frac{1}{\sigma_0} \right) \phi \left(\frac{h_i - \beta_0 x_i}{\sigma_0} \right) \right]^{1(h_i > 0)} \left[1 - \Phi \left(\frac{\beta_0 x_i}{\sigma_0} \right) \right]^{1(h_i = 0)}$$

Log-likelihood for observation i

$$l_i(\beta, \sigma) = 1(h_i > 0) \log \left(\left(\frac{1}{\sigma} \right) \phi \left(\frac{h_i - \beta x_i}{\sigma} \right) \right) \\ + 1(h_i = 0) \log \left(1 - \Phi \left(\frac{\beta x_i}{\sigma} \right) \right)$$

Summary

- The Tobit model is like a mixture of the regression model and the Probit model
 - it is partly a Probit model because the participation decision is binary
 - it is partly a linear regression model because among those who work the hours worked can be considered continuous
- Estimating the model by OLS using those who choose to work will usually result in inconsistency because the selected sample is not iid (selection bias)
- The Tobit model can be consistently estimated by ML