Static Labor Force Participation

Chao Fu
University of Wisconsin-Madison

Beijing Summer School on Socioeconomic Inequality
2015

Static Model of Married Women's Labor Force Participation

Let

a = the age of the wife

C_a = household consumption

P_a = 1 if wife works, = 0 if wife does not work

Static Model of Married Women's Labor Force Participation

Let

a = the age of the wife

C_a = household consumption

P_a = 1 if wife works, = 0 if wife does not work

$$U_a = U_a(C_a, P_a)$$

$$U(0,0) = 0,\ U_C > 0,\ U_{CC} < 0, U(C,1) < U(C,0), \frac{\partial U}{\partial C}\bigg|_{P=0} \lesseqgtr \frac{\partial U}{\partial C}\bigg|_{P=1}$$

Marginal utility from consumption may increase/decrease with labor participation.

Static Model of Married Women's Labor Force Participation

Let

a = the age of the wife

C_a = household consumption

P_a = 1 if wife works, = 0 if wife does not work

$$U_a = U_a(C_a, P_a)$$

$$U(0,0) = 0,\ U_C > 0,\ U_{CC} < 0, U(C,1) < U(C,0), \frac{\partial U}{\partial C}\bigg|_{P=0} \lesseqgtr \frac{\partial U}{\partial C}\bigg|_{P=1}$$

Budget Constraint:

$$C_a = y_a + w_a P_a,$$

Alternative-Specific Utilities:

$$U^1 = U(y + w, 1)$$
$$U^0 = U(y, 0)$$

Alternative Specific Utilities:

$$U^1 = U(y + w, 1)$$
$$U^0 = U(y, 0)$$

Decision Rule (maximize utility subject to the budget constraint):

$$P = 1 \text{ iff } U(y + w, 1) \geqslant U(y, 0)$$
$$= 0 \text{ otherwise}$$

Alternative Specific Utilities:

$$U^1 = U(y + w, 1)$$
$$U^0 = U(y, 0)$$

Decision Rule (maximize utility subject to the budget constraint):

$$P = 1 \text{ iff } U(y + w, 1) \geqslant U(y, 0)$$
$$= 0 \text{ otherwise}$$

Participation Function:

$$P = P(y, w)$$

Are there testable implications of the model?

Are there testable implications of the model?

Assume a cross section of married couples with different ages of the wife for whom we have the following information:

$$P_{ai}, y_{ai}, w_{ai}, \quad i = 1, \ldots, I$$

Assume wage offer data are available even both for those women who work and for those who do not. Then,

Are there testable implications of the model?

Assume a cross section of *I* married couples with different ages of the wife for whom we have the following information:

$$P_{ai}, y_{ai}, w_{ai}, \quad i = 1, \ldots, I$$

Assume wage <u>offer</u> data are available even both for those women who work and for those who do not. Then,

(i) If $y_{aj} = y_{ak}$ and $w_{aj} = w_{ak}$ then $P_{aj} = P_{ak}$

Are there testable implications of the model?

Assume a cross section of *I* married couples with different ages of the wife for whom we have the following information:

$$P_{ai}, y_{ai}, w_{ai}, \quad i = 1, \ldots, I$$

Assume wage <u>offer</u> data are available even both for those women who work and for those who do not. Then,

(i) If $y_{aj} = y_{ak}$ and $w_{aj} = w_{ak}$ then $P_{aj} = P_{ak}$

(ii) If $y_{aj} = y_{ak},\ w_{aj} > w_{ak}$ then $P_{aj} \geqslant P_{ak}$

These may not be true in the data.

Why might the theory be rejected?

Why might the theory be rejected?

(i) Theory is correct, but *w*, *y*, and/or *P* are incorrectly measured.

Why might the theory be rejected?

(i) Theory is correct, but w, y, and/or P are incorrectly measured.

(ii) P depends on more than just w and y, e.g., preferences for the wife's leisure differ among households with the same y and w.

Why might the theory be rejected?

(i) Theory is correct, but w, y, and/or P are incorrectly measured.

(ii) P depends on more than just w and y, e.g., preferences for the wife's leisure differ among households with the same y and w.

(iii) The theory is wrong in a fundamental way – the assumption of static optimization is wrong, households do not optimize.

(i) Any theory can be reconciled with data if we allow for unrestricted measurement error.

Need external validation of the true values of y, w and P.

(i) Any theory can be reconciled with data if we allow for unrestricted measurement error.

Need external validation of the true values of y, w and P.

(ii) Allow for preferences to depend on other household characteristics that might affect the value of the wife's leisure, e.g., the number of young children in the household. The rejection criterion would condition on these characteristics.

But, there are likely be *unmeasured* characteristics that are related to household preferences.

How can we explicitly account for them?

One way is to explicitly account for the possibility that couples differ in the value they place on the wife's leisure in ways that the researcher cannot measure.

One way is to explicitly account for the possibility that couples differ in the value they place on the wife's leisure in ways that the researcher cannot measure.

Denoting the unobserved preference for leisure by ϵ, write

$$U = U(C, P; \epsilon(1 - P))$$

where

$$U_{\epsilon|P=0} > 0$$

Then, the participation decision is:

$$P = 1 \text{ iff } U(y+w,1) \geqslant U(y,0;\epsilon)$$

$$P = 0 \text{ otherwise}$$

Then, the participation decision is:

$$P = 1 \text{ iff } U(y+w,1) \geqslant U(y,0;\epsilon)$$

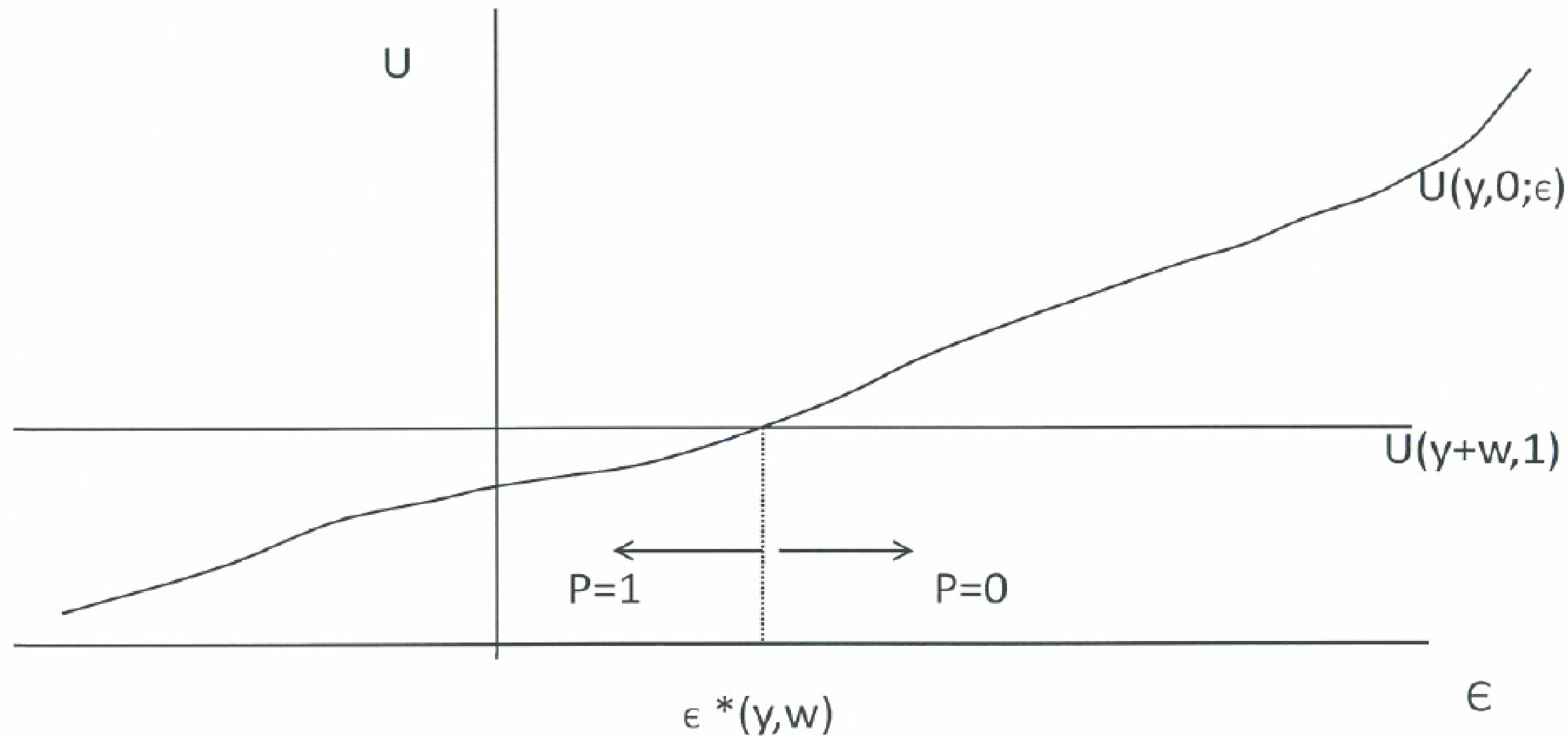
$$P = 0 \text{ otherwise}$$

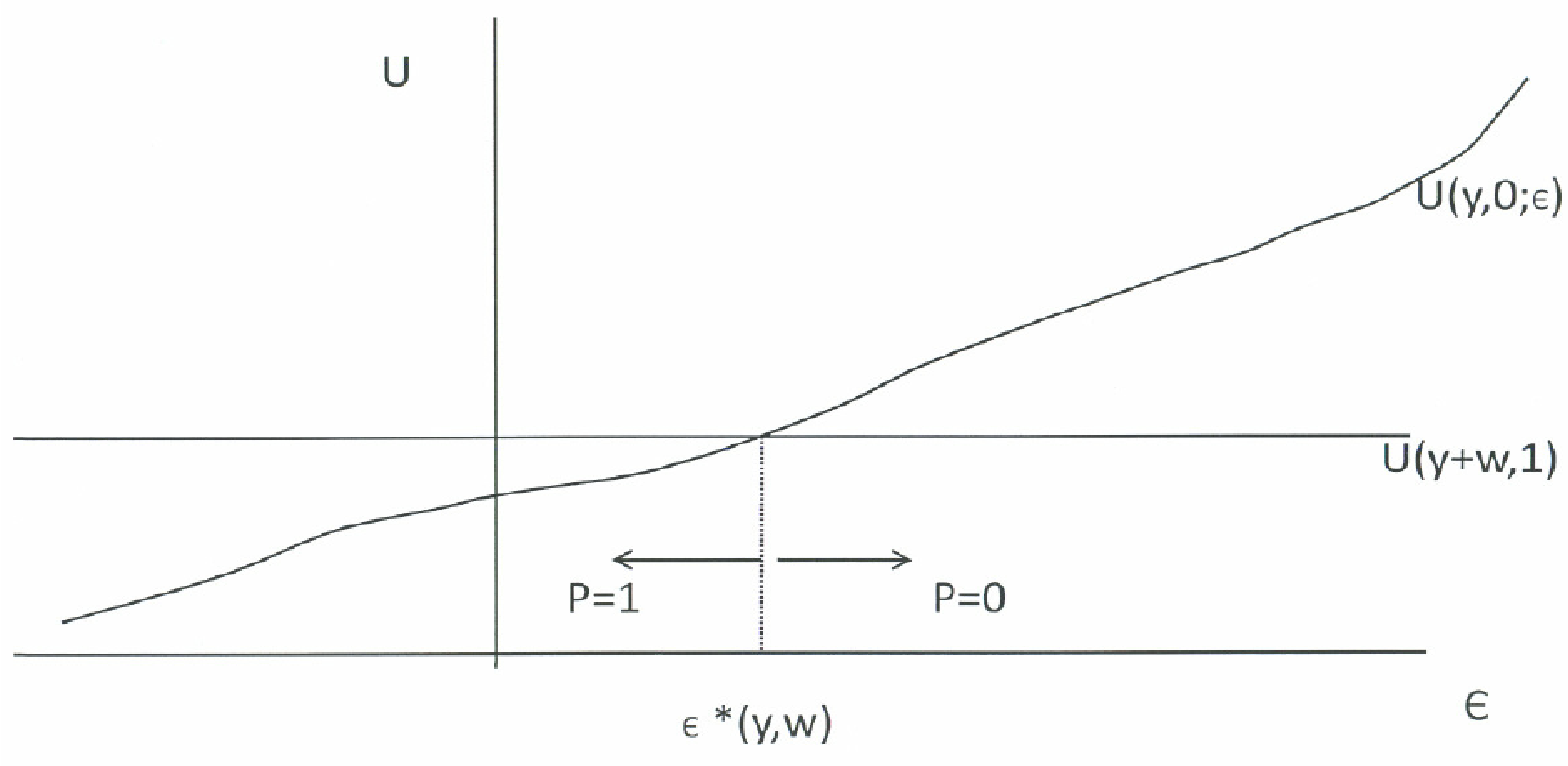
or,

$$P = 1 \text{ iff } \epsilon \leq \epsilon^*(y,w)$$

$$P = 0 \text{ otherwise}$$

where $\epsilon \leq \epsilon^*(y,w)$ is the threshold value of ϵ at which the couple is just indifferent between *P=1* and *P=0*.


U
U(y,0;ϵ)
U(y+w,1)
P=1
P=0
ϵ *(y,w)
ϵ



Properties of the cut-off value, $\epsilon^*(y,w)$:

$$\frac{\partial \epsilon^*(y,w)}{\partial w} > 0 \qquad \frac{\partial \epsilon^*(y,w)}{\partial y} \lessgtr 0$$

As y goes up, the value of wife staying at home may go up/down.

Let $\epsilon \sim F_{\epsilon|y,w}$. Then, participation is probabilistic from our perspective, though not from the household's.

$$\Pr(P = 1|y, w) = \int_{-\infty}^{\epsilon^*(y,w)} dF_{\epsilon|y,w}$$

$$= F_{\epsilon|y,w}(\epsilon^*(y, w)) = G(y, w)$$

Let $\epsilon \sim F_{\epsilon|y,w}$. Then, participation is probabilistic from our perspective, though not from the household's.

$$\Pr(P = 1|y, w) = \int_{-\infty}^{\epsilon^*(y,w)} dF_{\epsilon|y,w}$$

$$= F_{\epsilon|y,w}(\epsilon^*(y, w)) = G(y, w)$$

Note that *G(y, w)* is a composite of *F* and *U*, both of which depend on *y* and *w*. Thus, comparative static effects of *w* and *y* on the participation probability confound changes in the utility of participation and non-participation and changes in the distribution of unmeasured preferences.

Are there testable implications?

Are there testable implications?

If we assume that ϵ is distributed independently of w and y, that is, $F_{\epsilon|y,w} = F_{\epsilon}$

$$\frac{\partial \Pr(P = 1|y, w)}{\partial w} = G_w(y, w) = f_{\epsilon}(\epsilon^*(y, w))\frac{\partial \epsilon^*}{\partial w} > 0$$

$$\frac{\partial \Pr(P = 1|y, w)}{\partial y} = G_y(y, w) = f_{\epsilon}(\epsilon^*(y, w))\frac{\partial \epsilon^*}{\partial y} \lessgtr 0$$

Are there testable implications?

If we assume that ϵ is distributed independently of w and y, that is, $F_{\epsilon|y,w} = F_\epsilon$

$$\frac{\partial \Pr(P = 1|y,w)}{\partial w} = G_w(y,w) = f_\epsilon(\epsilon^*(y,w))\frac{\partial \epsilon^*}{\partial w} > 0$$

$$\frac{\partial \Pr(P = 1|y,w)}{\partial y} = G_y(y,w) = f_\epsilon(\epsilon^*(y,w))\frac{\partial \epsilon^*}{\partial y} \lesseqgtr 0$$

The model thus has the testable implications that the *probability* of working is increasing in the wage level. Testing that proposition requires estimation.

Estimation:

The participation function is thus $P(y, w; \epsilon)$. It is determined by the primitives *U* and *F* and the assumption of maximization. The participation probability function, *G(y, w)*, is itself not a primitive.

Estimation:

The participation function is thus $P(y, w; \epsilon)$. It is determined by the primitives *U* and *F* and the assumption of maximization. The participation probability function, *G(y, w)*, is itself not a primitive.

Questions:

1. What can we learn about the features of *U*, *F* and *G* given cross sectional data on *P*, *y* and *w*?

Estimation:

The participation function is thus $P(y, w; \epsilon)$. It is determined by the primitives *U* and *F* and the assumption of maximization. The participation probability function, *G(y, w)*, is itself not a primitive.

Questions:

1. What can we learn about the features of *U*, *F* and *G* given cross sectional data on *P*, *y* and *w*?

2. How does what we learn depend on *a priori* assumptions?

Approaches to Estimation – A General Taxonomy

Approaches to Estimation – A General Taxonomy

Parametric vs. Non-Parametric:

Parametric approaches impose either a functional form assumption on *U* or *G*, and/or a distributional assumption on *F*.

Non-Parametric approaches usually restrict the functions to be a member of a broad class, such as the class of continuous and differentiable functions.

Approaches to Estimation – A General Taxonomy

Parametric vs. Non-Parametric: *Parametric* approaches impose either a functional form assumption on U or G, and/or a distributional assumption on F. *Non-Parametric* approaches usually restrict the functions to be a member of a broad class, such as the class of continuous and differentiable functions.

Structural vs. Non-Structural:

Structural approaches recover the primitive functions, U and F.

Non-structural approaches recover G.

The following table describes the approaches that we will consider for the labor force participation model.

	nonparametric (NP)	parametric (P)
nonstructural (NS)	√	√
structural (S)	X	√

Note: See Matzkin (1992, Ecma) for the NP-S case.

NP-NS Approach:

NP-NS Approach

Note that

$$E(P|y,w) = \Pr(P = 1|y,w) \cdot 1 + [1 - \Pr(P = 1|y,w)] \cdot 0$$
$$= \Pr(P = 1|y,w)$$

NP-NS Approach

Note that

$$E(P|y,w) = \Pr(P = 1|y,w) \cdot 1 + [1 - \Pr(P = 1|y,w)] \cdot 0$$
$$= \Pr(P = 1|y,w)$$

Thus, $$P = E(P|y,w) + u = G(y,w) + u$$

where $E(u|y,w) = 0$ by construction (u=P-E(u|y,w)). The G function can be estimated by non-parametric regression.

NP-NS Approach

Note that

$$E(P|y,w) = \Pr(P = 1|y,w) \cdot 1 + [1 - \Pr(P = 1|y,w)] \cdot 0$$
$$= \Pr(P = 1|y,w)$$

Thus, $$P = E(P|y,w) + u = G(y,w) + u$$

where $E(u|y,w) = 0$ by construction. The G function can be estimated by non-parametric regression.

The model can be tested by checking whether

$$G_w(w,y) > 0 \text{ for all } w,y$$

P-NS Approach:

P-NS Approach

Choose a functional form for G. Examples are:

1. Linear

$$G(w,y) = \alpha_0 + \alpha_1 y + \alpha_2 w$$

which yields the linear probability model

$$P_i = \alpha_0 + \alpha_1 y_i + \alpha_2 w_i + u_i$$

P-NS Approach

Choose a functional form for G. Examples are:

1. Linear

$$G(w,y) = \alpha_0 + \alpha_1 y + \alpha_2 w$$

which yields the linear probability model

$$P_i = \alpha_0 + \alpha_1 y_i + \alpha_2 w_i + u_i$$

A test of the model is whether $\alpha_2 > 0$.

P-NS Approach

2. a. Probit

$$G = \Phi(\alpha_0 + \alpha_1 y_i + \alpha_2 w_i)$$

where Φ is the standard normal cumulative.

P-NS Approach

2. a. Probit

$$G = \Phi(\alpha_0 + \alpha_1 y_i + \alpha_2 w_i)$$

where Φ is the standard normal cumulative.

b. Logit

$$G = \Lambda(\alpha_0 + \alpha_1 y_i + \alpha_2 w_i)$$

where Λ is the logistic function $\exp(x)/(1 + \exp(x))$.

P-NS Approach

2. a. Probit

$$G = \Phi(\alpha_0 + \alpha_1 y_i + \alpha_2 w_i)$$

where Φ is the standard normal cumulative.

b. Logit

$$G = \Lambda(\alpha_0 + \alpha_1 y_i + \alpha_2 w_i)$$

where Λ is the logistic function $\exp(x)/(1 + \exp(x))$.

Denoting the parameters of the G function by Θ^G, the likelihood function is

$$L(\Theta^G; \text{data}) = \Pi_{i=1}^{I} G(y_i, w_i)^{P_i}[1 - G(y_i, w_i)]^{1-P_i}$$

P-S Approach:

P-S Approach

Denoting the parameters of the utility function by Θ^U and those of the distribution function by Θ^F,

$$L(\Theta^U, \Theta^F; \text{ data})$$

$$= \Pi_{i=1}^I F_\epsilon(\epsilon^*(y, w))^{P_i}[1 - F_\epsilon(\epsilon^*(y, w))]^{1-P_i}$$

P-S Approach

Denoting the parameters of the utility function by Θ^U and those of the distribution function by Θ^F,

$$L(\Theta^U, \Theta^F; \text{ data})$$

$$= \Pi_{i=1}^{I} F_\epsilon(\epsilon^*(y,w))^{P_i}[1 - F_\epsilon(\epsilon^*(y,w))]^{1-P_i}$$

We will consider three alternatives.

P-S Approach

1. Assume a functional form for *U*, which will imply a form for $\epsilon^*(y, w)$, but leave the distribution of *F* free.

P-S Approach

1. Assume a functional form for *U*, which will imply a form for $\epsilon^*(y, w)$, but leave the distribution of *F* free.

2. Assume a distribution for *F* and leave *U* free.

P-S Approach

1. Assume a functional form for *U*, which will imply a form for $\epsilon^*(y, w)$, but leave the distribution of *F* free.

2. Assume a distribution for *F* and leave *U* free.

3. Assume a functional form for both *U* and *F*.

P-S Approach

Cases 1 and 3: As an illustration, assume the utility function to be

$$U(C,P;\epsilon) = C + \alpha_1(1-P) + \alpha_2(1-P)C + \epsilon(1-P)$$

P-S Approach

Cases 1 and 3: As an illustration, assume the utility function to be

$$U(C, P; \epsilon) = C + \alpha_1(1-P) + \alpha_2(1-P)C + \epsilon(1-P)$$

Alternative-specific utilities:

$$U^1 = y + w$$

$$U^0 = y + \alpha_1 + \alpha_2 y + \epsilon$$

P-S Approach

Cases 1 and 3: As an illustration, assume the utility function to be

$$U(C, P; \epsilon) = C + \alpha_1(1 - P) + \alpha_2(1 - P)C + \epsilon(1 - P)$$

Alternative-specific utilities:

$$U^1 = y + w$$

$$U^0 = y + \alpha_1 + \alpha_2 y + \epsilon$$

Threshold value for $\epsilon,\ \epsilon^*(y, w)$:

$$\epsilon^*(y, w) = w - \alpha_2 y - \alpha_1$$

P-S Approach

Case 1:

L(;data) $= \Pi_{i=1}^{I} G(y_i, w_i)^{P_i} [1 - G(y_i, w_i)]^{1-P_i}$

where

$G(y_i, w_i) = F_{\epsilon}(w - \alpha_2 y - \alpha_1)$

P-S Approach

Case 3: Assume further that the preference for leisure is normally distributed $(\epsilon \sim N(0, \sigma_\epsilon^2))$. Then,

$$G(y, w) = \Phi\left(\frac{w - \alpha_2 y - \alpha_1}{\sigma_\epsilon}\right)$$

where Φ is the standard cumulative normal.

P-S Approach

Case 3: Assume further that the preference for leisure is normally distributed $(\epsilon \sim N(0, \sigma_\epsilon^2))$. Then,

$$G(y, w) = \Phi\left(\frac{w - \alpha_2 y - \alpha_1}{\sigma_\epsilon}\right)$$

where Φ is the standard cumulative normal.

What is the test of the model?

P-S Approach

Case 3: Assume further that the preference for leisure is normally distributed $(\epsilon \sim N(0, \sigma_\epsilon^2))$. Then,

$$G(y, w) = \Phi\left(\frac{w - \alpha_2 y - \alpha_1}{\sigma_\epsilon}\right)$$

where Φ is the standard cumulative normal.

A test of the model comes down to the condition that

$$\frac{\partial \Pr(P = 1|w, y)}{\partial w} = \sigma_\epsilon^{-1}\phi\left(\frac{w - \alpha_2 y - \alpha_1}{\sigma_\epsilon}\right) > 0 \text{ for all } y, w.$$

P-S Approach

Case 3: Assume further that the preference for leisure is normally distributed $(\epsilon \sim N(0, \sigma_\epsilon^2))$. Then,

$$G(y, w) = \Phi\left(\frac{w - \alpha_2 y - \alpha_1}{\sigma_\epsilon}\right)$$

where Φ is the standard cumulative normal.

A test of the model comes down to the condition that

$$\frac{\partial \Pr(P = 1 | w, y)}{\partial w} = \sigma_\epsilon^{-1} \phi\left(\frac{w - \alpha_2 y - \alpha_1}{\sigma_\epsilon}\right) > 0 \text{ for all } y, w.$$

This condition is equivalent to a test that the coefficient on the wage, σ_ϵ^{-1}, is positive.

P-S Approach

Case 2: Specify a distribution for ϵ and non-parametrically estimate $\epsilon^*(y, w)$.

For example, if $\epsilon \sim N(0, \sigma_\epsilon^2)$, it is possible to identify $\epsilon^*(y, w)$ up to scale, that is, up to a normalization of σ_ϵ.

P-S Approach

Case 2: Specify a distribution for ϵ and non-parametrically estimate $\epsilon^*(y,w)$.

For example, if $\epsilon \sim N(0,\sigma_\epsilon^2)$, it is possible to identify $\epsilon^*(y,w)$ up to scale, that is, up to a normalization of σ_ϵ.

But, non-parametric estimation $\epsilon^*(y,w)$ does not non-parametrically identify *U.*

How should we choose the estimation method, that is, one set of assumptions vs. another?

How should we choose the estimation method, that is, one set of assumptions vs. another?

What is the goal of estimation?

1. provide an understanding the determinants of the labor supply decision.
2. perform policy evaluation

How should we choose the estimation method, that is, one set of assumptions vs. another?

What is the goal of estimation?

1. provide an understanding the determinants of the labor supply decision.
2. perform policy evaluation

Conditional on achieving the goal of estimation, the fewer extra-theoretic assumptions the better. Are there tradeoffs between achieving a goal and reliance on extra-theoretic assumptions?

Ex-Ante Policy Evaluation

Ex-Ante Policy Evaluation

predicting the impact of policies that are outside the historical experience.

Ex-Ante Policy Evaluation

Ex-Ante Policy Evaluation

predicting the impact of policies that are outside the historical experience.

Is it possible within a non-parametric framework?

Ex-Ante Policy Evaluation

predicting the impact of policies that are outside the historical experience.

Is it possible within a non-parametric framework?

Consider, for example, the imposition of a wage tax, where none had existed before and suppose we non-parametrically estimate $G(y, w)$

NP-NS: Ex-Ante Policy Evaluation

A proportional wage tax changes the budget constraint:

$$C = y + (1 - \tau)wP$$
$$= y + w'P$$

What is the participation probability under the new tax regime?

NP-NS: Ex-Ante Policy Evaluation

A proportional wage tax changes the budget constraint:

$$\begin{aligned} C &= y + (1-\tau)wP \\ &= y + w'P \end{aligned}$$

The participation probability with the tax is $G(y, w')$.

NP-NS: Ex-Ante Policy Evaluation

A proportional wage tax changes the budget constraint:

$$\begin{aligned} C &= y + (1-\tau)wP \\ &= y + w'P \end{aligned}$$

The participation probability with the tax is $G(y, w')$.

Letting $h(w,y)$ be the sample density of w, y, the effect of the tax on the participation rate in the sample is

$$\int_{S_w} \int_{S_y} [G(w',y) - G(w,y)]h(w,y)dwdy$$

where S_w and S_y denote the support of w and y.

NP-NS: Ex-Ante Policy Evaluation

population tax effect = sample tax effect

only if

the sample density = population density.

NP-NS: Ex-Ante Policy Evaluation

population tax effect = sample tax effect only if the sample density = population density.

If, for example, the supports are not the same, then it will be necessary to extrapolate outside the range of support observed in the sample using a P-NS or P-S approach.

NP-NS: Ex-Ante Policy Evaluation

population tax effect = sample tax effect only if the sample density = population density.

If, for example, the supports are not the same, then it will be necessary to extrapolate outside the range of support observed in the sample using a P-NS or P-S approach.

Even if the supports are the same, there will always be values of *w* for which there are no observation with a wage of w'. For example, the minimum wage in the population has no analog in w' (The new minimum wage is lower than the minimum wage observable in the old regime).

P-S: Ex-Ante Policy Evaluation

P-S: The effect of the wage tax in the P-S approach is

$$\Phi\left(\frac{(1-\tau)w-\alpha_2 y-\alpha_1}{\sigma_\epsilon}\right)-\Phi\left(\frac{w-\alpha_2 y-\alpha_1}{\sigma_\epsilon}\right)$$

Identification of σ_ϵ is critical for performing the policy evaluation.

Ex-Ante Policy Evaluation

If one has sufficient data, non-parametric estimation of the policy effect is preferred.

The P-S approach provides a more precise estimate, but may not fit the observations well.