Dynamic Oligopoly: Additional Issues

Paul Schrimpf

UBC
Economics 565

March 5, 2020

References

- Reviews:
  - Aguirregabiria and Nevo (2010)
  - Aguirregabiria (2017) chapters
  - Ackerberg, Caves, and Frazer (2015) section 3
  - Aguirregabiria and Mira (2010)

Section 1

Introduction

Introduction

- There are not many applied papers that estimate dynamic games (less true now than 4 years ago)
- Reasons:
  1. Estimating dynamic games is computationally intensive
  2. Assumption that the only unobserved heterogeneity are i.i.d shocks is not plausible
     - Why bother estimating a complicated model if the results are not credible?
     - Should add some permanent and/or autocorrelated unobserved heterogeneity
- Today we will look at recent research addressing these two issues

Section 2

Computation

Computation 1

- Estimation involves maximizing some objective function subject to equilibrium conditions
- Estimation methods:
  - Maximum likelihood

$$\max_{\theta \in \Theta, \mathbf{P} \in [0,1]^N} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{i=1}^{N} \log \Lambda \left(a_{imt} | v_i^{\mathbf{P}}(\cdot, x_{mt}; \theta) \right)$$
$$\text{s.t.} \mathbf{P} = \Lambda(v^{\mathbf{P}}(\theta))$$

  - 2-step estimators: estimate $\hat{\text{P}}(a|x)$ from observed actions and then

$$\max_{\theta \in \Theta} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{i=1}^{N} \log \Lambda(a_{imt} | v_i^{\hat{\mathbf{P}}}(\cdot, x_{mt}; \theta))$$

Computation 2

- Nested pseudo likelihood (NPL) (Aguirregabiria and Mira, 2007): after 2-step estimator update $\hat{\mathbf{P}}^{(k)} = \Lambda(v^{\hat{\mathbf{P}}^{(k-1)}}(\hat{\theta}^{(k-1)}))$, re-maximize pseudo likelihood to get $\hat{\theta}^{(k)}$ and repeat
- Computation time: 2-step $<$ NPL $\leq$ MLE
- Possible reductions in computation
  - Improve calculation of $\Lambda(v^{\mathbf{P}}(\theta))$ (main problem is $v^{\mathbf{P}}$)
  - Improve maximization
    - Better maximization algorithm (MPEC Su and Judd (2012)); same issues with starting values and local optima as in BLP models
    - Bayesian method (MCMC) instead of maximization, Imai, Jain, and Ching (2009) and Gallant, Hong, and Khwaja (2012)

Improving calculation of value function 1

- For finite state space can compute $v^{\mathbf{P}}$ as
$$V^{\mathbf{P}}(\theta) = (I - \delta \mathbf{M}_c)^{-1} \mathbf{M}_c \left[\boldsymbol{\pi}(\theta) + g(\mathbf{p}, \theta) \right]$$
- Inverting matrix takes $O(S^3)$ operations where S is size of state space
    - Matrix inversion becomes prohibitively slow for surprisingly moderate S
    - On my desktop (AMD-FX8150 cpu) $S = 1000$ takes 1.15 seconds, $S = 2000$ takes 11.1 seconds, $S = 3000$ takes 38.6 seconds , $S = 4000$ takes 91.2 seconds
    - Faster hardware can cut these times by a constant factor, but still face cubic growth
    - Using GPU instead of CPU for matrix inversion can be much faster

Improving calculation of value function 2

- On my desktop the GPU (NVidia GeForce GTX 560) takes about a hundredth as long to invert large matrices as the CPU
- Scientific computing using GPUs is a new and active field
- Programming for GPUs can be difficult
- Fast GPUs are not part of most servers

- Inverting sparse matrices can take much less than $O(S^3)$ operations
  - $I - \delta \mathbf{M}_c$ is often sparse
  - Exact complexity of inversion depends on number of non-zero entries and their locations (sparsity pattern)
- Some papers iterate value equation instead of explicitly inverting

$$V^{\mathbf{P}}(\theta) = \mathbf{M}_c \left[\pi(\theta) + g(\mathbf{p}, \theta) + \delta V^{\mathbf{P}}(\theta) \right]$$

  - Simulation often used

Improving calculation of value function 3

- Still solving same equation, if solving accurately has to take $O(S^3)$
- If iterating is faster must be either (i) implicitly exploiting sparsity or (ii) solving inaccurately
- Estimation can proceed with approximate solutions that only become exact at estimated θ, e.g. Kasahara and Shimotsu (2011)

- **M** depends on **P**, so for 2-step methods only need to compute inverse once
- State space can be very large even for models that appear simple
    - E.g. entry/exit game with N firms whose identities matter $S = 2^N|X|$

Reducing the size of the state space 1

- Economically motivated restrictions can reduce the size of the state space

R1 assume homogenous players and symmetric equilibrium
  - E.g. entry game, assume:
    1. Only number of competitors and not their identities affects profits
    2. Firms have the same profit function
    3. Symmetric equilibrium

  then state space size is $2(N+1)|X|$

R2 Inclusive values (Nevo and Rossi, 2008)
  - Inclusive value = in discrete choice model the expected utility of a consumer from facing several options before observing the shocks (McFadden et al., 1978)

$$E\left[\max_j u_j + \epsilon_j\right]$$

Reducing the size of the state space 2

- Adjusted inclusive value $\approx$ inclusive value minus firm's marginal costs; denote by i_f
- With appropriate assumptions, profits can be written as function of adjusted inclusive values,

$$\pi_f(\text{all state variables}) = \pi_f(i_f, i_{-f})$$

- Assume strategies only depend on adjusted inclusive values,

$$\mathrm{P}\left(i_{f,t+1}, i_{-f,t+1} | \text{state}_t\right) = \mathrm{P}\left(i_{f,t+1}, i_{-f,t+1} \middle| i_{f,t}, i_{-f,t}, \sum \text{invest}_{f,t}\right)$$

possible justifications:

  - Strong assumptions about investment process
  - Limited information of firms
  - Bounded rationality: firms have as hard a time computing strategies as we do

Reducing the size of the state space 3

- Then value function only depends on inclusive values

R3 Oblivious equilibrium (Weintraub, Benkard, and Van Roy (2008), Weintraub, Benkard, and Van Roy (2010), Farias, Saure, and Weintraub (2012))

- Oblivious equilibrium: firms make decisions conditional only on their own state variables and long-run industry average state
- In Markov equilibrium, firms make decisions based on all state variables
- Weintraub, Benkard, and Van Roy (2008) show that oblivious equilibrium approximates Markov perfect equilibrium as number of firms increases
- Krusell and Smith (1998) use similar idea in dynamic macro model

Finite dependence 1

- Value function:

$$\begin{aligned} V^{\mathbf{P}}(\theta) =& \mathbf{M}_c \left[\pi(\theta) + g(\mathbf{p}, \theta) + \delta V^{\mathbf{P}}(\theta)\right] \\ =& \mathbf{M}_c \left[\pi(\theta) + g(\mathbf{p}, \theta) + \delta \mathbf{M}_c \left[\pi(\theta) + g(\mathbf{p}, \theta) + \delta V^{\mathbf{P}}(\theta)\right]\right] \\ =& \left(\sum_{t=0}^{T} \mathbf{M}_c^{t+1} \delta^t\right) \left[\pi(\theta) + g(\mathbf{p}, \theta)\right] + \mathbf{M}_c^{T+2} \delta^{T+1} V^{\mathbf{P}}(\theta) \end{aligned}$$

- Arcidiacono and Miller (2011): if $\mathbf{M}_c^{T+2} \delta^{T+1} V^{\mathbf{P}}(\theta)$ is identical across actions then it will drop out of $v^p(a, x) - v^p(a', x)$, avoiding inversion
- Examples:
    - Renewal action: bus engine replacement
    - Terminal choice

Section 3

Unobserved heterogeneity

Unobserved heterogeneity 1

- Only unobservables in basic dynamic model are i.i.d. shocks
- More plausible to allow richer unobserved heterogeneity i.e. unobserved state variables
- Panel data can identify fairly rich unobserved heterogeneity
  - E.g. in linear models:
    - Random effects
    - Fixed effects
    - (Dynamic factors)
    - Fixed effects with additional autocorrelation (dynamic panel models as in Blundell and Bond (2000))
  - Variants of methods for linear models can be applied to dynamic games, but not straightforward because of
    - Nonlinearity – requires different identification arguments; complicates fixed effects estimation

Unobserved heterogeneity 2

- Computation – introducing unobserved state variables makes computing the model more complex
- Types of unobserved heterogeneity
  - Permanent firm or market unobserved heterogeneity
    - Similar to random or fixed effects
    - Aguirregabiria and Mira (2007), Collard-Wexler (2013), Aguirregabiria and Ho (2012)
  - Unobserved states that follow a controlled Markov process
    - Identification of transition probabilities: Kasahara and Shimotsu (2009), Hu and Shum (2012), Allman, Matias, and Rhodes (2009), Hu and Shum (2013)
    - Arcidiacono and Miller (2011), Kasahara and Shimotsu (2011)

Permanent unobserved heterogeneity 1

- Here we go over approach of Aguirregabiria and Mira (2007) as described in Aguirregabiria and Nevo (2010)
- Market specific random effect in profits

$$\pi_{imt} = \Pi(a_{mt}, x_{imt}, \theta) + \theta_i(a) + \sigma_i \xi_m + \epsilon_{imt}$$

  - ϵ_{imt} i.i.d.
  - ξ_m unobserved, discrete with finite support, known mean and variance (absorbed by θ_i and σ_i), known support, $\{\xi^\ell\}_{\ell=1}^L$, pmf λ
  - θ_i and σ_i varying with i requires large T (or large M and same firms across markets)
- Conditional choice probabilities different for each ℓ, denote by $\mathbf{P}_\ell$

Permanent unobserved heterogeneity 2

- Equilibrium for each ℓ:

$$\mathbf{P}^\ell = \Lambda(v^{\mathbf{P}_\ell}(\theta), \ell)$$

- Pseudo-likelihood integrates over distribution of ξ_m

$$\max_{\theta, \mathbf{P}^\ell, \lambda} \sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{i=1}^{N} \log \left(\sum_{\ell=1}^{L} \lambda_{\ell|x} \Lambda \left(a_{imt} | v_i^{\mathbf{P}_\ell}(\cdot, x_{mt}; \theta, \ell) \right) \right)$$

$$\text{s.t.} \mathbf{P}_\ell = \Lambda(v^{\mathbf{P}_\ell}(\theta), \ell)$$

where

$$\lambda_{\ell|x} = \mathrm{P}(\xi_m = \xi^\ell | x_{m1})$$

- Initial conditions problem

Permanent unobserved heterogeneity 3

- ξ_m will be correlated with initial values of endogenous state variables (markets with high ξ_m will start with a large number of firms)
- One solution: assume stationary, find stationary distribution of $x|\xi$,

$$P(x_t|\xi_m = \xi^\ell) = \sum_{x_{t-1}} P(x_{t-1}|\xi_m = \xi^\ell) P(x_t|x_{t-1}, \xi_m = \xi^\ell)$$

Bayes' rule

$$\lambda_{\ell|x} = \frac{\lambda_\ell P(x|\xi_m = \xi^\ell)}{\sum_{j=1}^{L} \lambda_j P(x|\xi_m = \xi^j)}$$

- To apply 2-step estimators need to first consistently estimate transition probabilities conditional on unobserved ξ_m and $P^\ell(\cdot|x_{mt})$

Permanent unobserved heterogeneity 4

- Can use Kasahara and Shimotsu (2009), but Aguirregabiria and Nevo (2010) say estimation is difficult
- Identification argument is constructive, based on singular value decomposition, can mimic for estimation , e.g. Hu, Shum, and Tan (2010)
- Levine, Hunter, and Chauveau (2011)

- NPL can be used but must iterate to convergence unless started from consistent $\hat{P}^{\ell}(\cdot|x_{mt})$
  - Start with arbitrary $P^{\ell}(\cdot|x_{mt})$
  - Maximize pseudo likelihood to get $\hat{\theta}$, distribution of ξ_m
  - Update $P^{\ell}(\cdot|x_{mt})$
  - Repeat until convergence

Unobserved autocorrelated state variables 1

- Suppose state $x_{mt} = (x^o_{mt}, x^u_{mt})$ where only x^o_{mt} is observed
- Kasahara and Shimotsu (2009) (for finite) and Hu and Shum (2012) (for continuous) give conditions for identification of transition probabilities $P(\cdot|x^o_{mt}, x^u_{mt})$
- Given consistent $\hat{P}(\cdot|x^o_{mt}, x^u_{mt})$ can apply 2-step estimator or NPL
    - Estimation is difficult

Arcidiacono and Miller (2011)

- Computationally tractable estimation with unobserved state variables
- Two innovations:
  - Avoid matrix inversion in value function computation through finite dependence
  - Modified EM algorithm to integrate out distribution of unobserved states

EM algorithm 1

- "Expectation-Maximization"
  - Expectation: conditional probabilities of unobserved state given observables and parameters updated
  - Maximization: maximize likelihood as though unobserved state observed
  - Repeated until convergence
- Setup: observe x, missing s, complete $z = (x, s)$
- Joint likelihood $p(x, z|\theta)$
- Marginal likelihood $L(\theta; x) = p(x|\theta) = \mathrm{E}[p(x, z|\theta)|x, \theta]$
- Difficulty:

$$\mathrm{E}[p(x, s|\theta)|x, \theta] = \int_S p(x, s|\theta)p(s|x; \theta)ds$$

  might be hard to compute
- Steps:

EM algorithm 2

- Initial θ^0
- Expectation: calculate $p(s|x;\theta^0)$
- Maximization: $\theta^1 = \arg\max_\theta \int_S p(x,s|\theta)p(s|x;\theta^0)ds$
- Iterate to convergence

- Pros: stable – each iteration guaranteed to increase likelihood
- Cons: slow?
- Arcidiacono and Miller (2011):
  - Not slow when using finite dependence so that maximization step fast
  - Show 2-step version of EM algorithm possible, i.e. can estimate $p(s|x;\theta)$ without estimating θ

$$p(s|x;\theta) = \frac{p(x,s|\theta)}{\sum_{s'} p(x,s'|\theta)}$$

Can use empirical probabilities in place of model probabilities within EM algorithm

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6):2411–2451. URL http://dx.doi.org/10.3982/ECTA13408.

Aguirregabiria, Victor. 2017. "Empirical Industrial Organization: Models, Methods, and Applications." URL http://www.individual.utoronto.ca/vaguirre/courses/eco2901/teaching_io_toronto.html.

Aguirregabiria, Victor and Chun-Yu Ho. 2012. "A dynamic oligopoly game of the US airline industry: Estimation and policy experiments." *Journal of Econometrics* 168 (1):156 – 173. URL http://www.sciencedirect.com/science/article/pii/S030440761100176X. <ce:title>The Econometrics of Auctions and Games</ce:title>.

Aguirregabiria, Victor and Pedro Mira. 2007. "Sequential Estimation of Dynamic Discrete Games." *Econometrica* 75 (1):pp. 1–53. URL http://www.jstor.org/stable/4123107.

———. 2010. “Dynamic discrete choice structural models: A survey.” *Journal of Econometrics* 156 (1):38 – 67. URL http://www.sciencedirect.com/science/article/pii/S0304407609001985.

Aguirregabiria, Victor and Aviv Nevo. 2010. “Recent developments in empirical IO: dynamic demand and dynamic games.” MPRA Paper 27814, University Library of Munich, Germany. URL http://ideas.repec.org/p/pra/mprapa/27814.html.

Allman, Elizabeth S., Catherine Matias, and John A. Rhodes. 2009. “IDENTIFIABILITY OF PARAMETERS IN LATENT STRUCTURE MODELS WITH MANY OBSERVED VARIABLES.” *The Annals of Statistics* 37 (6A):pp. 3099–3132. URL http://www.jstor.org/stable/25662188.

Arcidiacono, Peter and Robert A. Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity." *Econometrica* 79 (6):1823–1867. URL http://dx.doi.org/10.3982/ECTA7743.

Blundell, R. and S. Bond. 2000. "GMM estimation with persistent panel data: an application to production functions." *Econometric Reviews* 19 (3):321–340. URL http://www.tandfonline.com/doi/pdf/10.1080/07474930008800475.

Collard-Wexler, Allan. 2013. "Demand Fluctuations in the Ready-Mix Concrete Industry." *Econometrica* 81 (3):1003–1037. URL http://dx.doi.org/10.3982/ECTA6877.

Farias, Vivek, Denis Saure, and Gabriel Y. Weintraub. 2012. "An approximate dynamic programming approach to solving dynamic oligopoly models." *The RAND Journal of Economics* 43 (2):253–282. URL http://dx.doi.org/10.1111/j.1756-2171.2012.00165.x.

Gallant, A Ronald, Han Hong, and Ahmed Khwaja. 2012. "Bayesian Estimation of a Dynamic Game with Endogenous, Partially Observed, Serially Correlated State." Tech. rep., Duke University, Department of Economics. URL http://econpapers.repec.org/paper/dukdukeec/12-01.htm.

Hu, Yingyao and Matthew Shum. 2012. "Nonparametric identification of dynamic models with unobserved state variables." *Journal of Econometrics* 171 (1):32 – 44. URL http://www.sciencedirect.com/science/article/pii/S0304407612001479.

———. 2013. “Identifying dynamic games with serially-correlated unobservables.” *Structural Econometric Models* 31:97À113. URL http://www.econ2.jhu.edu/people/hu/AIE-Hu-Shum-2013.pdf.

Hu, Yingyao, Matthew Shum, and Wei Tan. 2010. “A simple estimator for dynamic models with serially correlated unobservables.” Working papers // the Johns Hopkins University, Department of Economics 558, Baltimore, Md. URL http://hdl.handle.net/10419/49872.

Imai, Susumu, Neelam Jain, and Andrew Ching. 2009. “Bayesian estimation of dynamic discrete choice models.” *Econometrica* 77 (6):1865–1899. URL http://onlinelibrary.wiley.com/doi/10.3982/ECTA5658/abstract.

Kasahara, Hiroyuki and Katsumi Shimotsu. 2009. “Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices.” *Econometrica* 77 (1):135–175. URL http://dx.doi.org/10.3982/ECTA6763.

———. 2011. “Sequential Estimation of Dynamic Programming Models with Unobserved Heterogeneity.” Tech. rep.

Krusell, Per and Anthony A. Smith. 1998. “Income and Wealth Heterogeneity in the Macroeconomy.” *Journal of Political Economy* 106 (5):pp. 867–896. URL http://www.jstor.org/stable/10.1086/250034.

Levine, M., D. R. Hunter, and D. Chauveau. 2011. “Maximum smoothed likelihood for multivariate mixtures.” *Biometrika* URL http://biomet.oxfordjournals.org/content/early/2011/04/28/biomet.asq079.abstract.

McFadden, Daniel et al. 1978. *Modelling the choice of residential location*. Institute of Transportation Studies, University of California.

Nevo, Aviv and Federico Rossi. 2008. “An approach for extending dynamic models to settings with multi-product firms.” *Economics Letters* 100 (1):49–52.

Su, C.L. and K.L. Judd. 2012. “Constrained optimization approaches to estimation of structural models.” *Econometrica* 80 (5):2213–2230. URL http://onlinelibrary.wiley.com/doi/10.3982/ECTA7925/abstract.

Weintraub, Gabriel Y., C. Lanier Benkard, and Benjamin Van Roy. 2008. “Markov Perfect Industry Dynamics With Many Firms.” *Econometrica* 76 (6):1375–1411. URL http://dx.doi.org/10.3982/ECTA6158.

Weintraub, Gabriel Y, C Lanier Benkard, and Benjamin Van Roy. 2010. “Computational methods for oblivious equilibrium.” *Operations research* 58 (4-Part-2):1247–1265. URL http://or.journal.informs.org/content/58/4-Part-2/1247.short.