CHAPTER 2: APPLICATIONS OF DISCRETE AND DYNAMIC CHOICE MODELS

Joan Llull

Structural Empirical Methods for Labor Economics (and Industrial Organization)

IDEA PhD Program
INTRODUCTION
Introduction

This chapter:

- **Spatial equilibrium**: Diamond (2016)
- Married woman’s labor force **participation**.
- **Human capital accumulation**: Heckman, Lochner, and Taber (1998).
SPATIAL EQUILIBRIUM: DIAMOND (2016)
Spatial Equilibrium Models

Date back from Rosen (1979) and Roback (1982).

The Rosen-Roback model is static (similar to the random utility models described in Chapter I, one time migration decisions).

These models typically feature (some of) the following:

- Perfectly mobile labor market (with moving costs?).
- Fixed land and endogenous housing markets.
- Local amenities.
- Productivity differentials across cities.
- Local price differences.

They are at the intersection of urban economics and labor economics (cities vs individuals as the subject of interest).
Diamond (2016)

I am going to present a stylized version of Diamond (2016) as a canonical example.

Spatial equilibrium model to determine causes and welfare consequences of increased skill sorting.

The model features:

- Perfectly mobile labor market (one time settlement).
- Fixed land and endogenous housing markets.
- Endogenous amenities (main novelty).
- Productivity differentials across cities.
**Motivation**

Panel A

\[ \Delta \text{In college employment ratio, 1980–2000} \]

\[ \Delta \text{In college employment ratio, 1980} \]

\[ \beta = 0.17 \]

\[ (0.027) \]

\[ R^2 = 0.13 \]

Panel B

\[ \Delta \text{In rent, 1980–2000} \]

\[ \Delta \text{In college employment ratio, 1980–2000} \]

\[ \beta = 0.70 \]

\[ (0.044) \]

\[ R^2 = 0.49 \]

Panel C

\[ \Delta \text{In noncollege wage, 1980–2000} \]

\[ \Delta \text{In college employment ratio, 1980–2000} \]

\[ \beta = 0.237 \]

\[ (0.026) \]

\[ R^2 = 0.24 \]

Panel D

\[ \Delta \text{In college wage, 1980–2000} \]

\[ \Delta \text{In college employment ratio, 1980–2000} \]

\[ \beta = 0.30 \]

\[ (0.025) \]

\[ R^2 = 0.36 \]
Model

Aggregate firm in each city:

\[ Y_j = K_j^{1-\alpha}(\theta U_j U_j^\rho + \theta S_j S_j^\rho)^{\frac{\alpha}{\rho}}. \]

Wage rates: workers’ marginal product \( W_{kj} \) for \( k \in \{S, U\} \).

Workers of skill \( k \) choose consumption of housing services and goods, as well as location to maximize utility:

\[
\max_{\{j,c,h\}} \zeta \ln c + (1 - \zeta) \ln h + G(A_j(U_j, S_j), X_j, z) + \varepsilon_j
\]

s.t. \( P_c + R_j h \leq W_{kj} \).

Housing supply:

\[ R_j = F(C_j, L_j). \]
Estimation

Except for amenity parameters associated to $z$, all relevant variation in the model is at the city-education group level.

Two-stage estimation:

1. **MLE/Logit** estimation of parameters for $z$, collapsing city-education group-specific parts in city-education dummies.

2. **Joint estimation** using two-step GMM (in diffs) of all other parameters, using city level variation.

**Instruments**: exogenous variables plus Bartik shocks and their interactions with variables that affect the housing supply elasticity.
Amenities: construct an amenity index $A_j$ related to endogenous supplies as $\Delta A_j = \gamma \Delta \ln \frac{S_j}{U_j} + \epsilon_j$.

Production function: standard expressions from FOCs.

Housing: equilibrium expression linking changes in rents with regulation and land availability indexes, interest rates, and endogenous housing demand variables.

Location choice: once the combination of all city-education group-specific elements of the utility are estimated as $\delta_{jk}$, a linear expression relates them to primitives.
Results

Panel A. Predicted change in ln college ratio due only to productivity changes

Panel B. Predicted change in ln college ratio due to observed wage and rent changes

Panel C. Predicted change in ln college ratio due to observed changes in wage, rent, and endogenous amenities
Married woman’s labor force participation
Utility

Unitary model of the household to describe female labor force participation decisions.

Single decision unit that takes into account the utilities of the two members of the couple in the decision process.

The couple’s utility is: \( U(c, d, n, x, \varepsilon(1 - d)) \), with:

- \( \frac{\partial U}{\partial c} > 0 \).
- \( \frac{\partial^2 U}{\partial c^2} < 0 \).
- \( U(c, 1, n, x, \varepsilon) > U(c, 0, n, x, 0) \) for some values of \( \varepsilon \).
- Typically, \( U(c, 1, n, x, \varepsilon) > U(c, 1, n', x, \varepsilon) \) for \( n > n' \) as well.
Budget constrain and distr. of unobservables

The **husband** is assumed to work, generating income $y$.

The **wife** receives a wage offer $\omega(x, v)$ and decides whether to work or not accordingly.

If the wife works, the household incurs in child care cost of $\pi$ per child.

The **budget constraint** is:

$$c = y + [\omega(x, v) - \pi n]d.$$  

Unobservables $\varepsilon$ and $\nu$ are **serially uncorrelated** and are **jointly distributed** as $F(\varepsilon, \nu|y, x, n)$.

The probability that the wife participates is:

$$\Pr(d = 1|x, n, y)$$

$$= \int 1\{U(y + \omega(x, v) - \pi n, 1, n, x, \varepsilon) - U(y, 1, n, x, 0) > 0\}dF(\varepsilon, \nu|y, x, n)$$

$$\equiv G(y, x, n).$$
Estimation approaches

Primitives to recover: $U(\cdot), \omega(\cdot)$, and $F(\cdot)$.

Four estimation methods: structural vs non-structural + parametric vs non-parametric.

We evaluate the convenience of these approaches considering three goals, to what extent the following elements affect participation:

1. wages.
2. husband’s earnings.
3. childcare costs.
Non-structural approaches

In a non-structural non-parametric approach, we do not need to make further assumptions: estimate $G(\cdot)$ non-parametrically.

First goal requires further assumptions: exclusion restriction.

Let $x_1$ denote the partition of the vector $x$ that affects wages but does not enter the utility function directly.

The effect of wage changes on participation can be inferred from $\partial G/\partial z'$.

The second goal is clearly feasible (within sample) without further assumptions, since $\partial G/\partial y$ is identified.

The third goal, on the contrary, is unfeasible without further assumptions, because $G$ and $\pi$ cannot be separately identified.

Parametric specification of $G(\cdot)$ (e.g. probit or logit): similar results but $\partial G/\partial y$ is also identified out of sample.
Non-parametric structural

The non-parametric structural approach requires identifying $U(\cdot)$, $\omega(\cdot)$, and $F(\cdot)$ separately without imposing additional assumptions about functional forms.

This is infeasible provided that wages are only observed for the individuals who work.

With further non-parametric assumptions and data on wages for the women who work, one could go a bit further.

For example, if $\omega(\cdot)$ is assumed to be additively separable $\Rightarrow$ deterministic part of the wage function is identified ($y$, $n$, and potentially some elements in $\mathbf{x}$ not included in $\mathbf{x}_1$, denoted by $\mathbf{x}_2$ are exclusion restrictions).

Further assumptions on $F(\cdot)$ could also lead to partial identification of $U(\cdot)$. 
Parametric structural

Consider the following very standard parametric assumptions:

\[ U(c, d, n, x, n, \varepsilon(1 - d)) \equiv c + (1 - d)[x'\beta + \gamma n + \varepsilon], \]

\[ \omega(x, \upsilon) = x'\delta + \upsilon, \]

and:

\[ (\varepsilon, \upsilon)'|y, x, n \sim \mathcal{N}(0, \Sigma). \]

Given this parameterization, the difference in utilities is:

\[ U(y + \omega(x, \upsilon) - \pi n, 1, n, x, \varepsilon) - U(y, 1, n, x, 0) > 0 \} \text{d}F(\varepsilon, \upsilon|y, x, n) \]

\[ = x'\delta - [\pi + \gamma]n - x'\beta + \upsilon - \varepsilon. \]
**Parametric structural**

Data on choices \( \Rightarrow \pi + \gamma, \delta_1, \) and \( \delta_2 - \beta, \) where \( \delta_1 \) and \( \delta_2 \) are the partitions of \( \delta \) associated, respectively, to \( x_1 \) and \( x_2. \)

Further data on wages for women who work \( \Rightarrow \beta \) and \( \delta \) are separately identified (Heckman selection approach).

Wage data and exclusion restrictions \( \Rightarrow \Sigma \) could also be identified (\( \sigma_{\nu\varepsilon} \) from Heckman, \( \sigma_{\nu}^2 \) from variance in wages, and \( \sigma_{\varepsilon}^2 \) from coefficient of \( x_1\delta_1 \) the probit).

**Goals:**

1. Elasticity of labor supply with respect to wages is only identified if there are exclusion restrictions.

2. Effect of husband’s income is assumed to be zero in this case; other utility functions would lead to different effects.

3. Effect of changing the cost of child care can also be identified, even though only \( \gamma + \pi \) is identified.
Human capital accumulation: Heckman, Lochner, and Taber (1998)
Equilibrium model for wage inequality

Literature on inequality: partial equilibrium.

Heckman, Lochner, and Taber (1998):

- General equilibrium.
- Overlapping generations.
- Human capital accumulation (education and on the job).
- New methods for estimating these models.

Using their estimated model, evaluate the mechanisms behind increasing wage inequality.
Model

Consider the following life-cycle maximization problem:

\[
V(h_a, b_a, e, i_t, r_{et}) \equiv \max_{c,g} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta V(h_{a+1}, b_{a+1}, e, i_{t+1}, r_{et+1}) \right\},
\]

s.t. \( b_{a+1} \leq b_a [1 + (1 - \tau)i_t] + (1 - \tau)r_{et} h_a (1 - g) - c. \)

On-the-job human capital accumulates as:

\[
h_{a+t}(\omega, e) = \omega g^{\eta_e} h_a (\omega, e)^{\psi_e} + (1 - \delta) h_a (\omega, e),
\]

with \( 0 < \eta_e < 1 \) and \( 0 < \psi_e < 1 \) for \( e \in \{S, U\} \).

Discrete distribution for \( \omega \), with eight points of support (four observable types, denoted by \( k \), and based on quartiles of AFQT test, two education groups).

Individuals are assumed to have perfect foresight of future prices and interest rates in equilibrium (no aggregate shocks in the economy).

Individuals work until age \( a_R \), when are forced to retire, and afterwards live until age \( \bar{a} \) without perceiving labor income.
**Model**

**Education decision** is taken at the front end:

$$\max_e \left[V^E(\omega, e, t) - \pi_e + \varepsilon_e\right].$$

**Output** $Y_t$ is determined by the following nested CES technology:

$$Y_t = \left\{ \alpha K^\phi_t + (1 - \alpha)[\theta_t L^\rho_{St} + (1 - \theta_t)L^\rho_{Ut}]^{\phi/\rho} \right\}^{1/\phi}.$$

Skill-biased technical change, determined by the evolution of $\theta_t$ is given by:

$$\ln \left( \frac{\theta_t}{1 - \theta_t} \right) = \ln \left( \frac{\theta_0}{1 - \theta_0} \right) + \varphi t.$$

**Equilibrium** is given by the sequence of interest rates $\{i_t\}_{t=0}^\infty$ and skill prices $\{r_{Ut}, r_{St}\}_{t=0}^\infty$ that **clear the market** subject to aggregate firm profit maximization, and workers’ lifetime utility maximization.
Estimation

They assume values for $\beta$, $\gamma$, $\delta(= 0)$, and $\tau$.

The tuition costs $\pi_e$ are estimated from the data.

The estimation of the remaining parameters is carried with a step-wise procedure.

**First step:** production function.

At old ages, say $a > a^*$ for some $a^*$, individuals no longer invest in human capital (that is, $g \approx 0$). Therefore:

$$w(a^* + 1, t + 1, h_{a^*+1}) \equiv r_{et+1}h_{a^*+1} = r_{et+1}h_{a^*}(1 - \delta),$$

which implies:

$$\frac{w(a^* + \ell, t + \ell, h_{a^*+\ell})}{w(a^*, t, h_{a^*})} = \frac{r_{et+\ell}(1 - \delta)\ell}{r_{et}}.$$

Normalizing $r_{e0} = 1$, skill prices are identified up to a scale $(1 - \delta)^t$. 
**Estimation**

Given these skill prices, the aggregate stocks of skill units can be recovered from the skill prices:

\[
\frac{\text{wage bill}_{et}}{r_{et}(1-\delta)^t} = \frac{L_{et}}{(1-\delta)^t}.
\]

**Relative demands** of the two labor inputs give:

\[
\ln \frac{r_{St}}{r_{Ut}} = \frac{\theta_t}{1-\theta_t} + (\phi - 1) \ln \frac{L_{St}}{L_{Ut}} = \ln \left( \frac{\theta_0}{1-\theta_0} \right) + \varphi t + (\phi - 1) \ln \frac{L_{St}}{L_{Ut}}.
\]

⇒ $\theta_0$, $\varphi$, and $\phi$ can be recovered by OLS.

The remaining aggregate PF parameters estimated analogously.

**Second step**: lifetime maximization problem.

NLS for wages ($g$ is unobserved, replaced by the solution of the dynamic problem). This solution is computed by backwards induction.

To estimate $h_0(k, e)$, parameterize $h_{aR}(k, e)$ and recover backwards.

---

Chapter 2: Applications of discrete and dynamic choice models
Estimation

Third step: education decision.

First estimate an auxiliary probit model to recover a non-parametric estimate of \( \{(1 - \tau)[V^E(\omega, S, t) - V^E(\omega, U, t)] + \mu_k\}/\sigma \) and a coefficient associated to estimated tuition costs \( \pi_e \).

Then, they recover the structural parameters from these estimates:

- \( \sigma \) is recovered as the coefficient associated to \( \pi_e \).
- \( \mu_k \) is recovered comparing the non-parametric estimates to the values predicted by the model (given the parameters estimated in steps one and two).