Sequential Estimation of Structural Models with a Fixed Point Constraint

Hiro Kasahara[1] Katsumi Shimotsu[2]

[1]Department of Economics
University of British Columbia

[2]Department of Economics
Hitotsubashi University

Outline

Introduction: structural economic models

- Many markets are characterized as a competition among a few firms with differentiated products
- Between monopoly and perfect competition
- In such markets, strategic interaction between firms become important.
- Many economists want to understand how firms compete with each other in such markets.
- Implications for competition policy, entry regulation, etc.
- "Structural economic models" explicitly model firms' strategic interaction (and dynamic choice)

Structural model 1: game-theoretic model of entry

- Two potential entrants to a market (for example, 5th generation smartphone, Docomo and KDDI).
- If both firms enter, both firms earn positive profit. If only one firm enters, it earns larger profit.
- Payoff matrix

		firm 2	
		in	out
firm 1	in	$(1,1)$	$(3,0)$
	out	$(0,3)$	$(0,0)$

- Nash equilibrium: given the other player's action, I have no incentive to change my action $\Rightarrow$ (in,in)

Games with private information

- But firm 1 may not know everything about firm 2.
- Payoff matrix with a random component

		firm 2	
		in	out
firm 1	in	$(\varepsilon_1 - \theta, \varepsilon_2 - \theta)$	$(\varepsilon_1, 0)$
	out	$(0, \varepsilon_2)$	$(0, 0)$

  ε_1 and ε_2 are drawn independently from uniform distribution$[0, 1]$. Both firms know θ.

- Only firm 1 observes ε_1, and only firm 2 observes ε_2 (private information).

How to determine the equilibrium

		firm 2	
		in	out
firm 1	in	$(\varepsilon_1 - \theta, \varepsilon_2 - \theta)$	$(\varepsilon_1, 0)$
	out	$(0, \varepsilon_2)$	$(0, 0)$

- Suppose firm 1 thinks that firm 2 enters with probability P_2 = firm 1's subjective probability ("belief")
- Firm 1 enters the market if $P_2(\varepsilon_1 - \theta) + (1 - P_2)\varepsilon_1 > 0 \Rightarrow \varepsilon_1 > \theta P_2$
- Because $\varepsilon_1 \sim U[0, 1]$, firm 1 enters the market with probability $1 - \theta P_2$ when his belief is P_2.

How to determine the equilibrium

- Firm 1 enters the market with probability $1 - \theta P_2$ when his belief is P_2.
- Firm 2 enters the market with probability $1 - \theta P_1$ when his belief is P_1.
- Bayesian perfect equilibrium: both players' belief must be consistent with their action.

$$P_1 = 1 - \theta P_2, \quad P_2 = 1 - \theta P_1 \Rightarrow P_1 = P_2 = 1/(1+\theta)$$

- Best response mapping: $[0,1]^2 \to [0,1]^2$
  $\Psi(\theta, P) = (1 - \theta P_2, 1 - \theta P_1)'$.
  Then BPE is a fixed point of $\Psi : P^* = \Psi(\theta, P^*)$

Estimation of θ in this model

- Suppose we have iid data of the entry decision (a_{1i}, a_{2i}) for $i = 1, \ldots, n$, and we want to estimate the value of θ.
- If we assume the data are in the equilibrium, we can estimate θ by

$$\hat{\theta} = \frac{1}{2}\left[\frac{1-\hat{P}_1}{\hat{P}_1} + \frac{1-\hat{P}_2}{\hat{P}_2}\right].$$

Do people actually behave according to the game theory?

Many experimental results

See, for example, The Handbook of Experimental Economics (1995), John H. Kagel and Alvin E. Roth, eds., Princeton University Press.

- Overall conclusion: yes and no
- Game of market: theoretical monopoly price $>$ price in experiments $>$ theoretical no-monopoly price
- Provision of public goods: less free-riding than predicted by theory

A game may have multiple equilibria

- Payoff matrix with the entry cost of -2

		firm 2	
		in	out
firm 1	in	$(-1,-1)$	$(1,0)$
	out	$(0,1)$	$(0,0)$

- Nash equilibria: (in,out), (out,in)

- We assume $\Psi(\theta, P)$ has a unique fixed point for the moment.

Structural model 2: dynamic discrete game

- N firms = potential entrants
- Entry/exit choice: $a_{it} \in A = \{0, 1\}$.
- Firm i's profit in period t:

$$\tilde{\Pi}_i(a_t, S_t, a_{i,t-1}, \epsilon_t; \theta)$$

  - All the firms' current decision: $a_t = (a_{1t}, ..., a_{Nt})'$
  - Market demand condition: S_t (observable)
  - Past entry decision: $a_{i,t-1}$
  - Private shocks: $\epsilon_t = (\epsilon_{1t}, ..., \epsilon_{Nt})'$.
- θ: parameter of interest

Dynamic discrete game example (continued)

- Profit function of firm *i*:

$$\tilde{\Pi}_i(a_t, S_t, a_{t-1}, \epsilon_{it}; \theta) = \theta_{RS} \ln S_t - \theta_{FC,i} - \theta_{EC}(1 - a_{i,t-1}) - \theta_{RN} \ln(1 + \sum_{j \neq i} a_{jt}) \; + \; \epsilon_{it1}$$

- θ_{RS}: Revenue parameter
- $\theta_{FC,i}$: Operating cost
- θ_{EC}: Entry cost
- θ_{RN}: Degree of strategic substitution

Structural model 2: dynamic discrete game

- Dynamic optimization by firm i

$$\max_{a_{i1}, a_{i2}, \ldots} \quad E\left[\sum_{t=0}^{\infty} \beta^t \Pi_i(a_t, S_t, a_{i,t-1}, \epsilon_t; \theta) \middle| S_t, a_{t-1}; \theta\right]$$

- Assume the state variable follows a Markov process
- Markov decision problem given his belief
- Stationary solution

Structural model 2: dynamic discrete game

- Empirical implication: firm i's conditional choice probabilities = $P_i(a|x)$.

- $P_i = \{P_i(a|x)\}_{(a,x)}$: firm i's conditional choice probabilities for all possible x

- The conditional choice probabilities of all the firms: $P = (P_1, \ldots, P_N)$

- For a given θ, an equilibrium is characterized by a fixed point of the best response mapping

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

P can be large

For example,

- 5 potential entrants
- all the firms' entry status in the previous period: $2^5 = 32$ support points
- market condition takes 10 different values
- x_t takes $10 \times 32 = 320$ different values
- Length of $P = 320 \times 5 = 1600$

$\Rightarrow$ finding a fixed point of Ψ can be computationally costly

Economic models with a fixed point constraint

- When $P = P(a|x)$ is the choice probability of a discrete action a conditional on x, the log-likelihood function is

$$Q_n(\theta) = \sum_{i=1}^{n} \ln P(a_i|x_i) \quad s.t. \quad P = \Psi(\theta, P)$$

Approach 1: nested fixed point approach (Rust, 1987)

$$\hat{\theta}_{ML} = \arg\max_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \ln P(a_i | x_i)$$

$$s.t. \qquad P = \Psi(\theta, P)$$

Starting from an initial value of θ,

- (Inner Loop) For each candidate parameter θ, find the fixed point: $P_\theta = \Psi(\theta, P_\theta)$.
- (Outer Loop) Maximize the log-likelihood function by a quasi-Newton method.

This algorithm produces $\hat{\theta}_{ML}$ but is very computationally intensive.

Approach 2: constraint optimization approach (Su and Judd, 2010)

- Write down the Lagrangian

$$\mathcal{L}(\theta, P, \lambda) = \frac{1}{n}\sum_{i=1}^{n} \ln P\left(a_i | x_i\right) + \lambda(P - \Psi\left(\theta, P\right))$$

- Solve the first-order condition

$$\begin{aligned} \nabla_\theta \mathcal{L}(\theta^*, P^*, \lambda^*) &= 0, \\ \nabla_P \mathcal{L}(\theta^*, P^*, \lambda^*) &= 0, \\ P^* - \Psi(\theta^*, P^*) &= 0 \end{aligned}$$

- According to the authors, one can solve this problem using the “NEOS Server, a free internet service which gives the user access to several state-of-the-art solvers."
- Applications?

Approach 3: NPL estimator (Aguirregabiria and Mira, 2007)

- $\tilde{P}_0$: initial guess of P
- **Step 1:** Given $\tilde{P}_{k-1}$,

$$\frac{1}{n}\sum_{i=1}^{n} \ln[\Psi(\theta, \tilde{P}_{k-1})](a_i|x_i).$$

  can be viewed as a pseudo log-likelihood function. So, estimate θ by maximizing this objective function $\Rightarrow \tilde{\theta}_k$

- **Step 2:** Given $\tilde{\theta}_k$, update the estimate of P by

$$\tilde{P}_k = \Psi(\tilde{\theta}_k, \tilde{P}_{k-1}).$$

- Iterate Steps 1-2: $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$

NPL Fixed Points: $\{\check{\theta}, \check{P}\}$

- If the sequence $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$ converges, its limit satisfies

$$\begin{aligned} \check{\theta} &= \arg\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ln[\Psi(\theta, \check{P})](a_i|x_i), \\ \check{P} &= \Psi(\check{\theta}, \check{P}). \end{aligned}$$

- The NPL estimator is defined as $(\check{\theta}, \check{P})$ that achieves the highest pseudo-likelihood value.
- Easy to implement: standard optimization and policy iteration

What empirical researchers want

- Estimation methods that work
- Easy-to implement, easy-to understand methods
- Not all the economists/econometricians want to spend a lot of time on writing computer codes.

Convergence of the NPL algorithm?

- Convergence of $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$?
- Empirical researchers report non-convergence of the NPL algorithm.
- For example, it can exhibit a 2-period cycle.
- Little is known when it works and when it doesn't, except for a few counter-examples.

Insight: a population NPL sequence

- Suppose we have infinitely many observations.

$$\tilde{\theta}_k = \tilde{\theta}(\tilde{P}_{k-1}) = \arg\max_{\theta} E \ln[\Psi(\theta, \tilde{P}_{k-1})](a_i|x_i).$$

$$\tilde{P}_k = \Psi(\tilde{\theta}(\tilde{P}_{k-1}), \tilde{P}_{k-1}) \equiv \phi_0(\tilde{P}_{k-1}).$$

- True probability P^0 is an NPL fixed point:

$$P^0 = \phi_0(P^0).$$

- $|$ dominant eigenvalue of $(\nabla_{P'} \phi_0(P)|_{P=P^0})\,| < 1$
  $\Rightarrow$ stability at P^0.

Property of NPL updating (Kasahara and Shimotsu, 2010)

- In a neighborhood of P^0,

$$\tilde{\theta}_j - \hat{\theta}_{NPL} = O_p(||\tilde{P}_{j-1} - \hat{P}_{NPL}||),$$

$$\tilde{P}_j - \hat{P}_{NPL} = M_{\Psi_\theta}\Psi_P(\tilde{P}_{j-1} - \hat{P}_{NPL})$$

$$+ O_p(n^{-1/2}||\tilde{P}_{j-1} - \hat{P}_{NPL}||) + O_p(||\tilde{P}_{j-1} - \hat{P}_{NPL}||^2),$$

  where $M_{\Psi_\theta} = I - \Psi_\theta(\Psi'_\theta\Delta_P\Psi_\theta)^{-1}\Psi'_\theta\Delta_P$

- $\nabla_{P'}\phi_0(P)|_{P=P^0} = M_{\Psi_\theta}\Psi_P$.

Property of NPL updating (Kasahara and Shimotsu, 2010)

- The convergence of $\tilde{P}_k$ depends on the eigenvalues of $M_{\Psi_\theta}\Psi_P$.
- The dominant eigenvalue of AB is not necessarily bounded by that of A and B.
- But $\rho(M_{\Psi_\theta}\Psi_P) \approx \rho(\Psi_P)$ holds if $\dim(P) \gg \dim(\theta)$.

Dynamic Game Example (continued)

- Profit function of firm i:

$$\tilde{\Pi}_i(a_t, S_t, a_{t-1}, \epsilon_{it}; \theta) = \theta_{RS} \ln S_t - \theta_{FC,i} - \theta_{EC}(1 - a_{i,t-1})$$
$$- \theta_{RN} \ln(1 + \sum_{j \neq i} a_{jt}) + \epsilon_{it1}$$

- θ_{RS}: Revenue parameter
- $\theta_{FC,i}$: Operating cost
- θ_{EC}: Entry cost
- θ_{RN}: Degree of strategic substitution

Dominant eigenvalue of Ψ_P and $M_{\Psi_\theta}\Psi_P$

θ_{RN}	$\rho(\Psi_P)$	$\rho(M_{\Psi_\theta}\Psi_P)$
1	0.337	0.292
2	0.693	0.595
4	1.184	1.180
6	1.479	1.478

Note: $\dim(P) = 144$ and $\dim(\theta) = 2$.

Strong strategic substitutability

$\rightarrow \{\tilde{\theta}_k, \tilde{P}_k\}$ diverges away from (θ^0, P^0).

What if Ψ is not locally contractive around P^0?

(1) Relaxation method:

$$[\Lambda(\theta, P)](a|x) = \{[\Psi(\theta, P)](a|x)\}^{\alpha} P(a|x)^{(1-\alpha)}, \alpha \in (0, 1)$$

— easy, works in some cases

(2) Recursive Projection Method (borrowed from Shroff and Keller, 1993, SIAM)

(1) Locally contractive mapping (relaxation method)

$$[\Lambda(\theta, P)](a|x) = \{[\Psi(\theta, P)](a|x)\}^{\alpha} P(a|x)^{(1-\alpha)}$$

- The same fixed point: $P_\theta = \Psi(\theta, P_\theta) = \Lambda(\theta, P_\theta)$
- Eigenvalues of $\nabla_{P'}\Lambda(\theta, P) = \alpha \times$(eigenvalues of Ψ_P) + $(1 - \alpha)$
- We can choose α so that $\rho(\Lambda_P) < 1$ if all the eigenvalues of Ψ_P are smaller than 1.

Convergence of Ψ and Λ for Dynamic Game Example

	$\theta_{RN} = 1$	$\theta_{RN} = 2$	$\theta_{RN} = 4$	$\theta_{RN} = 6$
$\rho(\Psi_P)$	0.337	0.693	1.184	1.479
$\rho(\Lambda_P)$	0.257	0.495	0.802	0.916
$\rho(M_{\Psi_\theta}\Psi_P)$	0.292	0.595	1.180	1.478
$\rho(M_{\Lambda_\theta}\Lambda_P)$	0.258	0.494	0.805	0.915

Alternative Sequential Estimation Procedure

- $\tilde{P}_0$: an initial consistent estimator
- Use $\Lambda(\theta, P)$ in place of $\Psi(\theta, P)$ and update θ and P sequentially.
- The algorithm converges locally if $\rho(\Lambda_P) < 1$.

(2) Recursive Projection Method (RPM) (Shroff and Keller, 1993, SIAM)

The eigenvalues of $\nabla_{P'}\Psi(\theta, P_\theta)$: $\{\lambda_1, ..., \lambda_L\}$

$$|\lambda_1| \geq ... \geq |\lambda_m| > 1 - \delta \geq |\lambda_{m+1}| \geq ... \geq |\lambda_L|.$$

Typically, m is small. Let

$$Z_\theta = \underbrace{[z_1, ..., z_m]}_{\text{eigenvectors}} \;, \quad \Pi_\theta \equiv \underbrace{Z_\theta (Z_\theta' Z_\theta)^{-1} Z_\theta'}_{\text{projection}}.$$

Decompose P_{j-1} as

$$u_{j-1} = \Pi_\theta P_{j-1}, \quad v_{j-1} = (I - \Pi_\theta) P_{j-1}.$$

RPM (cont.)

- Define
$$
\begin{aligned}
f(u,v,\theta) &\equiv \Pi_\theta \Psi(\theta, u+v), \\
g(u,v,\theta) &\equiv (I - \Pi_\theta)\Psi(\theta, u+v).
\end{aligned}
$$
- $P_j = \Psi(P_{j-1},\theta) = \underbrace{f(u_{j-1}, v_{j-1}, \theta)}_{\text{not contracting}} + g(u_{j-1}, v_{j-1}, \theta)$.
- Newton step on $u = f(u,v,\theta)$: $u_j = h(u_{j-1}, v_{j-1}, \theta)$
$$P_j = \Gamma(P_{j-1},\theta) = h(u_{j-1}, v_{j-1}, \theta) + g(u_{j-1}, v_{j-1}, \theta)$$

Sequential Estimation Procedure by approximate RPM

- $(\tilde{P}_0, \tilde{\theta}_0)$: an initial consistent estimator

- **Step 1:** Given $\tilde{P}_{k-1}$, estimate θ by

$$\tilde{\theta}_k = \arg\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ln[\Gamma(\theta;\ \tilde{P}_{k-1}, \tilde{\theta}_{k-1})](a_i|x_i).$$

- **Step 2:** Given $\tilde{\theta}_k$, update the estimate of P by

$$\tilde{P}_k = \Gamma(\tilde{\theta}_k;\ \tilde{P}_{k-1}, \tilde{\theta}_{k-1}).$$

- Iterate Steps 1-2: $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$

- The algorithm converges locally if $\rho(M_{\Gamma_\theta}\Gamma_P) < 1$, where $M_{\Gamma_\theta} = I - \Gamma_\theta(\Gamma_\theta' \Delta_P \Gamma_\theta)^{-1}\Gamma_\theta' \Delta_P$.

Stable Case: $\theta_{RN} = 2$ and $|\rho(M_{\Psi_\theta}\Psi_P)| < 1$

	Estimator	$n = 500$		$n = 2000$		$n = 8000$	
		Bias	$\sqrt{\text{MSE}}$	Bias	$\sqrt{\text{MSE}}$	Bias	$\sqrt{\text{MSE}}$
$\hat{\theta}_{RN}$	NPL-Ψ	-0.0467	0.4705	-0.0009	0.2339	-0.0095	0.1130
	NPL-Λ	-0.0467	0.4705	-0.0009	0.2339	-0.0095	0.1130
	RPM	-0.0544	0.4642	-0.0102	0.2274	-0.0111	0.1116
	Ψ ($k = 1$)	-0.7895	0.9604	-0.2565	0.3949	-0.0828	0.1687

Unstable Case: $\theta_{RN} = 4$ and $|\rho(M_{\Psi_\theta}\Psi_P)| > 1$

	Estimator	$n = 500$		$n = 2000$		$n = 8000$	
		Bias	$\sqrt{\text{MSE}}$	Bias	$\sqrt{\text{MSE}}$	Bias	$\sqrt{\text{MSE}}$
$\hat{\theta}_{RN}$	NPL-Ψ	-0.1417	0.2572	-0.1414	0.2314	-0.0918	0.1612
	NPL-Λ	0.0241	0.1424	-0.0001	0.0739	0.0013	0.0352
	RPM	0.0249	0.1604	-0.0003	0.0841	0.0014	0.0342
	Ψ $(k = 1)$	-0.7713	0.9094	-0.1964	0.2599	-0.0462	0.0937

Note: With 1000 simulated samples. We set $q = 4$ for q-NPL.

Unstable Case: MSE of $\hat{\theta}_{RN}$ for $\theta_{RN} = 4$ and $n = 8000$

$\sqrt{MSE}$ of $\hat{\theta}_{RN}$:

	k=2	k=5	k=10	k=15	k=20	k=25
NPL-Ψ	0.0676	0.0719	0.0784	0.1219	0.1302	0.1568
NPL-Λ	0.0697	0.0369	0.0355	0.0354	0.0354	0.0354
RPM	0.0599	0.0361	0.0353	0.0350	0.0349	0.0349

Note: With 1000 simulated samples.

Open questions

- Multiple equilibria (1): how to find all the fixed points on $\Psi(\theta, P)$
- Multiple equilibria (2): how to determine which equilibria are played by the data

References

Aguirregabiria, V. and Mira. P. (2010).“Dynamic discrete choice structural models: A survey.” Journal of Econometrics 156, 38-67.

Aguirregabiria, V. and Mira, P. (2007). “Sequential estimation of dynamic discrete games.” Econometrica 75, 1-53.

Imai, S., Jain N., and Ching, A. (2009). “Bayesian Estimation of Dynamic Discrete Choice Models." Econometrica 77, 1865-1899.

Rust, J. (1987). “Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher.” Econometrica 55, 999-1033.

Shroff, G. M. and Keller, H. B. (1993). “Stabilization of unstable procedures: the recursive projection method.” SIAM Journal of Numerical Analysis, 30, 1099-1120.

Su, Che-Lin and Judd, K. L. (2010). Constrained optimization approaches to estimation of structural models. Working Paper, University of Chicago.