Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with 2-period Finite Dependence

Yu Hao [1] Hiroyuki Kasahara [2]

[1]The University of Hong Kong

[2]The University of British Columbia

Department of Decisions, Operations and Technology
CUHK Business School

Bellman equation and the Conditional Choice Probabilities

$$\mathrm{V}(x) = \int \max_{d\in\{0,1\}} \{\mathrm{v}_\theta(x,d) + \epsilon(d)\}\, g(\epsilon|x)d\epsilon$$
$$:= [\Gamma(\theta, V)](x) \qquad \text{(Bellman operator)},$$

$$\mathrm{P}(d=1|x;\theta) = \frac{1}{1+\exp\left(\mathrm{v}_\theta(x,0) - \mathrm{v}_\theta(x,1)\right)} \qquad \text{(Type I EV)}$$
$$:= [\Lambda(\theta, V)](a|x).$$

where

$$\mathrm{v}_\theta(x,d) := u_\theta(x,d) + \beta \sum_{x^\dagger} \mathrm{V}(x^\dagger) f(x^\dagger|x,d).$$

Maximum Likelihood Estimator (e.g., Rust (1987))

$$\max_{\theta} \sum_{i=1}^{n} \ln[\Lambda(\theta, V_\theta)](d_i|x_i)$$

subject to

$$V_\theta = \Gamma(\theta, V_\theta).$$

Hotz and Miller's CCP estimator

- The Bellman equation: a fixed point problem in the space of value functions.

$$V = \Gamma(\theta, V).$$

- The model's restriction can also be characterized by a fixed point problem in the space of conditional choice probabilities:

$$P = \Psi(\theta, P).$$

Hotz and Miller's CCP estimator

$$V(x) = \sum_{d\in\{0,1\}} \mathrm{P}(d|x)\left\{ u_\theta(x,d) + \underbrace{E[\epsilon(d)|\mathrm{P}(d|x)]}_{-\ln \mathrm{P}(d|x)} + \beta \sum_{x^\dagger\in X} V(x')f(x'|x,d)\right\},$$

- The Policy iteration mapping:

$$\Psi_{HM}(\theta,P) := \Lambda(\theta,\varphi(\theta,P)).$$

where

$$V = (I-\beta E_P)^{-1} u_{\theta,P} := \varphi(\theta,P).$$

- CCP estimator:

$$\max_\theta \sum_{i=1}^{n} \ln[\Psi_{HM}(\theta,\hat{P})](d_i|x_i)$$

Alternative policy function mappings under finite dependence

- Finite dependence: there exist a sequence of <u>future</u> choices such that the subsequent continuation values do not depend on the current choice.

- Engine replacement at $t+1$ (i.e., $a_{t+1} = 1$) $\Rightarrow$ continuation value at $t+2$ does not depend on the current choice a_t.

Example of finite dependence (Bus engine replacement model)

$$\max_{d_1, d_2, \ldots} E\left[\sum_{t=1}^{\infty} \beta^t \left\{u_\theta(x_t, d_t) + \epsilon_t(d_t)\right\}\right],$$

where

$$\mathrm{u}_\theta(x_t, d_t) = \begin{cases} 0, & \text{if } d_t = 1 \\ \theta_0 + \theta_1 x_t, & \text{if } d_t = 0, \end{cases}$$

- x_t: mileage
- $f(x_{t+1} | x_t, d_t)$: transition probability with $f(0|x_t, 1) = 1$
- $\epsilon_t = (\epsilon_t(0), \epsilon_t(1))' \overset{iid}{\sim}$ Type I EV

An estimator by [Arcidiacono and Miller, 2011]

$$\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \ln[\Psi_{FD}(\theta, \hat{P}, \hat{f})](d_i|x_i).$$

$$[\Psi_{FD}(\theta, P, f)](d = 1|x)$$
$$= \frac{1}{1 + \exp\left(\theta_0 + \theta_1 x + \beta \sum_{x^\dagger} \log \mathrm{P}(1|x^\dagger)\left(f(x^\dagger|x, 1) - f(x^\dagger|x, 0)\right)\right)}$$

→ Given $\hat{P}$ and $\hat{f}$, **estimating** $(\theta_0, \theta_1, \beta)$ **is as easy as estimating a logit model!**

CCP under finite dependence

$$P_\theta(1|x) = \frac{1}{1 + \exp(\mathrm{v}_\theta(x,0) - \mathrm{v}_\theta(x,1))},$$

where

$$\mathrm{v}_\theta(x,d) := \mathrm{u}_\theta(x,d) + \beta \sum_{x^\dagger} \mathrm{V}(x^\dagger) f(x^\dagger|x,d).$$

We can show that

$$\begin{aligned} \mathrm{V}(x) &= \mathrm{v}_\theta(x,1) - \log \mathrm{P}(1|x) = \mathrm{v}_\theta(x,0) - \log \mathrm{P}(0|x) \\ &= w(\mathrm{v}_\theta(x,1) - \log \mathrm{P}(1|x)) + (1-w)(\mathrm{v}_\theta(x,0) - \log \mathrm{P}(0|x)). \end{aligned}$$

CCP under finite dependence

$$\mathrm{v}_\theta(x,d) = \mathrm{u}_\theta(x,d) + \beta \sum_{x^\dagger} \mathrm{V}(x^\dagger) f(x^\dagger | x, d)$$

$$\begin{aligned}
\mathrm{V}(x^\dagger) &= \mathrm{v}_\theta(x^\dagger, 1) - \log \mathrm{P}(1|x^\dagger) \\
&= u_\theta(x^\dagger, 1) - \log \mathrm{P}(1|x^\dagger) + \underbrace{\beta \sum_{x^{\dagger\dagger}} \mathrm{V}(x^{\dagger\dagger}) f(x^{\dagger\dagger}|x^\dagger, 1)}_{=\mathrm{V}(0)}.
\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}
\mathrm{v}_\theta(x,d) = \mathrm{u}_\theta(x,d) &+ \beta \sum_{x^\dagger} \left\{ u_\theta(x^\dagger, 1) - \log \mathrm{P}(1|x^\dagger) \right\} f(x^\dagger | x, d) \\
&+ \beta^2 \mathrm{V}(0) \sum_{x'} f(x^\dagger | x, d).
\end{aligned}$$

CCP under decision-specific finite dependence [Arcidiacono and Miller, 2011]

$$\begin{aligned} \mathrm{v}_\theta(x,0) - \mathrm{v}_\theta(x,1) &= u_\theta(x,0) - u_\theta(x,1) \\ &\quad + \beta \sum_{x^\dagger} \left\{ u_\theta(x^\dagger,1) - \log \mathrm{P}(1|x^\dagger) \right\} \tilde{f}(x^\dagger|x) \\ &\quad + \beta^2 \mathrm{V}(0) \underbrace{\sum_{x^\dagger} \tilde{f}(x^\dagger|x)}_{=0} \end{aligned}$$

where

$$\tilde{f}(x^\dagger|x) := f(x^\dagger|x,0) - f(x^\dagger|x,1).$$

$$\begin{aligned} P_\theta(1|x) &= \frac{1}{1 + \exp(\mathrm{v}_\theta(x,0) - \mathrm{v}_\theta(x,1))} \\ &= \frac{1}{1 + \exp\left(\theta_0 + \theta_1 x_t - \beta \sum_{x^\dagger} \log \mathrm{P}(1|x^\dagger) \tilde{f}(x^\dagger|x)\right)}. \end{aligned}$$

CCP without Finite Dependence

$$[\Psi(\theta, P, f)](d = 1|x) = \frac{1}{1 + \exp(\tilde{v}_\theta(x))}$$

$$\tilde{v}_\theta(x) = \tilde{u}_\theta(x) + \beta \sum_{x'} \log P(1|x')\tilde{f}(x'|x)$$

$$+ \beta^2 \sum_{x''} V(x'') \sum_{x'} f(x''|x', 1)\tilde{f}(x'|x)$$

But:

$$\tilde{v}_\theta(x, d) = \tilde{u}_\theta(x, d) + \beta \sum_{x'} V(x')\tilde{f}(x'|x, d)$$

where $V(x') = w(v_\theta(x', 1) - \log P(1|x'))$

$$+ (1 - w)(v_\theta(x', 0) - \log P(0|x'))$$

$$= u_\theta^w(x') - p^w(x') + \sum_{x''} V(x'')f^w(x''|x')$$

where $p^w(x) = w \log P(1|x) + (1 - w) \log P(0|x)$

Near Finite Dependence

- **Goal:** Select weights $\{w_{x,d}(x^\dagger, d^\dagger)\}$ to satisfy for each (x, d):

$$\sum_{x^\dagger} \left(\sum_{d^\dagger \in \{0,1\}} w_{x,d}(x^\dagger, d^\dagger) f(x^{\dagger\dagger} | x^\dagger, d^\dagger) \right) \tilde{f}(x^\dagger | x, d) \approx 0.$$

- **Advancement:** In [Arcidiacono and Miller, 2019], weights $\{w(x^\dagger, d^\dagger)\}$ are not (x, d)-specific.
- **Impact:** This choice grants $(|XD|)^2$ degrees of freedom—versus $(|XD|)$—for finer norm minimization.

Our proposed estimator

$$\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \ln[\Psi_{AFD}(\theta, \hat{P}, \hat{f})](d_i | x_i).$$

$$[\Psi_{AFD}(\theta, P, f)](d = 1|x) = \frac{1}{1 + \exp\left(\mathrm{v}_\theta(x, 0) - \mathrm{v}_\theta(x, 1)\right)}$$

where,

$$\begin{aligned}
\mathrm{v}_\theta(x, 0) - \mathrm{v}_\theta(x, 1) = & \ u_\theta(x, 0) - u_\theta(x, 1) \\
& + \beta \sum_{x^\dagger} \left(u_\theta^{w_{x,x^\dagger}}(x^\dagger) - p^{w_{x,x^\dagger}}(x^\dagger) \right) \tilde{f}(x^\dagger | x) \\
& + \beta^2 \sum_{x^{\dagger\dagger}} \mathrm{V}(x^{\dagger\dagger}) \underbrace{\sum_{x'} f^{w_{x,x^\dagger}}(x^{\dagger\dagger} | x^\dagger) \tilde{f}(x^\dagger | x)}_{\approx 0}
\end{aligned}$$

Computing $\mathbf{w}$ for 1-period finite dependence I

The objective is to minimize the weights, which can be represented as the following:

$$\check{\mathbf{W}}_{t+1}\tilde{\mathbf{F}}_{t+1} + \tilde{\mathbf{F}}_t \mathbf{F}_{0,t+1} = 0,$$

Here, $\tilde{\mathbf{F}}$ represents the difference between two Markov transition matrices:

$$\tilde{\mathbf{F}} = \mathbf{F}_{1,t} - \mathbf{F}_{0,t}$$

where

$$\mathbf{F}_d := \begin{bmatrix} f_t(x^{(1)}_{\tau+1}|x^{(1)}_{\tau}, d) & \cdots & f_t(x^{(X)}_{\tau+1}|x^{(1)}_{\tau}, d) \\ \vdots & \ddots & \vdots \\ f_t(x^{(1)}_{\tau+1}|x^{(X)}_{\tau}, d) & \cdots & f_t(x^{(X)}_{\tau+1}|x^{(X)}_{\tau}, d) \end{bmatrix} \quad \text{for } d = 1, 0$$

and

$$\check{\mathbf{W}}_{t+1} = \begin{bmatrix} \tilde{f}(x^{(1)}_{t+1}|x^{(1)}_{t}, 1)\mathrm{w}_{t+1}(x^{(1)}_{t+1}|x^{(1)}_{t}) & \cdots & \tilde{f}(x^{(X)}_{t+1}|x^{(1)}_{t}, 1)\mathrm{w}_{t+1}(x^{(X)}_{t+1}|x^{(1)}_{t}) \\ \vdots & \ddots & \vdots \\ \tilde{f}(x^{(1)}_{t+1}|x^{(X)}_{t}, 1)\mathrm{w}_{t+1}(x^{(1)}_{t+1}|x^{(X)}_{t}) & \cdots & \tilde{f}(x^{(X)}_{t+1}|x^{(X)}_{t}, 1)\mathrm{w}_{t+1}(x^{(X)}_{t+1}|x^{(X)}_{t}) \end{bmatrix}.$$

Computing $\mathbf{w}$ for 1-period finite dependence II

We derive a closed-form solution for $\mathbf{w}$, as shown below:

$$\check{\mathbf{W}}_{t+1} = -\tilde{\mathbf{F}}_t \mathbf{F}_{0,t+1} (\tilde{\mathbf{F}}_{t+1})^+ \qquad (1)$$

where $(\tilde{\mathbf{F}})^+_{t+1}$ denotes the Moore-Penrose pseudo-inverse of $\tilde{\mathbf{F}}_{t+1}$.
Let $\tilde{\mathbf{F}}^{(1)}(\check{\mathbf{W}}_{t+1}) = \check{\mathbf{W}}_{t+1}\tilde{\mathbf{F}}_{t+1} + \tilde{\mathbf{F}}_t \mathbf{F}_{0,t+1}$,
the value difference $\mathrm{v}_1 - \mathrm{v}_0$ is written as

$$\tilde{\mathbf{v}}_t = \tilde{\mathbf{u}}_t + \beta(\check{\mathbf{W}}_{t+1}\tilde{\mathbf{u}}_{t+1} + \tilde{\mathbf{u}}_t \mathbf{F}_{0,t+1}) + \beta^2 \tilde{\mathbf{F}}^{(1)}(\check{\mathbf{W}}_{t+1})\mathrm{V}_{t+2}$$

the norm of $\tilde{\mathbf{F}}^{(1)}$ is the impact of V_{t+2} on the current CCP P_t.

Extend to ρ-period: Initial Steps

Sequential Approach to Simplify Weight Minimization:

1. Obtain initial weights $\check{\mathbf{W}}^{(1)}_{t+1}$ by solving:

$$\check{\mathbf{W}}^{(1)}_{t+1} = \arg\min_{\mathbf{W}} \mathbf{W}\tilde{\mathbf{F}}_{t+1} + \tilde{\mathbf{F}}_{t}\mathbf{F}_{0,t+1}$$

2. Minimize future decision influences by deriving $\check{\mathbf{W}}^{(2)}_{t+2}$:

$$\check{\mathbf{W}}^{(2)}_{t+2} = \arg\min_{\mathbf{W}} \mathbf{W}\tilde{\mathbf{F}}_{t+2} + \tilde{\mathbf{F}}^{(1)}(\check{\mathbf{W}}_{t+1})\mathbf{F}_{0,t+2}$$

Impact and Optimization in ρ-period

Evaluating the Extended Impact:
Value difference $\mathrm{v}_1 - \mathrm{v}_0$ and norm of $\tilde{\mathbf{F}}^{(2)}$:

$$\tilde{\mathbf{v}}_t = \tilde{\mathbf{u}}_t + \beta(\check{\mathbf{W}}_{t+1}\tilde{\mathbf{u}}_{t+1} + \tilde{\mathbf{u}}_t\mathbf{F}_{0,t+1}) + \beta^2\left(\check{\mathbf{W}}^{(2)}_{t+2}\tilde{\mathbf{u}}_{t+2} + \tilde{\mathbf{F}}^{(1)}(\check{\mathbf{W}}^{(1)}_{t+1})\mathbf{u}_{0,t+2}\right) + \beta^3\tilde{\mathbf{F}}^{(2)}(\check{\mathbf{W}}^{(1)}_{t+1}, \check{\mathbf{W}}^{(2)}_{t+2})\mathrm{V}_{t+3}$$

the norm of $\tilde{\mathbf{F}}^{(2)}$ is the impact of V_{t+2} on the current CCP P_t.

Optimization of 2-Period Finite Dependence I

- The second-period weights $\mathbf{W}^{(2)}$ are functionally dependent on $\mathbf{W}^{(1)}$, enabling effective optimization over two periods:

$$\mathbf{W}^{(2)}(\mathbf{W}^{(1)})$$

- The resulting expression for $\tilde{\mathbf{F}}^{(2)}$ is:

$$\begin{aligned}\tilde{\mathbf{F}}^{(2)}(\check{\mathbf{W}}^{(1)}_{t+1}, \check{\mathbf{W}}^{(2)}_{t+2}) &= \left(-\left(\check{\mathbf{W}}_{t+1}\tilde{\mathbf{F}} + \tilde{\mathbf{F}}\mathbf{F}_0\right)\mathbf{F}_0\tilde{\mathbf{F}}^{+}\tilde{\mathbf{F}} + \check{\mathbf{W}}_{t+1}\tilde{\mathbf{F}}\mathbf{F}_0 + \tilde{\mathbf{F}}\mathbf{F}_0^2\right) \\ &= \mathbf{w}_{t+1}\tilde{\mathbf{F}}\mathbf{F}_0\left(\mathbf{I} - \tilde{\mathbf{F}}^{+}\tilde{\mathbf{F}}\right) + \tilde{\mathbf{F}}\mathbf{F}_0^2\left(\mathbf{I} - \tilde{\mathbf{F}}^{+}\tilde{\mathbf{F}}\right).\end{aligned}$$

Optimization of 2-Period Finite Dependence II

- Define the projection matrix onto the null space of $\tilde{\mathbf{F}}$:

$$\mathcal{P}_{\tilde{\mathbf{F}}} = \left(\mathbf{I} - \tilde{\mathbf{F}}^{+}\tilde{\mathbf{F}}\right).$$

- In terms of optimality, when determining the 1-period weight matrix, the first-period weight matrix should be selected optimally as:

$$\check{\mathbf{W}}^{(2)*}_{t+1} = -\tilde{\mathbf{F}}\mathbf{F}_0^2\mathcal{P}_{\tilde{\mathbf{F}}}\left(\tilde{\mathbf{F}}\mathbf{F}_0\mathcal{P}_{\tilde{\mathbf{F}}}\right)^{+}.$$

$$\tilde{\mathbf{F}}^{(2)} = \tilde{\mathbf{F}}\mathbf{F}_0\mathbf{F}_0\mathcal{P}_{\tilde{\mathbf{F}}}\left(\mathbf{I} - (\tilde{\mathbf{F}}\mathbf{F}_0\mathcal{P}_{\tilde{\mathbf{F}}})^{+}(\tilde{\mathbf{F}}\mathbf{F}_0\mathcal{P}_{\tilde{\mathbf{F}}})\right).$$

Optimality and Singular Value Decomposition I

- **Singular Value Decomposition (SVD) of $\tilde{\mathbf{F}}$:**
  - Let $\tilde{\mathbf{F}} \in \mathbb{R}^{(D)X \times X}$.
  - SVD: $\tilde{\mathbf{F}} = \mathbf{U}_{\tilde{F}} \mathbf{S}_{\tilde{F}} \mathbf{V}_{\tilde{\mathbf{F}}}^{\top}$,
  - Components:
    - ★ $\mathbf{U}_{\tilde{F}} \in \mathbb{R}^{(D)X \times (D)X}$: Left singular vectors.
    - ★ $\mathbf{S}_{\tilde{F}} \in \mathbb{R}^{(D)X \times X}$: Diagonal matrix with singular values.
    - ★ $\mathbf{V}_{\tilde{\mathbf{F}}} \in \mathbb{R}^{X \times X}$: Right singular vectors.
  - Rank and partitioning of $\mathbf{S}_{\tilde{F}}$:

$$\mathbf{S}_{\tilde{F}} = \begin{bmatrix} \mathbf{S}_{\tilde{F},00} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{S}_{\tilde{F},00}$ is a rank matrix.

Optimality and Singular Value Decomposition II

- Transformation Using Unitary Matrices
  - Define the transformation matrix $\mathbf{S}_0$:

$$\mathbf{S}_0 = \mathbf{V}_{\tilde{\mathbf{F}}}^\top \mathbf{F}_0 \mathbf{V}_{\tilde{\mathbf{F}}}$$

  - Since $\mathbf{V}_{\tilde{\mathbf{F}}}^\top$ is unitary, we relate $\mathbf{F}_0$ back to the original space:

$$\mathbf{F}_0 = \mathbf{V}_{\tilde{\mathbf{F}}} \mathbf{S}_0 \mathbf{V}_{\tilde{\mathbf{F}}}^\top$$

  - Partitioning of $\mathbf{S}_0$ into submatrices:

$$\mathbf{S}_0 = \begin{bmatrix} \underbrace{\mathbf{S}_{0,00}}_{\text{rank}(\tilde{\mathbf{F}}) \times \text{rank}(\tilde{\mathbf{F}})} & \underbrace{\mathbf{S}_{0,01}}_{\text{rank}(\tilde{\mathbf{F}}) \times \text{nullity}(\tilde{\mathbf{F}})} \\ \underbrace{\mathbf{S}_{0,10}}_{\text{nullity}(\tilde{\mathbf{F}}) \times \text{rank}(\tilde{\mathbf{F}})} & \underbrace{\mathbf{S}_{0,11}}_{\text{nullity}(\tilde{\mathbf{F}}) \times \text{nullity}(\tilde{\mathbf{F}})} \end{bmatrix}$$

Optimality and Singular Value Decomposition III

- **Projection Matrix $\mathcal{P}_{\tilde{\mathbf{F}}}$:**
  - Utilizing Markov matrix properties and the non-full rank of $\tilde{\mathbf{F}}$.
  - Projection onto the null space of $\tilde{\mathbf{F}}$:

$$\mathcal{P}_{\tilde{\mathbf{F}}} = \mathbf{V}_{\tilde{\mathbf{F}}} \Sigma_{\mathrm{Null}(\tilde{F})} \mathbf{V}_{\tilde{\mathbf{F}}}^{\top},$$

$$\text{where } \boldsymbol{\Sigma}_{\mathrm{Null}(\tilde{F})} = \mathrm{diag}\left(\underbrace{0, \ldots, 0}_{\mathrm{rank}(\tilde{\mathbf{F}}) \text{ zeros}}, 1, \ldots, 1 \right).$$

Characterization of the Transition Difference

By re-organizing the matrix, we show that:

- $\tilde{\mathbf{F}}^{(1)} = \mathbf{U}_{\tilde{F}} \begin{bmatrix} \mathbf{0} & \mathbf{S}_{\tilde{F},00}\mathbf{S}_{0,01} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}_{\tilde{\mathbf{F}}}^{\top}.$
- $\tilde{\mathbf{F}}^{(2)} = \mathbf{U}_{\tilde{F}} \begin{bmatrix} \mathbf{0} & \mathbf{S}_{\tilde{F},00}\left(\mathbf{S}_{0,01}\mathbf{S}_{0,11}(\mathbf{I} - \mathbf{S}_{0,01}^{+}\mathbf{S}_{0,01})\right) \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}_{\tilde{\mathbf{F}}}^{\top}.$

Finite Dependence Propositions

- **Proposition for 1-Period Dependence:**

$$\text{If } \mathbf{S}_{0,01} = 0, \text{ then } \tilde{\mathbf{F}}^{(1)} = 0.$$

  - This condition is rarely met, except in special cases such as terminal or absorbing states.

- **Proposition for 2-Period Dependence:**

$$\text{If } \mathbf{S}_{0,01}\mathbf{S}_{0,11}(\mathbf{I} - \mathbf{S}^{+}_{0,01}\mathbf{S}_{0,01}) = 0, \text{ then } \tilde{\mathbf{F}}^{(2)} = 0.$$

  - This condition is generally satisfied across all model classes, encompassing scenarios like terminal or absorbing states.
  - It also includes cases where $\mathbf{S}_{0,01}$ has full rank.

Dimensionality Reduction

Objective: Preventing the curse of dimensionality in sequential mapping by avoiding recursive multiplication of large matrices.
Strategy: Partition the state into two parts:

- ω - states affected by the decision
- z - exogenous state unaffected by the decision

By defining $\mathbf{w}^{(1)}$:

$$\mathbf{w}^{(1)} = -\tilde{\mathbf{F}}\mathbf{F}_0(\tilde{\mathbf{F}})^{+} = \underbrace{-(\tilde{\mathbf{F}}_\omega \mathbf{F}_{\omega,0} \tilde{\mathbf{F}}_\omega^{+})}_{\mathbf{w}_\omega^{(1)}} \otimes \mathbf{F}_z,$$

$$\tilde{\mathbf{F}}\mathbf{F}(\mathbf{w}) = \tilde{\mathbf{F}}_\omega \mathbf{F}_{\omega,0}(\mathbf{I} - (\tilde{\mathbf{F}}_\omega^{+}\tilde{\mathbf{F}}_\omega)) \otimes \mathbf{F}_z.$$

Simulation Overview

- Dynamic entry/exit problem based on [Aguirregabiria and Magesan, 2016], incorporating history in firm profits with vector θ.
- State variables $x = (z_1, z_2, z_3, z_4, \omega, y)$ represent market conditions and firm-specific factors; firms decide to operate or exit based on (x, ϵ).
- Flow payoff $u(d_t, x_t; \theta)$:

$$
\begin{aligned}
u(d_t, x_t; \theta) &= d_t(VP_t - EC_t - FC_t) \\
\text{where } VP_t &= \exp(\omega)[\theta_0^{VP} + \theta_1^{VP} z_{1t} + \theta_2^{VP} z_{2t}] \\
FC_t &= [\theta_0^{FC} + \theta_1^{FC} z_{3t}] \\
EC_t &= (1 - y_t)[\theta_0^{EC} + \theta_1^{EC} z_{4t}].
\end{aligned}
$$

- Shocks are AR(1) processes, leading to a state space of dimension $X = 2 \cdot K_z^4 \cdot K_o$ with action-independent transitions.

Transition Probability Formulas

1. Entry cost and fixed cost shocks (z_{jt}): Independent of chosen action

$$f(z_j'|z_j)=\begin{cases}\Phi([z_j^{(1)}+(\omega_j^{(1)}/2)-\gamma_0^j-\gamma_1^j z]/\sigma_j) & z' = z_j^{(1)};\\ 1-\Phi([z_j^{(K-1)}+(\omega_j^{(K-1)}/2)-\gamma_0^j-\gamma_1^j z]/\sigma_j) & z' = z_j^{(K)}.\\ \Phi([z_j^{(k)}+(\omega_j^{(k)}/2)-\gamma_0^j-\gamma_1^j z]/\sigma_j)- & \\ \quad \Phi([z_j^{(k-1)}+(\omega_j^{(k-1)}/2)-\gamma_0^j-\gamma_1^j z]/\sigma_j) & \text{otherwise;}\end{cases}$$

2. Productivity shock (ω_t): Dependent on chosen action

$$f(\omega'|\omega,d)=\begin{cases}\Phi([\omega^{(1)}+(\omega^{(1)}/2)-\gamma_0^\omega-\gamma_1^\omega\omega-\gamma_d d]/\sigma) & \omega' = \omega^{(1)};\\ 1-\Phi([\omega^{(K-1)}+(\omega^{(K-1)}/2)-\gamma_0^\omega-\gamma_1^\omega\omega-\gamma_d d]/\sigma) & \omega' = \omega^{(K)}.\\ \Phi([\omega^{(k)}+(\omega^{(k)}/2)-\gamma_0^\omega-\gamma_1^\omega\omega-\gamma_d d]/\sigma)- & \\ \quad \Phi([\omega^{(k-1)}+(\omega^{(k-1)}/2)-\gamma_0^\omega-\gamma_1^\omega\omega-\gamma_d d]/\sigma) & \text{otherwise}\end{cases}$$

Weight Solving

N state	γ_a	Time				Singular Value	
		HM Inverse	$\mathbf{w}$	$\mathbf{H}(\cdot)$	$\mathbf{h}(\cdot)$	$\tilde{\mathbf{F}}^{(1)}$	$\tilde{\mathbf{F}}^{(2)}$
5184	0	11.250	0.050	0.120	0.100	1.93E-17	4.15e-17
5184	0.8	17.980	0.080	0.140	0.120	0.048	6.35e-17
5184	1.2	12.580	0.110	0.210	0.200	0.111	5.10e-17
7776	0	40.900	0.180	0.450	0.440	1.41e-16	1.88e-17
7776	0.8	55.520	0.140	0.320	0.250	0.164	8.32e-17
7776	1.2	50.890	0.140	0.320	0.280	0.341	5.59e-17
10368	0	123.260	0.210	0.560	0.480	3.68E-16	7.35e-17
10368	0.8	128.190	0.220	0.540	0.410	0.184	6.26e-16
10368	1.2	124.260	0.180	0.430	0.380	0.420	1.58e-15
12960	0	210.980	0.270	0.840	0.620	2.78E-16	5.06e-17
12960	0.8	244.600	0.330	0.800	0.680	0.197	2.64e-13
12960	1.2	245.770	0.300	0.680	0.600	0.411	6.90e-13

Define the sequence $\{\tilde{\mathbf{F}}^{(k)}\}_{k=0}^{\rho}$ in $\mathcal{F}$ recursively by

$$\tilde{\mathbf{F}}^{(k)} \equiv \begin{cases} \tilde{\mathbf{F}}, & \text{if } k = 0, \\ \tilde{\mathbf{F}}^{(k-1)}\mathbf{F}(\mathbf{w}^{(k)}), & \text{for } k = 1, 2, \ldots, \rho, \end{cases} \tag{2}$$

Parameter Estimates

Table: Non-stationary Model: $nM = 500, nT = 4, nMC = 100$

	VP0	VP1	VP2	FC0	FC1	EC0	EC1	Time	ρ_1	ρ_2
True θ	0.5	1.0	-1.0	0.5	1.0	1.0	1.0			
					X=2560, $\gamma_a = 0$					
NFD	0.501	1.009	-1.013	0.488	1.019	1.017	1.004	0.45	6.83E-17	9.50E-17
	(0.153)	(0.085)	(0.086)	(0.212)	(0.081)	(0.230)	(0.085)	0.40		
NFD2	0.500	1.001	-1.005	0.485	1.016	1.022	1.003	1.03	1.39E-16	3.17E-16
	(0.065)	(0.044)	(0.043)	(0.154)	(0.071)	(0.223)	(0.084)	0.99		
NFD	0.495	1.014	-1.017	0.451	0.997	1.029	0.993	27111.56	1.55E-16	2.55E-16
(no x_t-specific)	(0.152)	(0.085)	(0.086)	(0.214)	(0.078)	(0.232)	(0.085)	27111.52		
					X=2560, $\gamma_a = 1$					
NFD	0.510	1.010	-1.012	0.266	0.943	0.978	0.991	0.42	2.49E-01	2.52E-01
	(0.071)	(0.049)	(0.051)	(0.277)	(0.085)	(0.239)	(0.086)	0.38		
NFD2	0.526	1.007	-1.015	0.523	1.012	0.984	1.003	0.97	5.57E-10	1.50E-09
	(0.178)	(0.117)	(0.116)	(0.310)	(0.126)	(0.254)	(0.111)	0.93		
NFD	0.443	1.080	-1.087	0.059	0.955	0.984	1.029	7511.62	4.94E-01	5.92E-01
(no x_t-specific)	(0.174)	(0.126)	(0.128)	(0.500)	(0.098)	(0.245)	(0.105)	7511.57		

Parameter Estimates

Table: Non-stationary Model: $nM = 500, nT = 4, nMC = 100$

	VP0	VP1	VP2	FC0	FC1	EC0	EC1	Time	ρ_1	ρ_2
True θ	0.5	1.0	-1.0	0.5	1.0	1.0	1.0			
					X=2560, $\gamma_a = 0$					
NFD	0.501	1.009	-1.013	0.488	1.019	1.017	1.004	0.45	6.83E-17	9.50E-17
	(0.153)	(0.085)	(0.086)	(0.212)	(0.081)	(0.230)	(0.085)	0.40		
NFD2	0.500	1.001	-1.005	0.485	1.016	1.022	1.003	1.03	1.39E-16	3.17E-16
	(0.065)	(0.044)	(0.043)	(0.154)	(0.071)	(0.223)	(0.084)	0.99		
NFD	0.495	1.014	-1.017	0.451	0.997	1.029	0.993	27111.56	1.55E-16	2.55E-16
(no x_t-specific)	(0.152)	(0.085)	(0.086)	(0.214)	(0.078)	(0.232)	(0.085)	27111.52		
					X=2560, $\gamma_a = 1$					
NFD	0.510	1.010	-1.012	0.266	0.943	0.978	0.991	0.42	2.49E-01	2.52E-01
	(0.071)	(0.049)	(0.051)	(0.277)	(0.085)	(0.239)	(0.086)	0.38		
NFD2	0.526	1.007	-1.015	0.523	1.012	0.984	1.003	0.97	5.57E-10	1.50E-09
	(0.178)	(0.117)	(0.116)	(0.310)	(0.126)	(0.254)	(0.111)	0.93		
NFD	0.443	1.080	-1.087	0.059	0.955	0.984	1.029	7511.62	4.94E-01	5.92E-01
(no x_t-specific)	(0.174)	(0.126)	(0.128)	(0.500)	(0.098)	(0.245)	(0.105)	7511.57		

Non-Stationary Model

- We can make the model non-stationary by allowing γ_a to take different values across periods.
- The productivity shock $\tilde{\omega}_t$ as a function of past actions, following the process

$$\tilde{\omega}_t = \gamma^{\omega}_{0t} + d_{t-1}\gamma_a + \gamma^{\omega}_1 \tilde{\omega}_{t-1} + e_{jt},$$

  where the disturbance term e_{jt} is independently and identically distributed as $\mathcal{N}(0, \sigma_j^2)$.
  - **Stationary** context, γ^{ω}_{0t} is set to zero for all t.
  - **non-stationary** scenario γ^{ω}_{0t} varies with time, assuming values $[-0.8, 0.8, 0, -0.3]$ for $t = 1, 2, 3, 4$, respectively.

Key Contributions

- Developed a closed-form solution for weight characterization and formulated a new estimator based on it.
- We propose a new estimator that leverages this weight characterization that can be used to estimate non-stationary models.
- Propose the proposition to characterize 2-period finite dependence.
- Demonstrated the estimator's performance via Monte Carlo simulation.

References I

Aguirregabiria, V. and Magesan, A. (2016).
Solution and Estimation of Dynamic Discrete Choice Structural Models Using Euler Equations.
SSRN Electronic Journal.

Arcidiacono, P. and Miller, R. A. (2011).
Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity.
Econometrica, 79(6):1823–1867.

Arcidiacono, P. and Miller, R. A. (2019).
Nonstationary dynamic models with finite dependence.
Quantitative Economics, 10(3):853–890.