Dynamic Labor Force Participation

Chao Fu
University of Wisconsin-Madison

Beijing Summer School on Socioeconomic Inequality
2015

Dynamic Labor Force Participation

To fix ideas, we begin with the static labor force participation model.

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

C: consumption,

P: participation

n: number of young children

Let c be the cost of childcare if the woman works.

Let y be husband's income.

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

$$C_a = y_a + w_a P_a - cP_a n_a,$$

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

$$C_a = y_a + w_a P_a - cP_a n_a,$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a$$

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

$$C_a = y_a + w_a P_a - cP_a n_a,$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a$$

s: years of schooling,

a: age

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

$$C_a = y_a + w_a P_a - cP_a n_a,$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a$$

$$\xi_a \sim N(0, \sigma_\xi^2)$$

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

$$C_a = y_a + w_a P_a - cP_a n_a,$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a$$

$$\xi_a \sim N(0, \sigma_\xi^2)$$

Household's state space: $\Omega_a = \{y_a, n_a, s, a, \xi_a\}$

To fix ideas, we begin with the static labor force participation model.

$$U_a(C_a, P_a; x_a) = C_a + \alpha_1(1 - P_a) + \alpha_2 C_a(1 - P_a) + \alpha_3(1 - P_a)n_a$$

$$C_a = y_a + w_a P_a - cP_a n_a,$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a$$

$$\xi_a \sim N(0, \sigma_\xi^2)$$

Household's state space:

$$\Omega_a = \{ y_a, n_a, s, a, \xi_a \}$$

State space observable to the researcher:

$$\Omega_a^- = \{ y_a, n_a, s, a \}$$

Alternative-Specific Utilities:

$$U^1 = y_a - cn_a + \exp(\gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a),$$

$$U^0 = y_a + \alpha_1 + \alpha_2 y_a + \alpha_3 n_a.$$

Alternative-Specific Utilities:

$$U^1 = y_a - cn_a + \exp(\gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a),$$

$$U^0 = y_a + \alpha_1 + \alpha_2 y_a + \alpha_3 n_a.$$

Decision Rule:

P_a = 1 if $\xi_a \geq \log(\alpha_1 + \alpha_2 y_a + (\alpha_3 + c)n_a) - \gamma_0 - \gamma_1 s - \gamma_2 a + \gamma_3 a^2 = \xi_a^*(\Omega_a^-)$

= 0, otherwise.

Alternative-Specific Utilities:

$$U^1 = y_a - cn_a + \exp(\gamma_0 + \gamma_1 s + \gamma_2 a - \gamma_3 a^2 + \xi_a),$$

$$U^0 = y_a + \alpha_1 + \alpha_2 y_a + \alpha_3 n_a.$$

Decision Rule:

$$P_a = 1 \text{ if } \xi_a \geq \log(\alpha_1 + \alpha_2 y_a + (\alpha_3 + c)n_a) - \gamma_0 - \gamma_1 s - \gamma_2 a + \gamma_3 a^2 = \xi_a^*(\Omega_a^-)$$

$$= 0, \text{ otherwise.}$$

The woman works if, given her age, education, husband's income and number of children, her ξ_a exceeds the critical value $\xi_a^*(\Omega_a^-)$.

Data: Cross-section of $i = 1, \ldots, I$ married women

$$P_{ia}, w_{ia}P_{ia},\ y_{ia}, n_{ia}, s_i \text{ and } a_i$$

Data: Cross-section of $i = 1,\ldots,I$ married women

$$P_{ia}, w_{ia}P_{ia},\ y_{ia}, n_{ia}, s_i \text{ and } a_i$$

Model parameters:

$$\Theta = \{\alpha_1, \alpha_2,\ \alpha_3,\ c,\ \gamma_0,\ \gamma_1, \gamma_2, \gamma_3, \sigma_\xi^2\}$$

Data: Cross-section of $i = 1, \ldots, I$ married women

$$P_{ia}, w_{ia}P_{ia},\ y_{ia}, n_{ia}, s_i \text{ and } a_i$$

Model parameters:

$$\Theta = \{\alpha_1, \alpha_2,\ \alpha_3,\ c,\ \gamma_0,\ \gamma_1, \gamma_2, \gamma_3, \sigma_\xi^2\}$$

Likelihood function:

$$L(\Theta; data) = \Pi_{i=1}^{I}[\Pr(P_{ia} = 1, w_{ia}|\Omega_{ia}^{-})]^{P_{ia}}[\Pr(P_{ia} = 0)|\Omega_{ia}^{-})]^{1-P_{ia}}$$

Rewrite

$$\Pr(P_{ia} = 1, w_{ia} | \Omega_{ia}^{-}) = \Pr(P_{ia} = 1 | w_{ia}, \Omega_{ia}^{-}) \Pr(w_{ia} | \Omega_{ia}^{-}).$$

Rewrite

$$\Pr(P_{ia} = 1, w_{ia}|\Omega^-_{ia}) = \Pr(P_{ia} = 1|w_{ia}, \Omega^-_{ia}) \Pr(w_{ia}|\Omega^-_{ia}).$$

The first term, the probability of working conditional on the wage, is

$$\Pr(w_{ia} > \alpha_1 + \alpha_2 y_{ia} + (\alpha_3 + c) n_{ia}|w_{ia}, \Omega^-_{ia})$$

Rewrite

$$\Pr(P_{ia} = 1, w_{ia}|\Omega_{ia}^{-}) = \Pr(P_{ia} = 1|w_{ia}, \Omega_{ia}^{-}) \Pr(w_{ia}|\Omega_{ia}^{-}).$$

The first term, the probability of working conditional on the wage, is

$$\Pr(w_{ia} > \alpha_1 + \alpha_2 y_{ia} + (\alpha_3 + c) n_{ia}|w_{ia}, \Omega_{ia}^{-})$$

What values can this probability take?

Rewrite

$$\Pr(P_{ia} = 1, w_{ia} | \Omega_{ia}^{-}) = \Pr(P_{ia} = 1 | w_{ia}, \Omega_{ia}^{-}) \Pr(w_{ia} | \Omega_{ia}^{-}).$$

The first term, the probability of working conditional on the wage, is

$$\Pr(w_{ia} > \alpha_1 + \alpha_2 y_{ia} + (\alpha_3 + c) n_{ia} | w_{ia}, \Omega_{ia}^{-})$$

What values can this probability take?

Only 0 or 1. Why?

Either the condition inside the probability statement holds or it does not hold.

What happens to the likelihood?

Either the condition inside the probability statement holds or it does not hold.

The likelihood will not be degenerate (zero) only if there exists a set of parameters such that this condition is not violated (even for a single observation).

Either the condition inside the probability statement holds or it does not hold.

The likelihood will not be degenerate (zero) only if there exists a set of parameters such that this condition is not violated (even for a single observation).

Thus, the parameters must satisfy the set of restrictions that

$$w_{ia} > \alpha_1 + \alpha_2 y_{ia} + (\alpha_3 + c) n_{ia}$$ for all *i*.

Either the condition inside the probability statement holds or it does not hold.

The likelihood will not be degenerate (zero) only if there exists a set of parameters such that this condition is not violated (even for a single observation).

Thus, the parameters must satisfy the set of restrictions that

$$w_{ia} > \alpha_1 + \alpha_2 y_{ia} + (\alpha_3 + c) n_{ia} \quad \text{for all } i.$$

In particular, it must be satisfied for the person with the lowest observed wage. Clearly that observation will have an extreme effect on the parameter estimates.

How can we avoid this?

How can we avoid this?

1. Add an additional error term to the decision model.

For example, we can assume that α_1 differs across people, i.e., replace α_1 with $\alpha_{1i} = \alpha_1 + \epsilon_i$

How can we avoid this?

1. Add an additional error term to the decision model.

   For example, we can assume that α_1 differs across people, i.e., replace α_1 with $\alpha_{1i} = \alpha_1 + \epsilon_i$

   What complication does this raise?

How can we avoid this?

1. Add an additional error term to the decision model.

   For example, we can assume that α_1 differs across people, i.e., replace α_1 with $\alpha_{1i} = \alpha_1 + \epsilon_i$

   What complication does this raise?

   We lose the analytical solution for the cutoff value.

$$\xi_a - \log(\alpha_1 + \epsilon_i + \alpha_2 y_a + (\alpha_3 + c) n_a) - \gamma_0 - \gamma_1 s - \gamma_2 a + \gamma_3 a^2 \lesseqgtr 0$$

2. Assume that the wage data are measured with error, for example, that the observed (reported) wage measures the true wage with a proportionate error,

$$\log w_{ia}^{o} = \log w_{ia} + \eta_{ia},$$

where

$\eta_{ia} \sim N(0, \sigma_{\eta}^{2})$ and $E(\xi_{ia}\eta_{ia}) = 0.$

2. Assume that the wage data are measured with error, for example, that the observed (reported) wage measures the true wage with a proportionate error,

$$\log w_{ia}^{o} = \log w_{ia} + \eta_{ia},$$

where

$\eta_{ia} \sim N(0, \sigma_{\eta}^{2})$ and $E(\xi_{ia}\eta_{ia}) = 0.$

What advantage does this have over the previous method?

2. Assume that the wage data are measured with error, for example, that the observed (reported) wage measures the true wage with a proportionate error,

$$\log w_{ia}^{o} = \log w_{ia} + \eta_{ia},$$

where

$$\eta_{ia} \sim N(0, \sigma_{\eta}^{2}) \text{ and } E(\xi_{ia}\eta_{ia}) = 0.$$

What advantage does this have over the previous method?

The decision rule is unchanged, so that the cutoff value has an analytical form. The parameter restriction that we had before only holds with respect to the true wage, not with respect to reported wages.

How does the likelihood function change?

$$\Pr(P_{ia} = 1, w_{ia}^o | \Omega_{ia}^-)$$

$$= (w_{ia}^o)^{-1} \Pr(\xi_{ia} \geq \xi_{ia}^*(\Omega_{ia}^-), u_{ia} = \log w_{ia}^o - (\gamma_0 + \gamma_1 s_i + \gamma_2 a_i - \gamma_3 a_i^2))$$

where $u_{ia} = \xi_{ia} + \eta_{ia}$

How does the likelihood function change?

$$\Pr(P_{ia} = 1, w_{ia}^o | \Omega_{ia}^-)$$

$$= (w_{ia}^o)^{-1} \Pr(\xi_{ia} \geq \xi_{ia}^*(\Omega_{ia}^-), u_{ia} = \log w_{ia}^o - (\gamma_0 + \gamma_1 s_i + \gamma_2 a_i - \gamma_3 a_i^2))$$

where $u_{ia} = \xi_{ia} + \eta_{ia}$

Given normality, this becomes

How does the likelihood function change?

$$\Pr(P_{ia} = 1, w_{ia}^o | \Omega_{ia}^-)$$

$$= (w_{ia}^o)^{-1} \Pr(\xi_{ia} \geq \xi_{ia}^*(\Omega_{ia}^-), u_{ia} = \log w_{ia}^o - (\gamma_0 + \gamma_1 s_i + \gamma_2 a_i - \gamma_3 a_i^2))$$

where $u_{ia} = \xi_{ia} + \eta_{ia}$.

Given normality, this becomes

$$\Pr(P_{ia} = 1, w_{ia}^o | \Omega_{ia}^-) = (1 - \Phi(\frac{\xi_{ia}^*(\Omega_{ia}^-) - \rho \frac{\sigma_\xi}{\sigma_u} u_{ia}}{\sigma_\xi \sqrt{1 - \rho^2}})\frac{1}{\sigma_u}\phi(\frac{u_{ia}}{\sigma_{u_{ia}}}),$$

4

$$\Pr(P_{ia} = 1, w_{ia}^o | \Omega_{ia}^-) = (1 - \Phi(\frac{\xi_{ia}^*(\Omega_{ia}^-) - \rho\frac{\sigma_\xi}{\sigma_u}u_{ia}}{\sigma_\xi\sqrt{1-\rho^2}}))\frac{1}{\sigma_u}\phi(\frac{u_{ia}}{\sigma_{u_{ia}}}),$$

where $\sigma_u = \sqrt{\sigma_\xi^2 + \sigma_u^2}$, $\rho = \sigma_\xi/\sigma_u$ and $1 - \rho^2$ is the fraction of the variance in the wage due to measurement error.

$$\Pr(P_{ia} = 1, w_{ia}^o | \Omega_{ia}^-) = (1 - \Phi(\frac{\xi_{ia}^*(\Omega_{ia}^-) - \rho \frac{\sigma_\xi}{\sigma_u} u_{ia}}{\sigma_\xi \sqrt{1 - \rho^2}})\frac{1}{\sigma_u}\phi(\frac{u_{ia}}{\sigma_{u_{ia}}}),$$

where $\sigma_u = \sqrt{\sigma_\xi^2 + \sigma_u^2}$, $\rho = \sigma_\xi / \sigma_u$ and $1 - \rho^2$ is the fraction of the variance in the wage due to measurement error.

The component of the likelihood function for non-workers is

$$\Pr(P_a = 0) = \Phi(\xi_{ia}^* / \sigma_\xi)$$

9

Dynamics:

How can we introduce dynamics into the model?

Dynamics:

How can we introduce dynamics into the model?

Generally, by allowing history to matter.

Dynamics:

How can we introduce dynamics into the model?

Generally, by allowing history to matter.

For example, consider modifying the wage function to include a measure of work experience.

Dynamics:

How can we introduce dynamics into the model?

Generally, by allowing history to matter.

For example, consider modifying the wage function to include a measure of work experience.

Specifically, assume that the wage a woman receives at age a depends on the number of periods the woman has worked prior to age *a*. The woman enters age *a* with K_{a-1} periods of work experience and faces the wage offer function

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

Given work experience, age is assumed to have no effect on wage offers.

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

Given work experience, age is assumed to have no effect on wage offers.

Assume that wage offers are serially independent, i.e., $E(\xi_a \xi_{a-1})$ = *0*.

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

Given work experience, age is assumed to have no effect on wage offers.

Assume that wage offers are serially independent, i.e., $E(\xi_a \xi_{a-1})$ = *0*.

Work experience evolves according to

$$K_a = \sum_{\tau=1}^{a} P_\tau = K_{a-1} + P_a$$

with K_0 = *0.*

Given that wage offers depend on prior work decisions, working at any age must also affect future wage offers.

Given that wage offers depend on prior work decisions, working at any age must also affect future wage offers.

A forward looking household would therefore take this effect into account in making any current work decision. We assume that the household is forward looking.

Given that wage offers depend on prior work decisions, working at any age must also affect future wage offers.

A forward looking household would therefore take this effect into account in making any current work decision. We assume that the household is forward looking.

Augmenting the state space to include current work experience,

$$\Omega_a = \{a, K_{a-1}, y_a, n_a, s, \xi_a\}$$

$$\Omega_a^- = \{a, K_{a-1}, y_a, n_a, s\}$$

Denoting $V_a(\Omega_a)$ as the maximum expected present discounted value of remaining lifetime utility at *a* given the state space and discount factor

$$V_a(\Omega_a) = \max_{P_a} E\left\{\sum_{\tau=a}^{A} \delta^{\tau-a}[U_\tau^1 P_\tau + U_\tau^0(1 - P_\tau]|\Omega_a\right\}$$

where A is the last decision age.

The value function, that is, the expected present discounted value of lifetime utility, can be written as the maximum over the two alternative-specific value functions, $V_t^k(\Omega_{it}),\ k \in \{0, 1\}$

$$V_a(\Omega_a) = \max(V_a^0(\Omega_a),\ V_a^1(\Omega_a)),$$

The value function, that is, the expected present discounted value of lifetime utility, can be written as the maximum over the two alternative-specific value functions, $V_t^k(\Omega_{it}), k \in \{0, 1\}$

$$V_a(\Omega_a) = \max(V_a^0(\Omega_a), V_a^1(\Omega_a)),$$

Each of the alternative-specific value functions obeys a Bellman equation

$$V_a^k(\Omega_a) = E(U_a^k \mid \Omega_a) + \delta E[V_{a+1}(\Omega_{a+1}) | \Omega_a, P_a = k] \text{ for } a < A$$

$$= E(U_A^k | \Omega_A) \text{ for } a = A.$$

Given the finite horizon, the dynamic programming problem can be solved recursively beginning at age A.

Given the finite horizon, the dynamic programming problem can be solved recursively beginning at age A.

But first we need to complete the model.

Given the finite horizon, the dynamic programming problem can be solved recursively beginning at age A.

But first we need to complete the model.

We need to make an assumption about how children and husband's income evolve.

Children:

In empirical applications n_a often represents the number of young children, say under the age of 6. To include that as a state variable in the model, we need to keep track of all births within the five years preceding each age.

Children:

In empirical applications n_a often represents the number of young children, say under the age of 6. To include that as a state variable in the model, we need to keep track of all births within the five years preceding each age.

If a period is a year and births may occur one year apart, then in order to update the number of children under the age of 6, one needs to keep track of all of the different ways to have had one child within a five year period plus all of the ways to have had two children in a five year period, etc.

In addition, we need to specify the probability that a child is born at each age.

In addition, we need to specify the probability that a child is born at each age.

One possibility would be to assume that the probability of having a child is a function of (a subset) of the state variables in the model, including for example, the history of prior births, the woman’s age, etc.

In addition, we need to specify the probability that a child is born at each age.

One possibility would be to assume that the probability of having a child is a function of (a subset) of the state variables in the model, including for example, the history of prior births, the woman's age, etc.

Another possibility would be to model the birth of a child as another choice.

In addition, we need to specify the probability that a child is born at each age.

One possibility would be to assume that the probability of having a child is a function of (a subset) of the state variables in the model, including for example, the history of prior births, the woman’s age, etc.

Another possibility would be to model the birth of a child as another choice.

We will circumvent these issues by modeling the labor force participation of women after 6 years after they become infecund, say from age 50 on.

Husband's income:

Assume that husband's income is given by

$$\log y_a = \beta_0 + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h + \epsilon_a,$$

where a_{ha} is the husband's age at the wife's age *a*, s_h is the husband's schooling and ϵ_a is a serially correlated normally distributed shock.

Husband's income:

Assume that husband's income is given by

$$\log y_a = \beta_0 + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h + \epsilon_a,$$

where a_{ha} is the husband's age at the wife's age *a*, s_h is the husband's schooling and ϵ_a is a serially uncorrelated normally distributed shock.

What should we assume about whether ϵ_a is in the information set at a?

We will assume that ϵ_a is *not* in the information set at a, for computational reasons to be discussed later.

Given that it is not in the information set, for $a' \leq a$

$$E_{a'}(y_a) = \ ?$$

We will assume that ϵ_a is *not* in the information set at a, for computational reasons to be discussed later.

Given that it is not in the information set, for $a' \leq a$

$$E_{a'}(y_a) = \exp(\beta_0 + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h)\exp(\tfrac{1}{2}\sigma_\epsilon^2)$$

$$= \overline{y}_a$$

Returning to the Bellman equations:

$$V_a^k(\Omega_a) = E(U_a^k \mid \Omega_a) + \delta E[V_{a+1}(\Omega_{a+1}) \mid \Omega_a, P_a = k] \text{ for } a < A$$

$$= E(U_A^k \mid \Omega_A) \text{ for } a = A.$$

The expectation of the future component of the value function is taken over the distribution of ξ_{a+1} conditional on the state space elements at age a.

We have assumed that

$$f(\xi_{a+1} \mid \xi_a, K_{a-1}, n_a, s) = f(\xi_{a+1})$$

is normal.

Given that we are taking the initial decision period to be *a=50*, the husband's age is given by

$$a_{ha} = a_{h,50} + (a - 50)$$

The state space now is:

$$\Omega_a = \{a, K_{a-1}, s, a_{h,50}, s_h, \xi\}$$

$$\Omega_a^- = \{a, K_{a-1}, s, a_{h,50}, s_h\}$$

Value Functions at A:

$$\begin{aligned} V_A^1(\Omega_A) &= E_A U_A^1(K_{A-1}, y_A, s, \xi_A) \\ &= \overline{y}_A + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2 + \xi_A), \end{aligned}$$

Value Functions at A:

$$
\begin{aligned}
V_A^1(\Omega_A) &= E_A U_A^1(K_{A-1}, y_A, s, \xi_A) \\
&= \overline{y}_A + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2 + \xi_A),
\end{aligned}
$$

$$
V_A^0(\Omega_A) = E_A U_A^0(y_A)
$$

$$
= \overline{y}_A + \alpha_1 + \alpha_2 \overline{y}_A.
$$

Value Functions at A:

$$V_A^1(\Omega_A) = E_A U_A^1(K_{A-1}, y_A, s, \xi_A)$$
$$= \overline{y}_A + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2 + \xi_A),$$

$$V_A^0(\Omega_A) = E_A U_A^0(y_A)$$
$$= \overline{y}_A + \alpha_1 + \alpha_2 \overline{y}_A.$$

$$V_A(\Omega_A) = \max(\overline{y}_A + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2 + \xi_A),$$
$$\overline{y}_A(1 + \alpha_2) + \alpha_1).$$

Decision Rule at A:

$$P_A(\Omega_A) = 1$$

if $\xi_A \geq \log(\alpha_1 + \alpha_2 \overline{y}_A) - \gamma_0 - \gamma_1 s - \gamma_2 K_{A-1} + \gamma_3 K_{A-1}^2$

Decision Rule at A:

$$P_A(\Omega_A) = 1$$

if $\xi_A \geq \log(\alpha_1 + \alpha_2 \overline{y}_A) - \gamma_0 - \gamma_1 s - \gamma_2 K_{A-1} + \gamma_3 K_{A-1}^2$

$$= \xi_A^{**}(\Omega_A^-)$$

Decision Rule at A:

$$P_A(\Omega_A) = 1$$

$$\text{if } \xi_A \geq \log(\alpha_1 + \alpha_2 \overline{y}_A) - \gamma_0 - \gamma_1 s - \gamma_2 K_{A-1} + \gamma_3 K_{A-1}^2$$
$$= \xi_A^{**}(\Omega_A^-)$$

$= 0$ otherwise

To solve for the decision rule at *A-1,* we will need to calculate the expectation of $V_A(\Omega_A)$ given Ω_{A-1}, $E_{A-1}V_A(\Omega_A)$.

To solve for the decision rule at *A-1,* we will need to calculate the expectation of $V_A(\Omega_A)$ given Ω_{A-1}, $E_{A-1}V_A(\Omega_A)$.

Taking the expectation of the maximum function

$$E_{A-1}V_A(\Omega_A) =$$

To solve for the decision rule at *A-1,* we will need to calculate the expectation of $V_A(\Omega_A)$ given $\Omega_{A-1},\ E_{A-1}V_A(\Omega_A)$.

Taking the expectation of the maximum function

$$E_{A-1}V_A(\Omega_A) =$$

$$(\overline{y}_A + E(\exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2 + \xi_A | \xi_A \geq \xi_A^{**}))) \bullet \Pr(\xi_A \geq \xi_A^{**})$$

$$+ (\overline{y}_A(1 + \alpha_2) + \alpha_1) \bullet \Pr(\xi_A < \xi_A^{**})$$

To solve for the decision rule at *A-1*, we will need to calculate the expectation of $V_A(\Omega_A)$ given Ω_{A-1}, $E_{A-1}V_A(\Omega_A)$.

Taking the expectation of the maximum function

$$E_{A-1}V_A(\Omega_A) =$$

$$(\overline{y}_A + E(\exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2 + \xi_A | \xi_A \geq \xi_A^{**}))) \cdot \Pr(\xi_A \geq \xi_A^{**})$$

$$+ (\overline{y}_A(1+\alpha_2) + \alpha_1) \cdot \Pr(\xi_A < \xi_A^{**})$$

$$= \overline{y}_A \Pr(\xi_A \geq \xi_A^{**})$$

$$+ \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2) E_{A-1}(e^{\xi_A}) | e^{\xi_A} \geq e^{\xi_A^{**}})) \Pr(e^{\xi_A} \geq e^{\xi_A^{**}})$$

$$+ ((1+\alpha_2)\overline{y}_A + \alpha_1) \cdot \Pr(\xi_A < \xi_A^{**})$$

Using the fact that with normality,

$$E(e^{\xi_A})|e^{\xi_A} \geq e^{\xi_A^{**}})) \Pr(e^{\xi_A} \geq e^{\xi_A^{**}}) = e^{0.5\sigma_\xi^2}[1 - \Phi(\frac{\xi_A^{**} - \sigma_\xi^2}{\sigma_\xi})],$$

Using the fact that with normality,

$$E(e^{\xi_A})|e^{\xi_A} \geq e^{\xi_A^{**}})) \Pr(e^{\xi_A} \geq e^{\xi_A^{**}}) = e^{0.5\sigma_\xi^2}[1 - \Phi(\frac{\xi_A^{**} - \sigma_\xi^2}{\sigma_\xi})],$$

and letting

$$\exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2) = X_A$$

Using the fact that with normality,

$$E(e^{\xi_A})|e^{\xi_A} \geq e^{\xi_A^{**}})) \Pr(e^{\xi_A} \geq e^{\xi_A^{**}}) = e^{0.5\sigma_\xi^2}[1 - \Phi(\frac{\xi_A^{**} - \sigma_\xi^2}{\sigma_\xi})],$$

and letting

$$\exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-1} - \gamma_3 K_{A-1}^2) = X_A$$

we get

$$E_{A-1} V_A(\Omega_A) = \overline{y}_A[1 - \Phi(\frac{\xi_A^{**}}{\sigma_\xi})] + X_A\, e^{0.5\sigma_\xi^2}[1 - \Phi(\frac{\xi_A^{**} - \sigma_\xi^2}{\sigma_\xi})]$$

$$+ ((1 + \alpha_2)\overline{y}_A + \alpha_1)\Phi(\frac{\xi_A^{**}}{\sigma_\xi^2})$$

Recall that ξ_A^{**} is a function of Ω_A^-. So, define

$$E_{A-1}V_A(\Omega_A) = \overline{y}_A[1-\Phi(\frac{\xi_A^{**}}{\sigma_\xi})] + X_A e^{0.5\sigma_\xi^2}[1-\Phi(\frac{\xi_A^{**}-\sigma_\xi^2}{\sigma_\xi})]$$

$$+((1+\alpha_2)\overline{y}_A+\alpha_1)\Phi(\frac{\xi_A^{**}}{\sigma_\xi^2})$$

$$= E\max(\Omega_A^-).$$

Having solved for $E_{A-1}V_A(\Omega_A)$ = $E\max(\Omega_A^-)$, the alternative-specific value functions for the previous period, A-1 are:

Having solved for $E_{A-1}V_A(\Omega_A) = E\max(\Omega_A^-)$, the alternative-specific value functions for the previous period, A-1 are

$$V_{A-1}^1(\Omega_{A-1}) = \overline{y}_{A-1} + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-2} - \gamma_3 K_{A-2}^2 + \xi_{A-1}) + \delta E V_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 1),$$

.

Having solved for $E_{A-1}V_A(\Omega_A) = E\max(\Omega_A^-)$, the alternative-specific value functions for the previous period, A-1 are

$$V_{A-1}^1(\Omega_{A-1}) = \overline{y}_{A-1} + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-2} - \gamma_3 K_{A-2}^2 + \xi_{A-1})$$

$$+ \delta EV_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 1),$$

$$V_{A-1}^0(\Omega_{A-1}) = \overline{y}_{A-1} + \alpha_1 + \alpha_2 \overline{y}_{A-1}$$

$$+ \delta EV_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 0)$$

Having solved for $E_{A-1}V_A(\Omega_A) = E\max(\Omega_A^-)$, the alternative-specific value functions for the previous period, A-1 are

$$V_{A-1}^1(\Omega_{A-1}) = \overline{y}_{A-1} + \exp(\gamma_0 + \gamma_1 s + \gamma_2 K_{A-2} - \gamma_3 K_{A-2}^2 + \xi_{A-1})$$

$$+ \delta EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 1),$$

$$V_{A-1}^0(\Omega_{A-1}) = \overline{y}_{A-1} + \alpha_1 + \alpha_2 \overline{y}_{A-1}$$

$$+ \delta EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 0)$$

Note that $EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 1)$ and $EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 0)$ are simply specific values of $E\max(\Omega_A^-)$ that have already been calculated.

Decision Rule:

$P_{A-1} = 1$ if

$$\xi_{A-1} \geq \log\{\alpha_1 + \alpha_2 \overline{y}_{A-1}$$

$$+ \delta[EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 0) - EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 1)]\}$$

$$- \gamma_0 - \gamma_1 s - \gamma_2 K_{A-2} + \gamma_3 K^2_{A-2}$$

Decision Rule:

$P_{A-1} = 1$ if

$$\xi_{A-1} \geq \log\{\alpha_1 + \alpha_2 \overline{y}_{A-1}$$

$$+ \delta[EV_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 0) - EV_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 1)]\}$$

$$- \gamma_0 - \gamma_1 s - \gamma_2 K_{A-2} + \gamma_3 K_{A-2}^2$$

$$= \xi_{A-1}^{**}(\Omega_{A-1}^{-})$$

Decision Rule:

$P_{A-1} = 1$ if

$$\xi_{A-1} \geq \log\{\alpha_1 + \alpha_2 \overline{y}_{A-1}$$

$$+ \delta[EV_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 0) - EV_A(\Omega_A | \Omega_{A-1}, P_{A-1} = 1)]\}$$

$$- \gamma_0 - \gamma_1 s - \gamma_2 K_{A-2} + \gamma_3 K_{A-2}^2$$

$$= \xi_{A-1}^{**}(\Omega_{A-1}^{-})$$

$= 0$ otherwise

Decision Rule:

$P_{A-1} = 1$ if

$$\xi_{A-1} \geq \log\{\alpha_1 + \alpha_2 \overline{y}_{A-1}$$

$$+ \delta[EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 0) - EV_A(\Omega_A|\Omega_{A-1}, P_{A-1} = 1)]\}$$

$$- \gamma_0 - \gamma_1 s - \gamma_2 K_{A-2} + \gamma_3 K_{A-2}^2$$

$$= \xi_{A-1}^{**}(\Omega_{A-1}^{-})$$

$= 0$ otherwise

or

$P_{A-1} = 1$ if $\xi_{A-1} \geq \xi_{A-1}^{**}(\Omega_{A-1}^{-})$, $= 0$ otherwise

We can continue to solve backwards, calculating $E\max\left(\Omega_a^-\right)$ and thus $\xi_a^*(\Omega_a^-)$ until we reach age 50. If we continued beyond that we would have to be explicit about how to model fertility.

Given $\xi_a^*(\Omega_a^-)$, we can determine the decision rule:

$$P_a = 1 \text{ if } \xi_a \geq \xi_a^{**}\left(\Omega_a^-\right),$$

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

Suppose we have longitudinal data on a sample of $i = 1, \ldots, I$ married women starting from age 50 to some final age, a_{iT_i}, which may or may not be the terminal age, *A*.

Suppose we have longitudinal data on a sample of $i = 1, \ldots, I$ married women starting from age 50 to some final age, a_{iT_i}, which may or may not be the terminal age, *A*.

We observe whether or not the woman works in each period, the wage in periods that she works, and all of the state variables at age 50, the woman's schooling, her husband's schooling and age and the number of prior periods the woman has worked. These are the initial conditions for the problem *given data we have available*.

Suppose we have longitudinal data on a sample of $i = 1, \ldots, I$ married women starting from age 50 to some final age, a_{iT_i}, which may or may not be the terminal age, *A*.

We observe whether or not the woman works in each period, the wage in periods that she works, and all of the state variables at age 50, the woman's schooling, her husband's schooling and age and the number of prior periods the woman has worked. These are the initial conditions for the problem *given data we have available*.

The likelihood contribution for household *i* is the probability of the sequence of participation decisions and wages for working periods over the observation periods.

The likelihood contribution for household *i* is the probability of the sequence of participation decisions and wages for working periods over the observation periods.

As in the static model, we need an additional error to avoid either degeneracy or extreme outlier effects. Assume that wages are measured with the same error structure as before:

$$\log w^o_{a_{it}} = \log w_{a_{it}} + \eta_{a_{it}};$$

$$\eta_{a_{it}} \sim N(0, \sigma^2_\eta)$$

Define an outcome in period $a_{it},\ O_{a_{it}}$ to be the pair $(P_{a_{it},}, w^o_{a_{it}})$ if the woman works and $P_{a_{it},}$ if the woman does not work. Then, the likelihood function is

$$L(\Theta; data) = \Pi_{i=1}^{i=I} \Pr(O_{a_{iT_i}}, O_{a_{iT_i-1}}, ..., O_{a_{i1}} | \Omega_{50}^{-}) \Pr(\Omega_{50}^{-})$$

where

$$\Theta = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2, \beta_3, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \sigma_\xi^2, \sigma_\eta^2, \sigma_\epsilon^2, \delta\}.$$

Given the assumption that the wage shocks are iid, for the purpose of estimation we can treat the predetermined state variable in Ω_{50}^{-} , work experience, as exogenous.

We can thus consistently estimate the parameters of the model dropping the marginal probability of the state space, that is, we can maximize the likelihood function

$$L(\Theta; data) = \Pi_{i=1}^{i=I} \Pr(O_{a_{iT_i}}, O_{a_{iT_i-1}}, ..., O_{a_{i1}} | \Omega_{50}^{-})$$

We can further write the likelihood function as the product of age-specific conditional probabilities, where each probability is conditioned on the relevant state space for that decision period:

$$L(\Theta; data) =$$

$$\Pi_{i=1}^{i=I} \Pr(O_{a_{iT_i}} | \Omega^{-}_{a_{iT_i}}) \Pr(O_{a_{iT_i-1}} | \Omega^{-}_{a_{iT_i-1}}) \cdots \Pr(O_{50} | \Omega^{-}_{50})$$

These conditional probabilities have the same representation as in the static model, except that the cut-off values of the wage error (ξ) that determine participation are the ones that solve the dynamic programming problem, that is,

These conditional probabilities have the same representation as in the static model, except that the cut-off values of the wage error (ξ) that determine participation are the ones that solve the dynamic programming problem, that is,

$$\Pr(P_{a_{it}} = 1, w^o_{a_{it}} | \Omega^-_{a_{it}}) = (1 - \Phi(\frac{\xi^{**}_{a_{it}}(\Omega^-_{a_{it}}) - \rho \frac{\sigma_\xi}{\sigma_u} u_{a_{it}}}{\sigma_\xi \sqrt{1-\rho^2}})) \frac{1}{\sigma_u} \phi(\frac{u_{a_{it}}}{\sigma_{u_{a_{it}}}}),$$

These conditional probabilities have the same representation as in the static model, except that the cut-off values of the wage error (ξ) that determine participation are the ones that solve the dynamic programming problem, that is,

$$\Pr(P_{a_{it}} = 1, w^{o}_{a_{it}} | \Omega^{-}_{a_{it}}) = (1 - \Phi(\frac{\xi^{**}_{a_{it}}(\Omega^{-}_{a_{it}}) - \rho \frac{\sigma_{\xi}}{\sigma_{u}} u_{a_{it}}}{\sigma_{\xi}\sqrt{1-\rho^{2}}}) \frac{1}{\sigma_{u}} \phi(\frac{u_{a_{it}}}{\sigma_{u_{a_{it}}}}),$$

$$\Pr(P_{a_{it}} = 0) = \Phi(\frac{\xi^{**}_{a_{it}}(\Omega^{-}_{a_{it}})}{\sigma_{\xi}}).$$

To perform the estimation:

1. Choose a set of parameters

2. Calculate the *Emax* functions at all state points by solving backwards. From the Emax functions, calculate the cutoffs, the $\xi_a^{**}(\Omega_a^-)$'s.

3. Calculate the likelihood value for each person and thus for the sample.

4. Find the parameters that maximize the (log) likelihood function.

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

$$\log w_a^o = \log w_a + \eta_a$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

$$\log w_a^o = \log w_a + \eta_a$$

$$\log y_a = \beta_{0i} + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h + \beta_4 s_h a_h + \epsilon_a$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

$$\log w_a^o = \log w_a + \eta_a$$

$$\log y_a = \beta_{0i} + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h + \beta_4 s_h a_h + \epsilon_a$$

$$\xi_a \sim N(0, \sigma_\xi^2), \eta_a \sim N(0, \sigma_\eta^2), \epsilon_a \sim N(0, \sigma_\epsilon^2)$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

$$\log w_a^o = \log w_a + \eta_a$$

$$\log y_a = \beta_{0i} + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h + \beta_4 s_h a_h + \epsilon_a$$

$$\xi_a \sim N(0, \sigma_\xi^2), \eta_a \sim N(0, \sigma_\eta^2), \epsilon_a \sim N(0, \sigma_\epsilon^2)$$

$$E(\xi_a \eta_a) = E(\xi_a \epsilon_a) = E(\eta_a \epsilon_a) = 0$$

Eckstein and Wolpin (1989)

$$U_a = C_a + \alpha_1 P_a + \alpha_2 C_a P_a + \alpha_3 P_a K_{a-1} + \alpha_4 P_a n_a + \alpha_5 P_a s,$$

$$C_a = y_a + w_a P_a - c n_a + b P_{a,}$$

$$\log w_a = \gamma_0 + \gamma_1 s + \gamma_2 K_{a-1} - \gamma_3 K_{a-1}^2 + \xi_a$$

$$\log w_a^o = \log w_a + \eta_a$$

$$\log y_a = \beta_{0i} + \beta_1 a_{ha} + \beta_2 a_{ha}^2 + \beta_3 s_h + \beta_4 s_h a_h + \epsilon_a$$

$$\xi_a \sim N(0, \sigma_\xi^2), \eta_a \sim N(0, \sigma_\eta^2), \epsilon_a \sim N(0, \sigma_\epsilon^2)$$

$$E(\xi_a \eta_a) = E(\xi_a \epsilon_a) = E(\eta_a \epsilon_a) = 0$$

δ given

Note the following:

(1) the utility function is not time separable in participation - α_3 reflects the effect of having worked in the past on the current disutility of participation;

Note the following:

(1) the utility function is not time separable in participation - α_3 reflects the effect of having worked in the past on the current disutility of participation;

(2) the cost of young children does not depend on participation - *c* is a "goods" cost;

Note the following:

(1) the utility function is not time separable in participation - α_3 reflects the effect of having worked in the past on the current disutility of participation;

(2) the cost of young children does not depend on participation - *c* is a "goods" cost;

(3) there is a fixed cost of working, *b*;

Note the following:

(1) the utility function is not time separable in participation - α_3 reflects the effect of having worked in the past on the current disutility of participation;

(2) the cost of young children does not depend on participation - *c* is a "goods" cost;

(3) there is a fixed cost of working, *b*;

(4) husband's income has a household-specific constant, β_{0i}.

Data: NLSMW

Subsample of 318 white women age 39 to 44 in 1967.

As few as 4 annual observations and as many as 16, with 60 percent having at least 11.

Participation defined as having worked at least one week during the calendar year.

Experience	All Ages	Age39-42	Age51-58
0	.098	.132	.143
1-5	.244	.306	.294
6-10	.385	.493	.362
11-15	.729	.731	.886
16-20	.742	.725	.893
21-25	.754	.832	.860
26+	.929	.800	.957

Participation rises steeply with work experience.

Model Fit:

	Actual			Predicted		
Experience	All Ages	Age39-42	Age51-58	All Ages	Age39-42	Age51-58
0	.098	.132	.143	.139	.140	.112
1-5	.244	.306	.294	.226	.234	.214
6-10	.385	.493	.362	.430	.497	.383
11-15	.729	.731	.886	.636	.719	.586
16-20	.742	.725	.893	.754	.804	.715
21-25	.754	.832	.860	.820	.829	.826
26+	.929	.800	.957	.885	.881	.894

The model provides two ways for persistence in participation to arise:

(1) that there is habit persistence in the sense that the disutility of work falls with prior work experience

But, the estimates imply the opposite, that the disutility of work increases with work experience, $\alpha_3 < 0$.

(2) Work experience increases future wage offers.

The estimates imply that the wage-experience relationship exerts a strong effect on participation.

For example, at the estimated value of the wage-experience slope, a 39 year old woman with 10 years of experience would work 16 additional years up to age 59 relative to a women with no work experience.

Schooling also operates similarly to work experience.

Although additional schooling increases the disutility of work, implying that women with more schooling would work less, it increases wage offers sufficiently that more schooled women participate more.