Dynamic discrete choice structural models: A survey

Aguirregabiria and Mira, J Econom, 2010

Milena Almagro, IO Reading Group

October 10, 2018

Dynamic discrete choice structural models: Road Map

- Single Agent
- Dynamic Discrete Games
- General Equilibrium

Single-agent Models

Set-up and Notation

- Agent i. Time t, with finite or infinite horizon
- State of the world: observable x_{it}, unobservable ϵ_{it}, action $a_{it} \in \{0, ..., J\}$
    - Sometimes observable payoff variables y_{it}
- State Markov transition $F(x_{it+1}, \epsilon_{it+1} | a_{it}, x_{it}, \epsilon_{it})$
- Objective at t $\max_{a \in A} \mathbb{E}\Big(\sum_{j=0}^{T-t} \beta^j U(a_{it}, x_{it}, \epsilon_{it}) | a_{it}, x_{it}, \epsilon_{it}\Big)$
- Value function

$$V(x_{it}, \epsilon_{it}) = \max_{a \in A} \Big\{ U(ax_{it+j}, \epsilon_{it+j}) + \beta \int V(x_{it+1}, \epsilon_{it+1}) dF(x_{it+1}, \epsilon_{it+1} | a_{it}, x_{it}, \epsilon_{it}) \Big\}$$

- Choice-specific value function

$$v(a, x_{it}, \epsilon_{it}) = U(a, x_{it}, \epsilon_{it}) + \beta \int V(x_{it+1}, \epsilon_{it+1}) dF(x_{it+1}, \epsilon_{it+1} | a_{it}, x_{it}, \epsilon_{it})$$

- Optimal decision rule $\alpha(x_{it}, \epsilon_{it}) = \arg\max_{a \in A} v(a, x_{it}, \epsilon_{it})$

Outline

Single-agent Models

Rust Model

Eckstein-Keane-Wolpin Models

Unobserved Heterogeneity

Dynamic Discrete Games

General Equilibrium Models

Revisiting Rust's Assumptions

- **Additive Separability (AS)**

$$U(a, x_{it}, \epsilon_{it}) = u(a, x_{it}) + \epsilon_{it}(a)$$

- **IID Unobservables (IID)** $\epsilon \sim G(\epsilon)$
- **Conditional Independence of X (CI-X)**

$$F(x_{it+1}|a_{it}, x_{it}, \epsilon_{it}) = F(x_{it+1}|a_{it}, x_{it})$$

- **Conditional Independence of Y (CI-Y)**

$$y_{it} = Y(a_{it}, x_{it}, \epsilon_{it}) = Y(a_{it}, x_{it})$$

- **CLOGIT** $\epsilon_{it}(a) \sim$ GEV, Type I
- **Discrete Support of x (DIS)**

$$x_{it} \in X = \{x^{(1)}, ..., x^{(|X|)}\} \text{ with } |X| < \infty$$

Likelihood under Rust's assumptions

With **CI-X** and **IID** we have Rust's Conditional Independence Assumption.
Under CIA, x_{it} sufficient statistic for the probability of the current choice.
Hence, likelihood is additively separable

$$l_i(\theta) = \sum_{t=1}^{T_i} \log \mathbb{P}(a_{it}|x_{it}, \theta) + \sum_{t=1}^{T_i} \log f_Y(y_{it}|a_{it}, x_{it}, \theta_Y) + \sum_{t=1}^{T_i} \log f_X(x_{it+1}|a_{it}, x_{it}, \theta_f) + \log \mathbb{P}(x_{i1}|\theta)$$

Using **CLOGIT**, also obtain closed form of Emax function

$$EV(x) = \log\left(\sum_{a=0}^{J} \exp\left\{u(a,x) + \beta \sum_{x'} EV(x') f_x(x'|a,x)\right\}\right)$$

and hence

$$v(a,x) = u(a,x) + \beta \sum_{x'} EV(x') f_x(x'|a,x)$$

Estimation Methods for Rust

- Nested Fixed Point Algorithm (skipped)
- Hotz-Miller's CCP Method
- Recursive CCP estimation (Nested Pseudo Likelihood)
- Simulation-based CCP estimator

Hotz-Miller (1993) overview

1. Estimate process governing evolution of x, $\hat{F}_x(x'|a, x)$ from the data

Hotz-Miller (1993) overview

1. Estimate process governing evolution of x, $\hat{F}_x(x'|a, x)$ from the data
2. Estimate CCP (frequency, kernel...) for every a and x, $\hat{P}(a, x)$

Hotz-Miller (1993) overview

1. Estimate process governing evolution of x, $\hat{F}_x(x'|a, x)$ from the data
2. Estimate CCP (frequency, kernel...) for every a and x, $\hat{P}(a, x)$
3. Recover value function from estimated CCP
    - With logit errors

$$\hat{v}(a, x) - \hat{v}(a', x) = \log \hat{P}(a, x) - \log \hat{P}(a', x)$$

Hotz-Miller (1993) overview

1. Estimate process governing evolution of x, $\hat{F}_x(x'|a,x)$ from the data
2. Estimate CCP (frequency, kernel...) for every a and x, $\hat{P}(a,x)$
3. Recover value function from estimated CCP
   - With logit errors

$$\hat{v}(a,x) - \hat{v}(a',x) = \log \hat{P}(a,x) - \log \hat{P}(a',x)$$

   - "Normalize" some action $\tilde{a}$, $u(\tilde{a},x) = 0$ and solve recursively for $\hat{v}(\tilde{a},x)$

$$v(\tilde{a},x) = \beta \int v(\tilde{a},x') dF_x(x'|\tilde{a},x) - \beta \int \log p(\tilde{a},x) dF(x'|\tilde{a},x)$$

Hotz-Miller (1993) overview

1. Estimate process governing evolution of x, $\hat{F}_x(x'|a,x)$ from the data
2. Estimate CCP (frequency, kernel...) for every a and x, $\hat{P}(a,x)$
3. Recover value function from estimated CCP
   - With logit errors

$$\hat{v}(a,x) - \hat{v}(a',x) = \log \hat{P}(a,x) - \log \hat{P}(a',x)$$

   - "Normalize" some action $\tilde{a}$, $u(\tilde{a},x) = 0$ and solve recursively for $\hat{v}(\tilde{a},x)$

$$v(\tilde{a},x) = \beta \int v(\tilde{a},x') dF_x(x'|\tilde{a},x) - \beta \int \log p(\tilde{a},x) dF(x'|\tilde{a},x)$$

   - Recover the rest of $\hat{v}(a,x)$ using $\hat{v}(\tilde{a},x)$

Hotz-Miller (1993) overview

1. Estimate process governing evolution of x, $\hat{F}_x(x'|a,x)$ from the data
2. Estimate CCP (frequency, kernel...) for every a and x, $\hat{P}(a,x)$
3. Recover value function from estimated CCP
   - With logit errors

$$\hat{v}(a,x) - \hat{v}(a',x) = \log \hat{P}(a,x) - \log \hat{P}(a',x)$$

   - "Normalize" some action $\tilde{a}$, $u(\tilde{a},x) = 0$ and solve recursively for $\hat{v}(\tilde{a},x)$

$$v(\tilde{a},x) = \beta \int v(\tilde{a},x')dF_x(x'|\tilde{a},x) - \beta \int \log p(\tilde{a},x)dF(x'|\tilde{a},x)$$

   - Recover the rest of $\hat{v}(a,x)$ using $\hat{v}(\tilde{a},x)$
4. Estimate θ based on $\hat{v}(x,a)$

Hotz and Miller: Linear Utility

- Linear function $u(a, x) = z(a, x)'\theta_u$
- Can be shown that

$$v(a, x) = \tilde{z}(a, x)'\theta_u + \tilde{e}(a, x)$$

where $\tilde{z}$ and $\tilde{e}$ are discounted sums of z and ϵ conditional on choice a

- $\tilde{z}$ and $\tilde{e}$ can be solved recursively

$$\tilde{z}(a, x_t) \equiv z(a, x_t) + \beta \sum_{x_{t+1}} f_x(x_{t+1}|a, x_t) \sum_{a' \in A} P(a'|x_{t+1})\tilde{z}(a', x_{t+1})$$

$$\tilde{e}(a, x_t) \equiv \beta \sum_{x_{t+1}} f_x(x_{t+1}|a, x_t) \sum_{a' \in A} P(a'|x_{t+1})[e(a', x_{t+1}) + \tilde{e}(a', x_{t+1})]$$

Hotz and Miller: Estimation

- Estimate via GMM with moment conditions

$$\sum_{i=1}^{N}\sum_{t=1}^{T_i} H(x_{it}) \begin{bmatrix} I\{a_{it=1}\} - \frac{\exp\{\hat{\tilde{z}}(1,x_{it})'\theta_u + \hat{\tilde{e}}(1,x_{it})\}}{\sum_{a=0}^{J}\exp\{\hat{\tilde{z}}(a,x_{it})'\theta_u + \hat{\tilde{e}}(a,x_{it})\}} \\ \vdots \\ I\{a_{it=J}\} - \frac{\exp\{\hat{\tilde{z}}(J,x_{it})'\theta_u + \hat{\tilde{e}}(J,x_{it})\}}{\sum_{a=0}^{J}\exp\{\hat{\tilde{z}}(a,x_{it})'\theta_u + \hat{\tilde{e}}(a,x_{it})\}} \end{bmatrix} \tag{1}$$

where $H(x_{it})$ are functions of instruments. For example

$$H(x_{it}) = (1, X_t, a_{it-1})$$

Hotz and Miller: Estimation

- Estimate via GMM with moment conditions

$$\sum_{i=1}^{N}\sum_{t=1}^{T_i} H(x_{it}) \begin{bmatrix} I\{a_{it=1}\} - \frac{\exp\{\hat{\tilde{z}}(1,x_{it})'\theta_u + \hat{\tilde{e}}(1,x_{it})\}}{\sum_{a=0}^{J}\exp\{\hat{\tilde{z}}(a,x_{it})'\theta_u + \hat{\tilde{e}}(a,x_{it})\}} \\ \vdots \\ I\{a_{it=J}\} - \frac{\exp\{\hat{\tilde{z}}(J,x_{it})'\theta_u + \hat{\tilde{e}}(J,x_{it})\}}{\sum_{a=0}^{J}\exp\{\hat{\tilde{z}}(a,x_{it})'\theta_u + \hat{\tilde{e}}(a,x_{it})\}} \end{bmatrix} \tag{1}$$

  where $H(x_{it})$ are functions of instruments. For example

$$H(x_{it}) = (1, X_t, a_{it-1})$$

- Where are moment conditions coming from?
- What about the non-linear case?

Pros

- Main advantage: computational simplicity. Only need to solve one recursive problem.
- For logit with linear utility, previous system of equations 1 has unique solution. Global search is not needed.

Hotz-Miller: Pros and Cons

Pros

- Main advantage: computational simplicity. Only need to solve one recursive problem.
- For logit with linear utility, previous system of equations 1 has unique solution. Global search is not needed.

Cons

- It is not efficient as opposed to NFXP-ML estimator.
- Aguirregabiria and Mira (2002) proposed a two-step pseudo-MLE.

$$\max_{\theta_u} \sum_{i=1}^{N} \sum_{t=1}^{T_i} \log \frac{\exp\{\hat{\tilde{z}}(a_{it}, x_{it})'\theta_u + \hat{\tilde{e}}(a_{it}, x_{it})\}}{\sum_{a=0}^{J} \exp\{\hat{\tilde{z}}(a, x_{it})'\theta_u + \hat{\tilde{e}}(a, x_{it})\}}$$

  Asymptotically equivalent to partial MLE but large finite sample bias.

- Cannot accommodate permanent unobserved heterogeneity

Recursive CCP estimation (Aguirregabiria and Mira (2002))

Step 1 Obtain an estimator of θ_u, $\hat{\theta}_u^1$

Step k
- Use $\hat{\theta}_u^{k-1}$ to form CCP's
$$\hat{P}(a|x)^{k-1}$$
- Obtain $\hat{v}(a,x)^{k-1}$ from $\hat{P}(a|x)^{k-1}$
- Obtain new estimate of θ_u
$$\hat{\theta}_u^K = \arg\max_{\theta_u} Q(\theta_u, \hat{P}^{k-1}, \hat{F}_x)$$

- Asymptotically equivalent to partial MLE and to tow-step PML
- Reduces finite sample bias

Simulation-based CCP estimator (Hotz et al. (1994))

HM impractical with large X, simulate to approximate values of CCP

- Take initial estimations of choice $\hat{P}$ and transition $\hat{F}$
- For every x_{it} in the sample and every $a \in A$, consider (a, x_{it}) as initial state
- Simulate R paths of future state variables and actions: for (a, x_{it}) draw x_{it+1} from $\hat{F}(x'|a, x_{it})$, then draw a_{t+1} from $\hat{P}(a|x_{it+1})$. And so on.
- Construct $\hat{z}(a, x_{it})$ and $\hat{e}(a|x_{it}) = \gamma - \log \hat{P}(a|x_{it}, \theta)$ and form

$$\tilde{z}_R^{\hat{P}}(a, x_{it}) = z(a, x_{it}) + \frac{1}{R}\sum_{r=1}^{R}\sum_{j=1}^{T^*}\beta^j z(a_{it+j}^r, x_{it+j}^r)$$

$$\tilde{e}_R^{\hat{P}}(a, x_{it}) = e(a, x_{it}) + \frac{1}{R}\sum_{r=1}^{R}\sum_{j=1}^{T^*}\beta^j e(a_{it+j}^r, x_{it+j}^r)$$

- Can we simulate v directly?
- Construct moment conditions

$$\mathbb{E}\left(h(x_{it})\Big[\log\left(\frac{P(a_{it}|x_{it})}{P(0|x_{it})}\right) - \left\{\tilde{z}_R^{\hat{P}}(a_{it}, x_{it}) - \tilde{z}_R^{\hat{P}}(0, x_{it})\right\}'\theta_u \right.$$

$$\left. - \left\{\tilde{e}_R^{\hat{P}}(a_{it}, x_{it}) - \tilde{e}_R^{\hat{P}}(0, x_{it})\right\}\Big]\right)$$

Outline

Example: Occupational Choice and Career Decisions

Each year, an individual chooses between staying at home $a = 0$, a white collar job, $a = 1$, a blue collar job $a = 2$ or attending school $a = 3$.

Per period utility

$$
\begin{aligned}
U(0, x_{it}) &= \omega_i(0) + \epsilon_{it}(0) \\
U(a, x_{it}) &= r_a \exp\{\omega_i(a) + \theta_{a1} s_{it} + \theta_{a1} e_{it} - \theta_{a3} e_{it}^2 + \epsilon_{it}(a)\} \\
U(3, x_{it}) &= \omega_i(3) - \theta_{tc} + \epsilon_{it}(3)
\end{aligned}
$$

Schooling s_{it} and experience evolve deterministically.

Both $\omega_i(a)$ and $\epsilon_{it}(a)$ are unobservable.

What assumptions are violated?

Issues

1. No **AS**: econometric model may not be saturated

   $\rightarrow$ allow for measurement error

2. No **CI-Y**: censured-choice payoff variable

   $\rightarrow$ have to use FIML, not two-step.

3. Unobserved heterogeneity: **agent types**

4. No **IID** over choices: unobservables correlated across choices

Outline

Single-agent Models

Rust Model

Eckstein-Keane-Wolpin Models

Unobserved Heterogeneity

Dynamic Discrete Games

General Equilibrium Models

Agent types: Mixture of Likelihoods (Heckman and Singer (2003))

Conditional on ω_i, $\{\epsilon_i(a)\}$ satisfy **IID**. If support of ω_i is $\Omega = \{\omega^1, ..., \omega^L\}$, then

$$l_i(\theta, \Omega, \pi) = \log\Big(\sum_{l=1}^{L} L_i(\theta, \omega^l)\pi_{l|x_{i1}}\Big)$$

where $\pi_{l|x} \equiv \mathbb{P}(\omega^l = \omega | x_{i1} = x)$ and

$$L_i(\theta, \omega^l) \equiv \Pi_{t=1}^{T_i}\mathbb{P}(a_{it}|x_{it}, \theta, \omega^l) f_Y(y_{it}|a_{it}, x_{it}, \theta, \omega^l)\Pi_{t=1}^{T_i-1} f_x(x_{it+1}|x_{it}, a_{it}, \theta_f, \omega^l)$$

- $\pi_{l|x_1}$ appears inside the log: cannot estimate f_Y and F_x separately
- Have to do FIML. Computationally very costly
- Moreover, #number recursive problems = #agent types

  $\rightarrow$ choose small L.

Sequential Expectation Maximization (Arcidiacono and Jones (2003))

- Recover additive separability with sequential **Expectation Maximization**
- For simplicity assume $\pi_{l|x_1} = \pi_l$, and CI-Y conditional on type
- Using Bayes

$$\mathbb{P}(l|a_i, x_i, y_i; \theta, \pi) = \frac{\pi_l \mathbb{P}(a_i, x_i, y_i|l, \theta)}{\mathbb{P}(a_i, x_i, y_i|\theta, \Omega, \pi)} = \frac{\pi_l L_i(\theta, \omega^l)}{\exp(l_i(\theta, \Omega, \pi))}$$

- Can be shown that FIML satisfies

$$\hat{\pi}_l = \sum_{i=1}^{N} \mathbb{P}(l|a_i, x_i, y_i; \hat{\theta}, \hat{\Omega}, \hat{\pi})$$

$$(\hat{\theta}, \hat{\Omega}) = \arg\max_{(\theta,\Omega)} \sum_{i=1}^{N} \sum_{l=1}^{L} \hat{\pi}_l \log L_i(\theta, \omega^l)$$

weights appear outside logs, then separable again!

Expectation Maximization Algorithm

Initialize at $\{\hat{\theta}_0, \hat{\Omega}_0, \hat{\pi}_0\}$

- Expectation step: compute

$$P_{il0} \equiv \mathbb{P}(l|a_i, x_i, y_i, \hat{\theta}_0, \hat{\Omega}_0, \hat{\pi}_0) = \frac{\hat{\pi}_{l0} L_i(\hat{\theta}_0, \hat{\omega}_0^l)}{\exp(l_i(\hat{\theta}_0, \hat{\Omega}_0, \hat{\pi}_0))}$$

- Maximization step: for P_{il0}, obtain

$$\hat{\pi}_{li} = \frac{1}{N}\sum_{i=1}^{N} P_{il0}$$

$$(\hat{\theta}_{x1}, \hat{\Omega}_{x1}) = \arg\max_{(\theta_x,\Omega_x)} \sum_{i=1}^{N}\sum_{l=1}^{L} P_{il0} \sum_{t=1}^{T_i-1} \log f_x(x_{it+1}|a_{it}, x_{it}, \theta_{x1}, \omega_{x1}^l)$$

$$(\hat{\theta}_{y1}, \hat{\Omega}_{y1}) = \arg\max_{(\theta_y,\Omega_y)} \sum_{i=1}^{N}\sum_{l=1}^{L} P_{il0} \sum_{t=1}^{T_i} \log f_y(y_{it+1}|a_{it}, x_{it}, \theta_{y1}, \omega_{y1}^l)$$

$$(\hat{\theta}_{u1}, \hat{\Omega}_{u1}) = \arg\max_{(\theta_u,\Omega_u)} \sum_{i=1}^{N}\sum_{l=1}^{L} P_{il0} \sum_{t=1}^{T_i} \log P(a_{it}|x_{it}, \theta_u, \hat{\theta}_{x1}, \hat{\theta}_{y1}, \omega_{u1}^l, \hat{\omega}_{x1}^l, \hat{\omega}_{y1}^l)$$

- If convergence, consistent and Asymptotically normal. Not efficient.

Keane-Wolpin's simulation and interpolation method

- Widely used for finite horizon, large X and correlated unobservables
- **CI-X** holds conditional on unobserved type, then likelihood $L_i(\theta, \omega^l)$ factors into transition and CCP $\rightarrow$ build Emax function
- Due to correlation, no closed form of Emax $\rightarrow$ numerical integration
- Finite horizon $\rightarrow$ Emax function solved by backward induction

$$\tilde{V}_{lt} = \frac{1}{R} \sum_{r=1}^{R} \max_{a \in A} \left\{ U_t(a, x, \epsilon_t^r, \omega^l, \theta) + \beta \sum_{x' \in X} \tilde{V}_{lt+1}(x') f_x(x'|a, x) \right\}$$

- Large state space $X \rightarrow$ Emax simulated only at some points of X, then interpolated using regression.
- Perform FIML on finite mixtures

$$l_i(\theta, \Omega, \pi) = \log \left(\sum_{l=1}^{L} L_i(\theta, \omega^l) \pi_{l|x_{i1}} \right)$$

Dynamic Discrete Games

Set-up

- Players $i \in \{1, ..., N\}$, simultaneously decide discrete action a_{it}
- Game is played independently at different locations m
- x_{it} common knowledge, ϵ_{it} private information
- Assume **AS, IID, CI-X**
- Choice-specific value function for strategy profile α, generating CCP P

$$\begin{aligned} v_i^P(a_i, x_t) =& \mathbb{E}_{\epsilon_{-i}}\Big[u_i(a_i, \alpha_{-i}(x_t, \epsilon_{-it})) \\ &+ \beta \int V_i^\alpha(x_{t+1}, \epsilon_{t+1}) dF_x(x_{t+1}|(a_i, \alpha_{-i}(x_t, \epsilon_{-it}), x_t)\Big] \end{aligned}$$

- Define best-response function probability function

$$\Lambda(a_i|v_i^P(., x_t)) \equiv \int I\Big\{a_i = \arg\max_{j \in A}\{v_i^P(j, x_t) + \epsilon_{it}(j)\}\Big\}$$

- For given θ, P is an MPE $\iff$ $P = \Lambda(v^P(\theta))$

Example: entry-exit model

N potential entrants. Payoff

$$U_{imt}(1) = \theta_{RS}\log(S_{mt}) - \theta_{RN}\log(1 + \sum_{j \neq i} a_{jmt}) - \theta_{FCi} - \theta_{ECi}(1 - a_{imt-1}) + \epsilon_{imt}(1)$$

$$U_{imt}(0) = \epsilon_{imt}(0)$$

Estimation first steps

- **ONE-MPE-Data**. Define $P^0_{mt} \equiv \{\mathbb{P}(a_{mt} = a | x_{mt=x}) : (a, x) \in A^N \times X\}$
  - For every (m, t), $P^0_{mt} = P^0$
  - Players expect P^0 to be played in future periods (out of sample)
  - $\{a_{mt}, x_{mt}\}$ are independent across markets and $\mathbb{P}(x_{mt} = x) > 0$ for all x
- Estimate θ_x directly from

$$\sum_{m=1}^{M} \sum_{t=1}^{T_m - 1} \log f_x(x_{mt+1} | a_{mt}, x_{mt}; \theta_x)$$

- Define pseudo-likelihood function for arbitrary P

$$Q(\theta, P) = \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{i=1}^{N} \ln \Lambda(a_{imt} | v_i^P(., x_{mt}, \theta))$$

- MLE is defined as

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \left\{ \sup_{P} Q(\theta, P) \quad \text{s.t. } P = \Lambda(v^P(\theta)) \right\}$$

  In practice very hard to implement, specially if multiple equilibria.
- Where are we imposing that P has to be consistent with data?

Two-step methods

- With **ONE-MPE-Data**, CCP $\sim$ beliefs about opponent behavior.
- Assume linear utility $u_i(a_{mt}, x_{mt}, \theta_u) = z_i(a_{mt}, x_{mt})'\theta_u$
- Then $v_i^P(a_i, x_t) = \tilde{z}_i(a_{mt}, x_{mt})'\theta_u + \tilde{e}_i(a_{mt}, x_{mt})$ with

$$\tilde{z}_i(a_i, x_t) \equiv \sum_{a_{-i}} \left(\Pi_{j\neq i} P_j(a_j|x_t)\right)\left[z_i(a_i, a_{-i}, x_t) + \beta \sum_{x_{t+1}} f_x(x_{t+1}|a_i, a_{-i}, x_t) W_{zi}^P(x_{t+1})\right]$$

$$W_{zi}^P(x_{t+1}) = \sum_{a_i \in A} P_i(a_i|x_{t+1})\tilde{z}_i(a_i, x_{t+1})$$

$$\tilde{e}_i(a_i, x_t) \equiv \beta \sum_{a_{-i}} \left(\Pi_{j\neq i} P_j(a_j|x_t)\right) \sum_{x_{t+1}} f_x(x_{t+1}|a_i, a_{-i}, x_t) W_{ei}^P(x_{t+1})$$

$$W_{ei}^P(x_{t+1}) = \sum_{a_i \in A} P_i(a_i|x_{t+1})[e(a_i, x_{t+1}) + \tilde{e}_i(a_i, x_{t+1})]$$

Estimation

- GMM: consistent and asymptotically normal

$$\sum_{i=1}^{M}\sum_{i=1}^{N}\sum_{t=1}^{T_m-1} H_i(x_{mt}) \begin{bmatrix} I\{a_{imt=1}\} - \Lambda\big(1|\tilde{z}^{\hat{P}}(.,x_{mt})'\theta_u + \tilde{e}^{\hat{P}}(.,x_{mt})\big) \\ \vdots \\ I\{a_{imt=J}\} - \Lambda\big(J|\tilde{z}^{\hat{P}}(.,x_{mt})'\theta_u + \tilde{e}^{\hat{P}}(.,x_{mt})\big) \end{bmatrix} = 0$$

- Minimum distance: consistent and asymptotically normal

$$\hat{\theta}_u = \arg\min_{\theta_u} \big[\hat{P} - \Lambda(v^{\hat{P}}(\theta))\big]' A_M \big[\hat{P} - \Lambda(v^{\hat{P}}(\theta))\big]$$

  Can achieve efficiency with a three step procedure

- Bajari, Benkard and Levin (2007): continuous state variables with monotonicity of policy functions

Pros and Cons of two-step procedures

Pros

- Computational simplicity

Cons

- Finite sample bias
- Ignores persistent unobservables
- Usually, inefficient

Sequential Estimation (Aguirregabiria and Mira (2007))

- Extend entry-exit model but include unobserved market heterogeneity

$$U_{imt}(1) = \theta_{RS}\log(S_{mt}) - \theta_{RN}\log(1 + \sum_{j\neq i} a_{jmt}) - \theta_{FCi} - \theta_{ECi}(1 - a_{imt-1}) + \omega_m + \epsilon_{imt}(1)$$

- ω_m has discrete and finite support and iid across markets with

$$\pi_l \equiv \mathbb{P}(\omega_m = \omega^l)$$

- Assume

$$P(x_{mt+1}|a_{mt}, x_{mt}, \omega_m) = f_x(x_{mt+1}|a_{mt}, x_{mt})$$

$\rightarrow$ can estimate transition probabilities from the data

- Relax **ONE-MPE-Data**

$$P^0_{mt} \equiv P^0_l$$

where l is the type of market m: one equilibrium per market type.

Sequential Estimation

Pseudo-likelihood has finite mixture form

$$Q(\theta_u, \{P_l\}) = \sum_{m=1}^{M} \Big(\sum_{l=1}^{L} \pi_{l|x_{m1}} \Pi_{t=1}^{T} \Pi_{i=1}^{N} \Lambda(a_{imt} | \tilde{z}^{\hat{P}_l}(., x_{mt})' \theta_u + \tilde{e}^{\hat{P}_l}(., x_{mt})) \Big)$$

Iterative procedure similar to recursive CCP

- Start with arbitrary choice probabilities: $\{\hat{P}_{l0} : l = 1, ..., L\}$
- Step 1: For every market l, obtain probabilities $\pi_{l|x_{m1}}$ assuming x_{m1} is drawn from the stationary distribution induced by the MPE.
- Step 2: Obtain pseudo-MLE

$$\hat{\theta}_{u1} = \arg\max Q(\theta_u, \{\hat{P}_{l0}\})$$

- Step 3: Update

$$\hat{P}_{l1} = \Lambda(v^{\hat{P}_{l0}}(\hat{\theta}_{u1}), \omega_l)$$

- Repeat 1-3 until convergence

General Equilibrium Models

Model Set Up (Lee and Wolpin (2010))

Embed the occupational choice model in a competitive labor market.

Supply side is an OLG version of such model.

Demand side characterize by Cobb-Douglas

$$Y_t = z_t S_{1t}^{\alpha_{1t}} S_{2t}^{\alpha_{2t}} K_t^{1-\alpha_{1t}-\alpha_{2t}}$$

Then

$$r_{at} = \frac{\alpha_{at}}{S_{at}} z_t S_{1t}^{\alpha_{1t}} S_{2t}^{\alpha_{2t}} K_t^{1-\alpha_{1t}-\alpha_{2t}} \qquad \text{for } a = 1, 2$$

$$S_{at} = \int_{x,\omega,\epsilon} k_{it}(a) I\{a = \alpha(x_{it}, \omega_i, \epsilon_{it}, X_t)\} di$$

z_t follows an AR(1). Evolution of skill-prices beliefs a la Krusell-Smith

$$\log r_{at+1} - \log r_{at} = \eta_{a0} + \sum_{k=1}^{2} \eta_{ak}(\log r_{kt} - \log r_{kt-1}) + \eta_{a3}(\log z_{t+1} - \log z_t)$$

Note that η is determined in equilibrium and is not a structural parameter

Solving the Model

Equilibrium in this model is a vector $\eta^*(\theta)$ that solves a fixed point.

- Step 1: Given η_0, individuals solve their problem.
- Step 2: Given z_0, r_{a0} and distribution of state variables,

  a Simulate sequence $\{z_t\}$

  b Guess skill prices using TFP draw and previous equation. Draw ϵ_{iat}. Obtain S_{at} for $t = 1$.

  c With S_{at}, obtain a new value of skill prices

  $$r'_{at} = \frac{\alpha_{at}}{S_{at}} z_t S_{1t}^{\alpha_{1t}} S_{2t}^{\alpha_{2t}} K_t^{1-\alpha_{1t}-\alpha_{2t}}$$

  d Repeat b-c until convergence of r_{at}

  e Repeat b-c-d for $t = 2, ..., T$

- Step 3: Use new sequence of $\{r_{at}\}$ and z_t to update η using OLS to obtain η_1
- Step 4: Repeat 1-3 until convergence of η.

Estimation

Estimation uses a Simulated Method of Moments

Criterion is weighted average distance between sample and simulated moments.

Estimation procedure is a nested solution-estimation algorithm:

- Outer iteration: looks for θ that minimizes criterion (Newton)
- Inner iteration: for given θ, solve for equilibrium as above. Given equilibrium beliefs, simulate data and calculate simulated moments.

Where is individual data used in this procedure?