Chapter 3: Dynamic Discrete Choice Models I: Full Solution Approaches

Joan Llull

Advanced Econometric Methods II
Barcelona GSE

Introduction

Dynamic discrete choice models

Discrete choice models seen in previous chapter are **static**.

This course: **dynamic** discrete choice $\Rightarrow$ individuals consider the effect of **today's decisions** on **tomorrow's outcomes**.

Many examples in economics of **forward-looking** individuals:

- Labor: human capital/career decisions/migration.
- Macro/finance: investment decisions.
- IO: engine replacement/patents/market entry-stay-exit.
- Family economics: marriage/fertility.
- Health: smoking/going on a diet.
- Micro: social interactions.

Seminal work by Miller (1984), Wolpin (1984), Pakes (1986), and Rust (1987).

General Framework

Model primitives and decision problem

Time is discrete, $t = 1, ..., T$ (with T finite or infinite).

Choices: $d_t = \{j : j \in \mathcal{D} = \{1, 2, ..., J\}\}$, and $d_{jt} = \mathbb{1}\{d_t = j\}$ with $\sum_{j \in \mathcal{D}} d_{jt} = 1$.

State variables: $\boldsymbol{s}_t = \{\boldsymbol{x}_t, \boldsymbol{\varepsilon}_t\}$, where $\boldsymbol{x}_t$ is observable (by the econometrician) and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, ..., \varepsilon_{Jt})'$ is unobservable.

State variables evolve as **choice-specific Markovian** process:

$$\boldsymbol{s}_{t+1} \sim F(\boldsymbol{s}_{t+1} | \boldsymbol{s}_t, d_t).$$

Model primitives and decision problem (cont'd)

Intertemporal payoffs:

$$\mathbb{E}_t\left[\sum_{l=0}^{T-t}\beta^l U(\boldsymbol{s}_{t+l}, d_{t+l})\right].$$

The **primitives of the model** $\{U, F, \beta\}$ are known by the econometrician up to a parameter vector $\boldsymbol{\theta}$.

Agents are **expected utility maximizers**:

$$d_t^*(\boldsymbol{s}_t) = \arg\max_{d_t \in \mathcal{D}} \mathbb{E}_t\left[\sum_{l=0}^{T-t}\beta^l U(\boldsymbol{s}_{t+l}, d_{t+l})\right].$$

Baseline assumptions

Assumption 1 (additive separability, AS):

$$U(d_t, \boldsymbol{x}_t, \boldsymbol{\varepsilon}_t) = u(d_t, \boldsymbol{x}_t) + \varepsilon_t(d_t).$$

where $\varepsilon_t(d_t) \equiv \sum_{j \in \mathcal{D}} d_{jt}\varepsilon_{jt}$. We also define $u(d_t, \boldsymbol{x}_t) \equiv \sum_{j \in \mathcal{D}} d_{jt} u_{jt}(\boldsymbol{x}_t)$.

Assumption 2 (iid unobservables):

$$\boldsymbol{\varepsilon}_t | \boldsymbol{x}_t \sim i.i.d.\ F_\varepsilon(\boldsymbol{\varepsilon}_t) \quad (i.i.d \text{ across individuals and over time}).$$

Assumption 3 (conditional independence of future x):

$$F_x(x_{t+1} | d_t, x_t, \boldsymbol{\varepsilon}_t) = F_x(x_{t+1} | d_t, x_t).$$

Assumptions 2+3 lead to **conditional independence (CI)**:

$$F(x_{t+1}, \boldsymbol{\varepsilon}_{t+1} | d_t, x_t, \boldsymbol{\varepsilon}_t) = F_x(x_{t+1} | d_t, x_t) F_\varepsilon(\boldsymbol{\varepsilon}_{t+1}).$$

Assumption 4 (conditional logit, CLOGIT):

$$\{\boldsymbol{\varepsilon}_{jt} : j \in \mathcal{D}\} \text{ Independent across alternatives } + \text{ Type I extreme value.}$$

Value function

Let $V_t(\boldsymbol{x}_t)$ denote the ex-ante value function in period t:

$$V_t(\boldsymbol{x}_t) \equiv \mathbb{E}_{t-1}\left[\sum_{l=0}^{T-t}\sum_{j\in\mathcal{D}}\beta^l d^*_{jt+l}(u_{jt+l}(\boldsymbol{x}_{t+l}) + \varepsilon_{jt+l})\middle|\boldsymbol{x}_t\right].$$

This function is sometimes referred to as **Emax**.

Appealing to **Bellman's optimality principle**:

$$V_t(\boldsymbol{x}_t) = \mathbb{E}_{t-1}\left[\sum_{j\in\mathcal{D}} d^*_{jt}\left(u_{jt}(\boldsymbol{x}_t) + \varepsilon_{jt} + \beta\int V_{t+1}(\boldsymbol{x}_{t+1})dF_x(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t, d^*_t)\right)\middle|\boldsymbol{x}_t\right]$$

$$= \sum_{j\in\mathcal{D}}\int d^*_{jt}\left(u_{jt}(\boldsymbol{x}_t) + \varepsilon_{jt} + \beta\int V_{t+1}(\boldsymbol{x}_{t+1})dF(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t, d^*_t)\right)dF_\varepsilon(\boldsymbol{\varepsilon}_t).$$

Conditional choice probabilities

Define the conditional value function $v_{jt}(\boldsymbol{x}_t)$ as:

$$v_{jt}(\boldsymbol{x}_t) \equiv u_{jt}(\boldsymbol{x}_t) + \beta \int V_{t+1}(\boldsymbol{x}_{t+1}) dF_x(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t, j).$$

The individual chooses j in period t if and only if:

$$v_{jt}(\boldsymbol{x}_t) + \varepsilon_{jt} \geq v_{kt}(\boldsymbol{x}_t) + \varepsilon_{kt} \quad \forall k \neq j.$$

Given CLOGIT, the **conditional choice probabilities** (CCP) $p_{jt}(\boldsymbol{x}_t)$ are conditional logit type:

$$p_{jt}(\boldsymbol{x}_t) = \frac{e^{v_{jt}(\boldsymbol{x}_t)}}{\sum_{h \in \mathcal{D}} e^{v_{ht}(\boldsymbol{x}_t)}}.$$

We need to **solve the model** to get $v_{jt}(\boldsymbol{x}_t)$ as a function of primitives (backwards induction or fixed point). CLOGIT implies:

$$V_{t+1}(x) = \ln \sum_{j \in \mathcal{D}} \exp\{v_{jt+1}(x)\} + \gamma,$$

where γ is the Euler-Mascheroni constant.

The likelihood function

We have **longitudinal data** $\{d_{it}, \boldsymbol{x}_{it}\}_{i=1,...,N}^{t=1,2,...,T_i}$.

The **log-likelihood** of this sample is given by:

$$\mathcal{L}_{\mathrm{N}}(\boldsymbol{\theta}) = \sum_{\mathrm{i=1}}^{\mathrm{N}} \ln \Pr(d_{i1}, d_{i2}, ..., d_{iT_i} \ , \ x_{i1}, x_{i2}, ..., x_{iT_i}; \boldsymbol{\theta}) \equiv \sum_{\mathrm{i=1}}^{\mathrm{N}} \ell_i(\boldsymbol{\theta}).$$

Given Markovian structure and CI, we can **factorize**:

$$\ell_i(\theta) = \sum_{t=1}^{T_i} \ln \Pr(d_{it}|\boldsymbol{x}_{it}; \boldsymbol{\theta}) + \sum_{t=2}^{T_i} \ln \Pr(\boldsymbol{x}_{it}|\boldsymbol{x}_{it-1}, d_{it-1}; \boldsymbol{\theta}) + \ln \Pr(\boldsymbol{x}_{i1}; \boldsymbol{\theta}).$$

Motivational Example: Rust's Engine Replacement Model

Rust (Econometrica 1987)

Analyzes the behavior of **Harold Zurcher**, superintendent of maintenance at Madison Metropolitan Bus Company (Madison, Wisconsin).

Decision: every month t, to replace or to keep the engine of each bus i:

$$d_t = \begin{cases} 1 & \text{if replaces} \\ 0 & \text{if keeps.} \end{cases}$$

Trade-off: replacing $\Rightarrow$ replacement cost, lower maintenance cost; keeping $\Rightarrow$ saves the replacement cost, larger maintenance cost:

$$U(d_t, x_t, \boldsymbol{\varepsilon}_t) = \begin{cases} -[\theta_R + \theta_M 0] + \varepsilon_{1t} & \text{if } d_t = 1 \\ -\theta_M x_t + \varepsilon_{0t} & \text{if } d_t = 0. \end{cases}$$

State variables: x_t is mileage, $\boldsymbol{\varepsilon}_t = (\varepsilon_{0t}, \varepsilon_{1t})'$ is a vector of state variables unobserved by the econometrician.

Transition probabilities

Support of x is **discrete** $\{x_t = x : x \in X; t = 1, ..., T\}$.

$F^1_{x_{t+1},x_t}$ is degenerate.

$F^0_{x_{t+1},x_t}$ is a **transition matrix** whose elements we **estimate**:

$$F^0_{x_{t+1},x_t} = \begin{pmatrix} \varphi_0 & \varphi_1 & \varphi_2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & \varphi_0 & \varphi_1 & \varphi_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \varphi_0 & \varphi_1 & \varphi_2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \varphi_0 & \varphi_1 & \varphi_2 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & \varphi_0 & 1-\varphi_0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}.$$

Value function

Baseline assumptions apply.

The **conditional value function** is:

$$v_{jt}(x_t) = u_{jt}(x_t) + \beta \sum_{x \in X} \ln \left(\sum_{h \in \mathcal{D}} \exp\{v_{ht+1}(x)\} \right) F^j_{x,x_t} + \beta\gamma,$$

which, given **infinite horizon** describes $v_{jt}(x_t) \equiv v_j(x_t)$ for all t as the solution of a **fixed point**:

$$v_j(x_t) = u_j(x_t) + \beta \sum_{x \in X} \ln \left(\sum_{h \in \mathcal{D}} \exp\{v_h(x)\} \right) F^j_{x,x_t} + \beta\gamma.$$

Estimation

Rust's NFXP Algorithm

Consider the division of the parameter vector in **two subsets**: $\boldsymbol{\theta} = (\boldsymbol{\theta}'_U, \boldsymbol{\theta}'_x)'$.

Recall the there are **three of components** of the likelihood:

- $\Pr(d_{it}|x_{it}; \boldsymbol{\theta}) = \Pr(d_{it}|x_{it}; \boldsymbol{\theta}_U, \boldsymbol{\theta}_x)$.
- $\Pr(x_{it}|x_{it-1}, d_{it-1}; \boldsymbol{\theta}) = \Pr(x_{it}|x_{it-1}, d_{it-1}; \boldsymbol{\theta}_x)$.
- $\Pr(x_{i1}; \boldsymbol{\theta})$: This term can be **ignored given CI**.

A **two-step** algorithm estimates the two subsets separately:

- $\hat{\boldsymbol{\theta}}_x = \arg\max_{\boldsymbol{\theta}_x} \sum_{\mathrm{i=1}}^{\mathrm{N}} \sum_{t=2}^{T_i} \ln \Pr(x_{it}|x_{it-1}, d_{it-1}; \boldsymbol{\theta}_x)$, (solution not required)
- $\hat{\boldsymbol{\theta}}_U = \arg\max_{\boldsymbol{\theta}_U} \sum_{\mathrm{i=1}}^{\mathrm{N}} \sum_{t=1}^{T_i} \ln \Pr(d_{it}|x_{it}; \boldsymbol{\theta}_U, \hat{\boldsymbol{\theta}}_x)$.

A third step with a **single iteration of BHHH** —see next slide— with $(\hat{\boldsymbol{\theta}}_U, \hat{\boldsymbol{\theta}}_x)$ gives results that are asymptotically equivalent to FIML.

Rust's NFXP Algorithm (cont'd): BHHH

Rust proposes a **nested fixed point algorithm** (a BHHH algorithm combined with the solution of the DP):

Inner loop: Solve the DP for each parameter evaluation $\boldsymbol{\theta}_U^m$.

Outer loop: A BHHH optimization routine iterates over $\boldsymbol{\theta}_U$ to maximize the log-likelihood of the sample.

The **BHHH** is similar to Newton-Raphson except that avoids computing the Hessian:

$$\boldsymbol{\theta}^{m+1} = \boldsymbol{\theta}^m - \left(\sum_{\mathrm{i}=1}^{\mathrm{N}} \sum_{t=1}^{T_i} \frac{\partial \ell_{it}(\boldsymbol{\theta}^m)}{\partial \boldsymbol{\theta}} \frac{\partial \ell_{it}(\boldsymbol{\theta}^m)}{\partial \boldsymbol{\theta}'} \right)^{-1} \left(\sum_{\mathrm{i}=1}^{\mathrm{N}} \sum_{t=1}^{T_i} \frac{\partial \ell_{it}(\boldsymbol{\theta}^m)}{\partial \boldsymbol{\theta}'} \right).$$

Why is it an **approximation**?

Results in the Rust example

Table: First Stage Estimation: Transition Function for Mileage

Parameter	Group 1, 2, 3	Group 4	Group 1, 2, 3, 4
φ_0	0.29	0.40	0.33
	(0.01)	(0.01)	(0.01)
φ_1	0.70	0.59	0.66
	(0.01)	(0.01)	(0.01)
φ_2	0.01	0.01	0.01
	(0.00)	(0.00)	(0.00)

Courtesy of José García-Louzao, Sergi Marin Arànega, Alex Tagliabracci, and Alessandro Ruggieri, who replicated Rust's paper for the replication exercise in the Microeconometrics IDEA PhD course in Fall 2014.

Results in the Rust example

Table: Second Stage Estimation: Cost Function Parameters

Method	Parameter	Group 1, 2, 3	Group 4	Group 1, 2, 3, 4
NFXP	θ_R	11.87	10.12	9.75
		(1.95)	(1.36)	(0.89)
	θ_M	5.02	1.18	1.37
		(1.40)	(0.28)	(0.24)

Courtesy of José García-Louzao, Sergi Marin Arànega, Alex Tagliabracci, and Alessandro Ruggieri, who replicated Rust's paper for the replication exercise in the Microeconometrics IDEA PhD course in Fall 2014.

Extensions/Departures from the Rust Framework

Unobserved Heterogeneity

Motivational example: **Keane and Wolpin** (1997).

They analyze **career decisions** of young U.S. male.

Every year individuals **decide** one of:

- Stay home ($d_t = 0$).
- Work in blue collar ($d_t = 1$), white collar ($d_t = 2$) or military ($d_t = 3$).
- Attend school ($d_t = 4$).

State variables are $\boldsymbol{z}_t \equiv (e_t, x_{1t}, x_{2t}, x_{3t})'$, $\boldsymbol{\omega}$, and $\boldsymbol{\varepsilon}_t \sim i.i.d.\, \mathcal{N}(\mathbf{0}, \Sigma)$.

Utilities are:

$$U(d_t, \boldsymbol{z}_t, \boldsymbol{\omega}, \boldsymbol{\varepsilon}_t) = \begin{cases} \omega_0 + \varepsilon_{0t} & \text{if } d_t = 0 \\ r_j \exp\{\omega_j + \theta_{1j} e_t + \theta_{2j} x_{jt} + \theta_{3j} x_{jt}^2 + \varepsilon_{jt}\} & \text{if } d_t = 1, 2, 3 \\ \omega_4 + \theta_4\, \mathbb{1}\{e_t \geq 12\} + \theta_5\, \mathbb{1}\{e_t \geq 16\} + \varepsilon_{4t} & \text{if } d_t = 4. \end{cases}$$

How does it depart from Rust?

Transitions of the observable state variables: Deterministic!

Implications?

Some **assumptions for the unobservables** are relaxed:

- **AS** (because of wage equations).
- **CLOGIT** ($\boldsymbol{\varepsilon}_t$ jointly normal+potentially correlated across alternatives).
- **IID** (over time correlation through $\boldsymbol{\omega}$).

The first two add complication to computate **Emax** and **CCPs**.

The third one implies maximizing the **integrated log-likelihood**, which integrates over $\boldsymbol{\omega}$, as $\boldsymbol{\varepsilon}_t$ satisfies IID (as in duration or RPL).

Maximum Likelihood Estimation

Define $\Omega \equiv \{\boldsymbol{\omega}^k : k = 1, 2, ..., K\}$.

The log-likelihood is be:

$$\mathcal{L}_{\mathrm{N}}(\boldsymbol{\theta}) = \sum_{\mathrm{i=1}}^{\mathrm{N}} \ln \left\{ \sum_{k=1}^{K} \Pr(d_{i1}, d_{i2}, ..., d_{iT_i} \ , \ \boldsymbol{z}_{i1}, \boldsymbol{z}_{i2}, ..., \boldsymbol{z}_{iT_i} | \boldsymbol{\omega}^k; \boldsymbol{\theta}) \pi_{k|\boldsymbol{z}_{i1}} \right\},$$

where $\pi_{k|\boldsymbol{z}_{i1}} \equiv \Pr(\boldsymbol{\omega}_i = \boldsymbol{\omega}^k | \boldsymbol{z}_{i1})$.

What are the two **inconveniences** generated by this complication?

- Computational burden.
- $\Pr(z_{i1}; \boldsymbol{\theta})$.

Estimation of competitive equilibrium models

Motivational example: **Lee and Wolpin** (2006).

As Keane-Wolpin, but r_j becomes an **equilibrium object** r_{jt}.

We add a **labor demand**, and r_{jt} clears the market.

Very connected to **macro GE heterogeneous agents** models.

Entails **several complications**:

- **Solution of DP** is a function of $\{r_{jt}\}_{j \in \mathcal{D}}$ (state space aug.).
- **Market clearing** with labor demand to find r_{jt}^* (equil. FXP).
- Individuals have to **forecast future skill prices** (aggregate shock).
- Estimation requires **lots of data** (given equilibrium and non-stationarity) $\Rightarrow$ Simulated Minimum Distance.

Using experimental data to validate the model

Motivational example: **Todd and Wolpin** (2006).

The **goal of the paper** is to estimate a model of child education and fertility to evaluate alternative subsidies.

Make use of **PROGRESA** randomized implementation.

Advantage with respect to treatment effects: evaluate alternative subsidies and predict long-run effects of the subsidy.

Empirical strategy: estimate the model using only the control group (individuals from villages in which the subsidy was not implemented).

Assumptions:

- Identification of the effect of the subsidy comes from wages of children and the structure of the model.
- Households in control villages do not anticipate the subsidy.

Application: Llull (2018)

Labor Market Impacts of Immigration

In **Llull (2018)**, I analyze how natives respond to inflows of immigrants, and what are the effects on wages.

Labor supply and human capital decisions in the model as follows:

- Individuals decide **yearly** on **participation**, **education** and **occupation** from age 16 (or upon entry) to 65 (no return migration).

- **Immigration** and **capital** process are specified outside of the model, but allowed to be **endogenous** to aggregate conditions.

- An **aggregate firm** combines labor skill units with capital to produce a single output.

- Labor **skill rental prices** are determined in **equilibrium**. The wage of an individual i at time t in occupation j:

$$w_{i,t}^{j} = r_t^{j} \times s_i \equiv price_t^{j} \times skill\ units_i.$$

Labor Supply

- **Choice set**:
  - Working in a **blue-collar** job ($d_a = B$)
  - Working in a **white-collar** job ($d_a = W$)
  - Attending **school** ($d_a = S$)
  - Staying at **home** ($d_a = H$)

- They are **not allowed to save**, so they consume all their net income each period.

- Imperfect forecasting of **future labor market conditions**.

- State variables include l, E, X_B, X_W, X_F, n, d_{a-1}, $\boldsymbol{\varepsilon}_a$, $\boldsymbol{r}_t$, and t.

Labor Supply

Individuals solve the following **dynamic programming problem**:

$$V_{a,t,l}(\Omega_{a,t}) = \max_{d_a} U_{a,l}(\Omega_{a,t}, d_a) + \beta E\left[V_{a+1,t+1,l}(\Omega_{a+1,t+1}) \mid \Omega_{a,t}, d_a, l\right]$$

$$U^j_{a,t,l} = w^j_{a,t,l} + \delta^{BW}_g \mathbb{1}\{d_{a-1} \neq \{B, W\}\}, \qquad w^j_{a,t,l} = r^j_t \times s^j_{a,l}, \qquad j = B, W$$

$$w^j_{a,t,l} = r^j_t \exp\{\omega^j_{0,l} + \omega^j_{1,is} E_a + \omega^j_2 X_{Ba} + \omega^j_3 X^2_{Ba} + \omega^j_4 X_{Wa} + \omega^j_5 X^2_{Wa} + \omega^j_6 X_{Fa} + \varepsilon^j_a\}$$

$$\begin{pmatrix} \varepsilon^B_a \\ \varepsilon^W_a \end{pmatrix} \sim i.i.\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} (\sigma^B_g)^2 & \rho^{BW}\sigma^B_g\sigma^W_g \\ \rho^{BW}\sigma^B_g\sigma^W_g & (\sigma^W_g)^2 \end{bmatrix}\right)$$

$$U^S_{a,l} = \delta^S_{0,l} - \delta^S_{1,g}\mathbb{1}\{d_{a-1} \neq S\} - \tau_1\mathbb{1}\{E_a \geq 12\} - \tau_2\mathbb{1}\{E_a \geq 16\} + \sigma^S_g\varepsilon^S_a$$

$$U^H_{a,t,l} = \delta^H_{0,l} + \delta^H_{1,g} n_a + \delta^H_{2,g} t + \sigma^H_g \varepsilon^H_a$$

Notation: $a \equiv$ age; $l \equiv$ **ability type (gender$\times$region of origin)**; $t \equiv$ time; $g \equiv$ **gender**; $is \equiv$ **immigrant/native**.

Labor Demand

The **labor demand** is given by an **aggregate production function**:

- **Aggregate firm** produces with the following technology:

$$Y_t = z_t K_{St}^{\lambda}\{\alpha S_{Bt}^{\rho} + (1-\alpha)[\theta S_{Wt}^{\gamma} + (1-\theta)K_{Et}^{\gamma}]^{\rho/\gamma}\}^{(1-\lambda)/\rho}.$$

- **Two types** of labor: blue- and white-collar. **Workers** within an occupation are also **heterogeneous** in skills.

- **Imperfect substitutability** between natives and immigrants is endogenously generated through individual choices.

- The **nested CES** is included to capture the **capital-skill complementarity** and **SBTC** (Krusell et al., 2000).

- z_t is an **aggregate productivity shock** assumed to **evolve** according to:

$$\ln z_{t+1} - \ln z_t = \phi_0 + \phi_1(\ln z_t - \ln z_{t-1}) + \varepsilon_{t+1}^{z}$$

$$\varepsilon_{t+1}^{z} \sim \mathcal{N}(0, \sigma^{z}).$$

Equilibrium

In **equilibrium**:

- **Demands** of skill units are given by the first order conditions on firm's problem.
- The aggregate **supply** of skill units is given by:

$$S_{jt} = \sum_{a=16}^{65} \sum_{i=1}^{N} s_{a,i}^{j} \mathbb{1}\{d_{a,i} = j\} \quad j = B, W$$

$\Rightarrow$ The **equilibrium** is given by the skill prices that equate the supply and the demand of skill units (**market clearing**).

- **Expectations** are approximated with a VAR rule, in line with Lee and Wolpin (2006, 2010), and in the same spirit of Krusell and Smith (1998) $\Rightarrow$ **fixed point**.

Results

Counterfactual: keep immigrants so that the share is constant to 1965 levels.

Two types of exercises: **fixed capital** and **fixed interest rates**.

Main **results**:

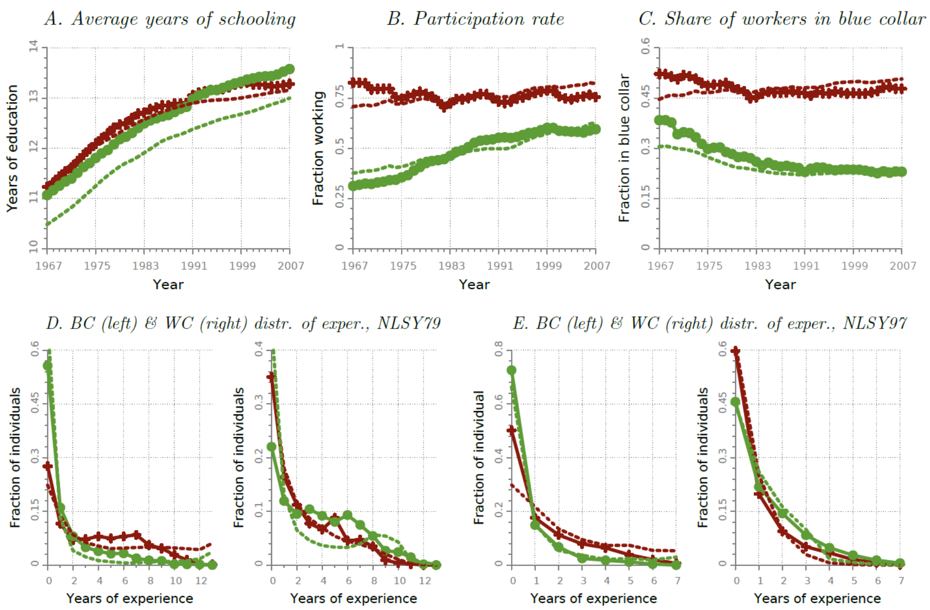
- **Equilibrium adjustments** are important to mitigate initial impacts on wages.
- Overall effects on **education** are very close to zero: strong heterogeneous effects that compensate each other.
- Participation margin matters for the effects **along the native wage distribution**.

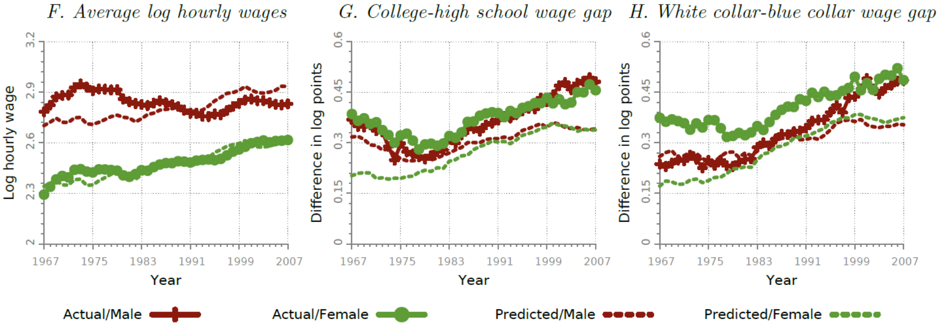
TABLE 4—EXPECTATION RULES FOR SKILL PRICES

	Blue-collar skill price		White-collar skill price	
Coefficient estimates:				
Constant (η_0)	0.002	(0.001)	0.002	(0.002)
Autoregressive term (η_j)	0.324	(0.046)	0.367	(0.048)
Δ Aggregate shock (η_z)	0.835	(0.046)	1.118	(0.065)
R-squared goodness of fit measures:				
Differences	0.870		0.858	
Levels	0.999		0.999	
Using predicted shock	0.221		0.222	

Note: The table includes estimates for the coefficients of expectation rules for aggregate skill prices —Equation (9). Goodness of fit measures are reported in the bottom panel. These measures are computed for the prediction of differences and levels for $j = B, W$. The last one uses the predicted increase in the aggregate shock obtained from Equation (7) instead of the actual increase. Standard errors (in parenthesis) are regression standard errors, and do not account for the error in the estimation of fundamental parameters.

Figure 2. Actual and Predicted Aggregates





Note: Panels A, B, C, F, G, and H are computed for individuals aged 25-54; actual data for these plots is obtained from March Supplements of the CPS (survey years from 1968 to 2008). In Panels D and E, experience is counted around 1993 (D) and (2006) for individuals in each cohort; sources for actual data in these plots are NLSY79 and NLSY97 as indicated.

TABLE 5—ACTUAL VS PREDICTED TRANSITION PROBABILITY MATRIX

	Choice in t							
	Blue collar		White collar		School		Home	
Choice in $t-1$	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
Blue collar	0.75	0.77	0.11	0.10	0.00	0.00	0.14	0.13
White collar	0.06	0.07	0.83	0.83	0.00	0.00	0.10	0.10
Home	0.11	0.08	0.13	0.13	0.01	0.01	0.76	0.79

Note: The table includes actual and predicted one-year transition probability matrix from blue collar, white collar, and home (rows) into blue collar, white collar, school, and home (columns) for individuals aged 25-54. Actual and predicted probabilities in each row add up to one. Actual data is obtained from one-year matched March Supplements of the CPS (survey years from 1968 to 2008).

TABLE 6—OUT OF SAMPLE FIT: ACT. VS PRED. STATISTICS FOR IMMIGRANTS

	Out-of-sample						In-sample	
	1970		1980		1990		1993-2007	
	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
A. Male								
Share with high school or less	0.67	0.69	0.57	0.61	0.52	0.55	0.55	0.56
Average years of education	10.8	11.1	11.4	11.8	11.7	12.1	11.9	12.1
Participation rate	0.77	0.56	0.68	0.61	0.63	0.66	0.75	0.72
Share of workers in blue collar	0.57	0.57	0.55	0.54	0.53	0.51	0.58	0.51
B. Female								
Share with high school or less	0.78	0.78	0.68	0.69	0.56	0.58	0.54	0.53
Average years of education	10.3	10.8	10.9	11.5	11.5	12.1	12.0	12.5
Participation rate	0.32	0.25	0.36	0.31	0.41	0.40	0.49	0.52
Share of workers in blue collar	0.46	0.45	0.45	0.44	0.39	0.43	0.41	0.43

Note: The table presents actual and predicted values of the listed aggregates for immigrants. Statistics for 1993-2007 are obtained from March Supplements of the CPS, and are used in the estimation. Data for 1970, 1980, and 1990 are from U.S. Census microdata samples and not used in the estimation.

Table 7—Estimated and Simulated Returns to Education

	Data		Simulation	
Least Squares (OLS)	0.096	(0.000)	0.096	(0.002)
Selection-corrected (Heckman, 1979)	0.123	(0.001)	0.114	(0.005)

Note: The table presents coefficients for years of education in OLS and Heckman (1979) selection-corrected regressions fitted on actual and simulated data. All regressions include dummies for potential experience (age minus education), gender, and year. In the selection-correction model, dummies for the number of children are included as exclusion restrictions. Actual data are obtained from the CPS. The sample period is 1967 to 2007. Random subsamples of 500,000 observations are drawn for both actual and simulated data. Nationally representative weights are used in the regressions. Standard errors, in parentheses, are calculated in the standard way in the left column, and are obtained from redrawing 100 times from the asymptotic distribution of the parameter estimates in the right column.

TABLE 8—PREDICTED ELASTICITY OF SUBSTITUTION BETWEEN IMMIGRANTS AND NATIVES

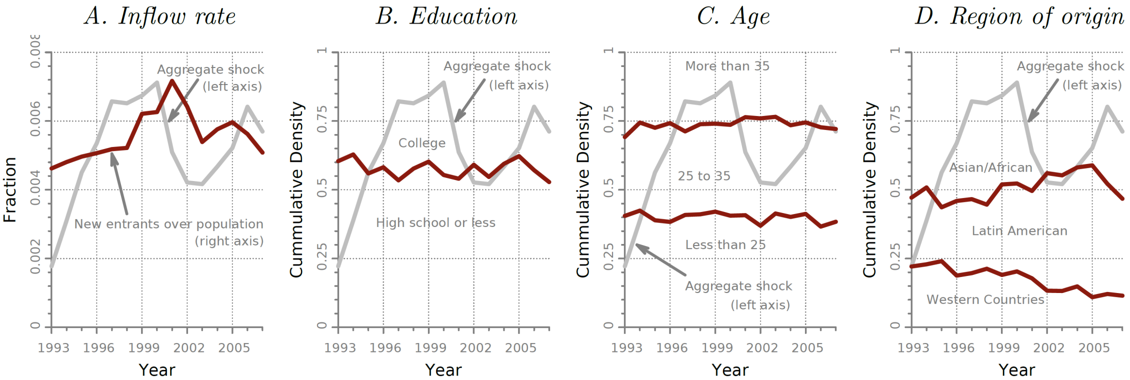
	Ottaviano and Peri (2012)		Simulations: Census years: 1970-2006		Simulations: Anual frequency: 1967-2007	
Baseline regression:						
Men	-0.048	(0.010)	-0.054	(0.011)	-0.050	(0.009)
Pooled Men and Women	-0.037	(0.012)	-0.065	(0.017)	-0.073	(0.014)
Men, Labor Supply is Employment	-0.040	(0.012)	-0.022	(0.012)	-0.008	(0.010)
Regression without cell and year dummies:						
Men	-0.063	(0.005)	-0.084	(0.015)	-0.083	(0.017)
Pooled Men and Women	-0.044	(0.006)	-0.137	(0.019)	-0.150	(0.020)
Men, Labor Supply is Employment	-0.066	(0.006)	-0.063	(0.022)	-0.060	(0.026)

Note: The table presents OLS estimates of $-1/\sigma_N$ from the following regression:

$$\ln\left(w_{Fkt}/w_{Dkt}\right) = \varphi_k + \varphi_t - 1/\sigma_N \ln\left(L_{Fkt}/L_{Dkt}\right) + u_{kt},$$

where $\{F, D\}$ indicate immigrants and natives respectively, k indicates education-experience cells, t indicates calendar year, w indicates average wages of skill cell k in year t, and L is labor supply in the corresponding cell. This regression corresponds to Equation (8) in Ottaviano and Peri (2012). The first column of the simulation results uses the same frequency as in Ottaviano and Peri (2012), excluding 1960; The second one include years 1967-2007 with annual frequency. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

Figure 3. Predicted Aggregate Shock and Recent Immigration



Note: Gray lines (right axis) plot the predicted values for the aggregate shock. Red lines (left axis) plot the share of new entrants over population (A), and, for them, the distribution of education (B), age (C), and country of origin (D). Data figures are smoothed with a 3-year moving average. Inflow rate is computed dividing the observed immigrants that entered over the preceding two (three) years divided by the number of years they refer to. *Source:* Current Population Survey, 1994-2008.

TABLE 9—EFFECTS ON SKILL PRICES AND THE ROLE OF EQUILIBRIUM

	No capital adjustment ($\partial K/\partial m = 0$):				Full capital adjustment ($\partial r_K/\partial m = 0$):			
	Blue collar		White collar		Blue collar		White collar	
No labor market adjustment	-4.92	(0.95)	-3.90	(0.60)	-1.76	(0.99)	0.86	(0.46)
Equilibrium effect	2.36	(0.78)	0.58	(0.72)	1.63	(1.00)	-0.86	(0.46)
Total effect	-2.56	(0.36)	-3.33	(0.39)	-0.13	(0.49)	-0.00	(0.15)

Note: The table compares baseline and counterfactual skill prices. Left and right panels correspond to different assumptions on counterfactual capital as indicated. "No labor market adjustment" indicates a scenario in which individuals are not allowed to adjust their human capital, occupational choice, and labor supply in response to immigration. "Equilibrium effect" is the difference between the total effect and the effect without labor market adjustment. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

Table 10—Wage Effects for Different Groups

	High school:				College:			
Age group:	25–39		40–54		25–39		40–54	
	No capital adjustment ($\partial K/\partial m = 0$):							
A. Male								
No labor market adjustment	-4.74	(0.72)	-4.55	(0.49)	-4.33	(0.22)	-4.18	(0.27)
Equilibrium effect	2.26	(0.82)	3.22	(0.90)	1.47	(0.70)	0.80	(0.44)
Total effect	-2.49	(0.28)	-1.33	(0.49)	-2.86	(0.66)	-3.38	(0.54)
B. Female								
No labor market adjustment	-4.47	(0.39)	-4.12	(0.38)	-4.08	(0.44)	-4.02	(0.48)
Equilibrium effect	1.33	(0.61)	2.48	(0.63)	0.43	(1.24)	0.43	(1.24)
Total effect	-3.14	(0.49)	-1.64	(0.50)	-3.65	(0.96)	-3.59	(1.00)
	Full capital adjustment ($\partial r_K/\partial m = 0$):							
A. Male								
No labor market adjustment	-1.32	(0.86)	-0.82	(0.64)	-0.24	(0.20)	0.15	(0.10)
Equilibrium effect	1.38	(0.93)	0.91	(0.73)	0.50	(0.47)	-0.01	(0.22)
Total effect	0.06	(0.17)	0.10	(0.20)	0.26	(0.30)	0.14	(0.22)
B. Female								
No labor market adjustment	-0.60	(0.58)	0.29	(0.42)	0.41	(0.54)	0.57	(0.49)
Equilibrium effect	0.69	(0.58)	-0.15	(0.40)	0.81	(1.03)	0.59	(0.94)
Total effect	0.09	(0.30)	0.14	(0.20)	1.22	(0.83)	1.16	(0.78)

Note: The table compares baseline and counterfactual average log wages for native males and females in different groups. In each panel, results are presented for different assumptions on counterfactual capital as indicated. "No labor market adjustment" indicates a scenario in which individuals are not allowed to adjust their human capital, occupation, and participation decisions. "Equilibrium effect" is the difference between the total effect and the effect without labor market adjustment. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

Table 11—Labor Supply Adjustments

Choice w/o immigr.	Fraction adjusting		Of which: Switch occ.		Stay home		Go to school	
			No capital adjustment ($\partial K/\partial m = 0$):					
A. Male								
Blue collar	8.8	(1.2)	55.8	(4.0)	38.5	(3.2)	5.6	(1.4)
White collar	7.7	(1.6)	52.3	(3.9)	40.1	(3.4)	7.6	(1.0)
Home	4.1	(0.7)	——		——		14.4	(2.9)
B. Female								
Blue collar	15.1	(4.2)	58.3	(7.9)	40.9	(7.3)	0.8	(1.0)
White collar	5.2	(2.0)	11.2	(7.1)	87.5	(6.4)	1.3	(1.4)
Home	4.0	(1.1)	——		——		1.7	(3.6)
			Full capital adjustment ($\partial r_K/\partial m = 0$):					
A. Male								
Blue collar	2.1	(1.4)	68.6	(5.3)	27.4	(5.2)	4.1	(1.3)
White collar	0.3	(0.6)	58.2	(9.7)	25.2	(8.4)	16.6	(4.3)
Home	1.6	(1.0)	——		——		9.1	(2.9)
B. Female								
Blue collar	6.1	(3.9)	68.7	(12.4)	30.7	(11.9)	0.6	(1.1)
White collar	0.7	(1.6)	46.6	(17.3)	52.1	(15.9)	1.3	(3.0)
Home	2.6	(1.5)	——		——		11.6	(5.6)

Note: The left column presents the percentage of native male and female individuals aged 25-54 that, in the cross-section of 2007, change their decisions in baseline and counterfactual simulations. The three remaining colums show the percentage of these individuals that do each of the adjustments indicated in the top row. Percentages are presented conditional on the choice made in the absence of immigration (counterfactual). Top and bottom panels make different assumptions a the counterfactual evolution of capital as indicated. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

Table 12—Education and Career Adjustments

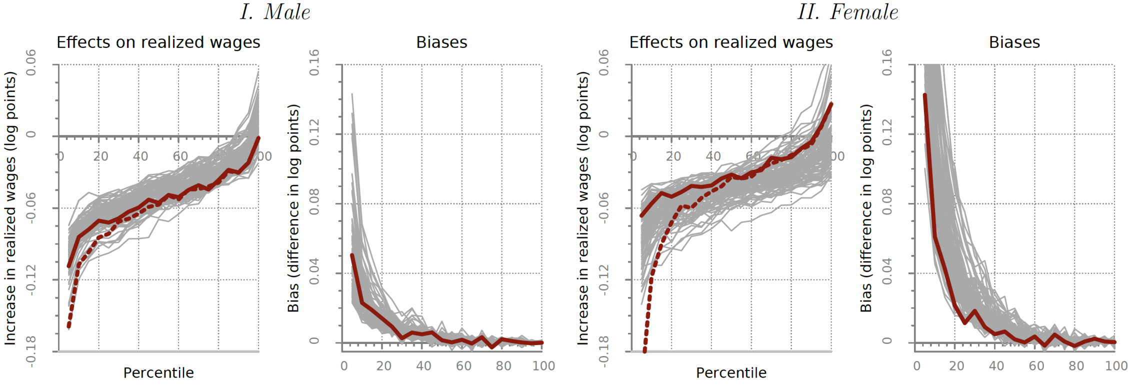
	All		Increase Education		Reduce Education		Keep educ. & change exp.	
			No capital adjustment ($\partial K/\partial m = 0$):					
A. Male								
Share of total	100.0		1.3	(0.4)	11.1	(2.4)	13.8	(1.6)
Average change in years of:								
Education	-0.28	(0.09)	3.20	(0.19)	-2.91	(0.12)	0.00	(0.00)
Experience in blue collar	-0.23	(0.16)	-7.65	(1.37)	2.33	(0.40)	-2.54	(0.37)
Experience in white collar	-0.07	(0.14)	3.32	(0.79)	-2.41	(0.31)	0.82	(0.34)
Time spent at home	0.58	(0.12)	1.13	(0.55)	2.99	(0.37)	1.72	(0.09)
B. Female								
Share of total	100.0		3.9	(1.4)	3.6	(3.0)	6.6	(0.7)
Average change in years of:								
Education	-0.00	(0.11)	2.71	(0.53)	-3.04	(0.30)	0.00	(0.00)
Experience in blue collar	-0.29	(0.15)	-4.26	(2.07)	1.04	(1.05)	-2.33	(1.01)
Experience in white collar	0.14	(0.21)	7.27	(2.37)	-4.17	(0.92)	-0.03	(0.77)
Time spent at home	0.15	(0.27)	-5.72	(2.09)	6.17	(0.80)	2.36	(0.72)

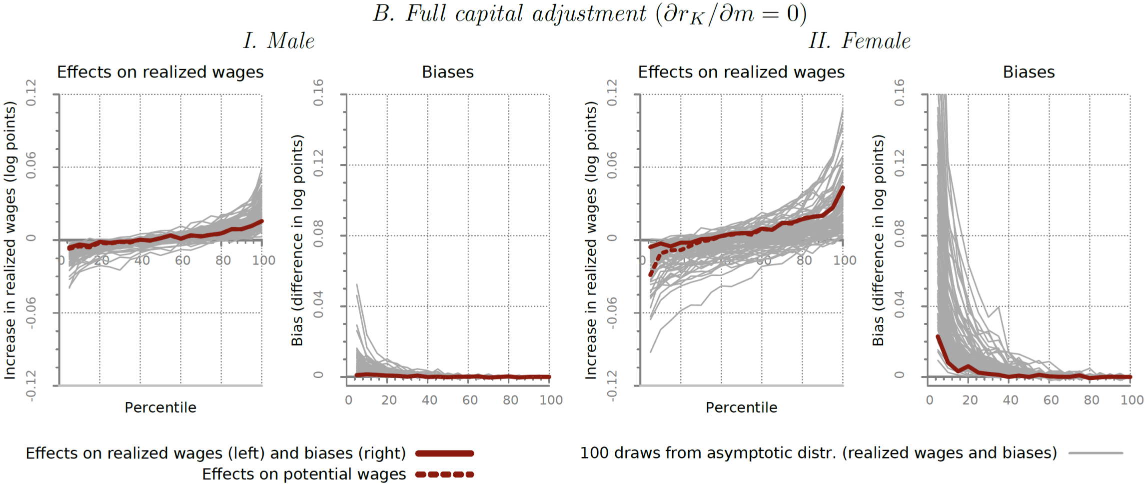
	Full capital adjustment ($\partial r_K/\partial m = 0$):							
A. Male								
Share of total	100.0		1.2	(0.7)	0.2	(0.8)	2.8	(1.6)
Average change in years of:								
Education	0.03	(0.03)	3.06	(0.17)	-2.90	(0.29)	0.00	(0.00)
Experience in blue collar	-0.13	(0.12)	-4.77	(1.55)	3.98	(1.88)	-2.40	(1.22)
Experience in white collar	0.08	(0.08)	3.20	(0.82)	-2.63	(1.28)	1.20	(0.69)
Time spent at home	0.02	(0.06)	-1.49	(0.95)	1.55	(0.79)	1.20	(0.67)
B. Female								
Share of total	100.0		3.8	(1.9)	0.4	(2.1)	1.7	(0.8)
Average change in years of:								
Education	0.08	(0.09)	2.34	(0.49)	-2.69	(0.55)	0.00	(0.00)
Experience in blue collar	-0.12	(0.06)	-1.71	(1.51)	-0.09	(1.57)	-2.91	(1.65)
Experience in white collar	0.14	(0.18)	3.81	(1.51)	-6.03	(1.49)	1.25	(1.18)
Time spent at home	-0.11	(0.26)	-4.44	(1.29)	8.81	(1.29)	1.66	(1.39)

Note: The top row of each panel indicates the fraction of individuals in each of the groups listed in the top row. The four rows at the bottom indicate the average change in the number of years in each of the alternatives accumulated by 2007. By construction, the sum of changes across alternatives in a given panel adds to zero. Different panels provide simulation results for the two genders in different capital scenarios as indicated. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

Figure 4. Wage Effects Along the Wage Distribution and Selection Biases

A. No capital adjustment ($\partial K \partial m = 0$)





Note: The figure plots the average differences in log hourly wages in baseline and counterfactual scenarios along the baseline wage distribution of native male and female aged 25-54 in 2007. The left figure in each pair represents wage effects on realized wages (solid black) and on potential wages (dashed black), and the right figure plots the difference between the two. Gray lines plot the effects on realized wages and the biases obtained for 100 random draws from the asymptotic distribution.