Extension to Dynamic Games

Robert A. Miller

Tilburg University

November 2023

A Class of Dynamic Markov Games

Players and choices

- Consider a dynamic infinite horizon game for finite I players.
- Thus $T = \infty$ and $I < \infty$.
- Each player $i \in I$ makes a choice $d_t^{(i)} \equiv \left(d_{t1}^{(i)}, \ldots, d_{tJ}^{(i)}\right)$ in period t.
- Denote the choices of all the players in period t by:

$$d_t \equiv \left(d_t^{(1)}, \ldots, d_t^{(I)}\right)$$

and denote by:

$$d_t^{(-i)} \equiv \left(d_t^{(1)}, \ldots, d_t^{(i-1)}, d_t^{(i+1)}, \ldots, d_t^{(I)}\right)$$

the choices of $\{1, \ldots, i-1, i+1, \ldots, I\}$ in period t, that is all the players apart from i.

A Class of Dynamic Markov Games

State variables

- Denote by x_t the state variables of the game that are not *iid*.
- For example x_t includes the capital of every firm. Then:
  - firms would have the same state variables.
  - x_t would affect rivals in very different ways.
- We assume all the players observe x_t.
- Denote by $F\left(x_{t+1}\left|x_t, d_t\right.\right)$ the probability of x_{t+1} occurs when the state variables are x_t and the players collectively choose d_t.
- Similarly let:

$$F_j\left(x_{t+1}\left|x_t, d_t^{(-i)}\right.\right) \equiv F\left(x_{t+1}\left|x_t, d_t^{(-i)}, d_{jt}^{(i)} = 1\right.\right)$$

  denote the probability distribution determining x_{t+1} given x_t when $\{1, \ldots, i-1, i+1, \ldots, I\}$ choose $d_t^{(-i)}$ in t and i makes choice j.

A Class of Dynamic Markov Games

Payoffs and information

- Suppose $\epsilon_t^{(i)} \equiv \left(\epsilon_{1t}^{(i)}, \ldots, \epsilon_{Jt}^{(i)}\right)$ is *iid* with density $g\left(\epsilon_t^{(i)}\right)$ that affects the payoffs of i in t.
- Also let $\epsilon_t^{(-i)} \equiv \left(\epsilon_t^{(1)}, \ldots, \epsilon_t^{(i-1)}, \epsilon_t^{(i+1)}, \ldots, \epsilon_t^{(I)}\right)$.
- The systematic component of current utility or payoff to player i in period t form taking choice j when everybody else chooses $d_t^{(-i)}$ and the state variables are z_t is denoted by $U_j^{(i)}\left(x_t, d_t^{(-i)}\right)$.
- Denoting by $\beta \in (0,1)$ the discount factor, the summed discounted payoff to player i throughout the course of the game is:

$$\sum_{t=1}^{T} \sum_{j=1}^{J} \beta^{t-1} d_{jt}^{(i)} \left[U_j^{(i)}\left(x_t, d_t^{(-i)}\right) + \epsilon_{jt}^{(i)}\right]$$

- Players noncooperatively maximize their expected utilities, moving simultaneously each period. Thus i does not condition on $d_t^{(-i)}$ when making his choice at date t, but only sees $\left(x_t, \epsilon_t^{(i)}\right)$.

Markov Perfect Equilibrium

Markov strategies

- This is a stationary environment and we focus on Markov decision rules, which can be expressed $d_j^{(i)}\left(x_t, \epsilon_t^{(i)}\right)$.
- Let $d^{(-i)}\left(x_t, \epsilon_t^{(-i)}\right)$ denote the strategy of every player but i:

$$\left(\begin{array}{r} d^{(1)}\left(x_t, \epsilon_t^{(1)}\right), \ldots, d^{(i-1)}\left(x_t, \epsilon_t^{(i-1)}\right), d^{(i+1)}\left(x_t, \epsilon_t^{(i+1)}\right), \\ d^{(i+2)}\left(x_t, \epsilon_t^{(i+2)}\right) \ldots, d^{(I)}\left(x_t, \epsilon_t^{(I)}\right) \end{array}\right)$$

- Then the expected value of the game to i from playing $d_j^{(i)}\left(x_t, \epsilon_t^{(i)}\right)$ when everyone else plays $d\left(x_t, \epsilon_t^{(-i)}\right)$ is:

$$V^{(i)}\left(x_1\right) \equiv$$
$$E\left\{\sum_{t=1}^{\infty} \sum_{j=1}^{J} \beta^{t-1} d_j^{(i)}\left(x_t, \epsilon_t^{(i)}\right)\left[U_j^{(i)}\left(z_t, d\left(x_t, \epsilon_t^{(-i)}\right)\right)+\epsilon_{jt}^{(i)}\right] | x_1\right\}$$

Markov Perfect Equilibrium

Choice probabilities generated by Markov strategies

- Integrating over $\epsilon_t^{(i)}$ we obtain the j^{th} conditional choice probability for the i^{th} player at t as $p_j^{(i)}(x_t)$:

$$p_j^{(i)}(x_t) = \int d_j^{(i)}\left(x_t, \epsilon_t^{(i)}\right) g\left(\epsilon_t^{(i)}\right) d\epsilon_t^{(i)}$$

- Let $P\left(d_t^{(-i)} \left| x_t\right.\right)$ denote the joint probability firm i's competitors choose $d_t^{(-i)}$ conditional on the state variables z_t.
- Since $\epsilon_t^{(i)}$ is distributed independently across $i \in \{1, \ldots, I\}$:

$$P\left(d_t^{(-i)} \left| x_t\right.\right) = \prod_{\substack{i'=1 \\ i' \neq i}}^{I} \left(\sum_{j=1}^{J} d_{jt}^{(i')} p_j^{(i')}(x_t)\right)$$

- The strategy $\left\{ d^{(i)}\left(x_t, \epsilon_t^{(i)}\right)\right\}_{i=1}^{I}$ is a Markov perfect equilibrium (MPE) if, for all $\left(i, x_t, \epsilon_t^{(i)}\right)$, the best response of i to $d^{(-i)}\left(x_t, \epsilon_t^{(-i)}\right)$ is $d^{(i)}\left(x_t, \epsilon_t^{(i)}\right)$ when everybody uses the same strategy thereafter.
- That is, suppose the other players collectively use $d^{(-i)}\left(x_t, \epsilon_t^{(-i)}\right)$ in period t, and $V^{(i)}(x_{t+1})$ is formed from $\left\{ d^{(i)}\left(x_t, \epsilon_t^{(i)}\right)\right\}_{i=1}^{I}$.
- Then $d^{(i)}\left(x_t, \epsilon_t^{(i)}\right)$ solves for i choosing j to maximize:

$$\sum_{d_t^{(-i)}} P\left(d_t^{(-i)} \,|x_t\right) \left\{ \begin{array}{l} U_j^{(i)}\left(x_t, d_t^{(-i)}\right) \\ +\beta \sum_{z=1}^{X} V^{(i)}(x) F_j\left(x \left| x_t, d_t^{(-i)}\right.\right) \end{array} \right\} + \epsilon_{jt}^{(i)}$$

Adapting Dynamic Games to the CCP Framework

Connection to individual optimization

- In equilibrium, the systematic component of the current utility of player i in period t, as a function of x_t, the state variables for game, and his own decision j, is:

$$u_j^{(i)}\left(x_t\right) = \sum_{d_t^{(-i)}} P\left(d_t^{(-i)} \left| x_t\right.\right) U_j^{(i)}\left(x_t, d_t^{(-i)}\right)$$

- Similarly the probability transition from x_t to x_{t+1} given action j by firm i is given by:

$$f_j^{(i)}\left(x_{t+1} \left| x_t^{(i)}\right.\right) = \sum_{d_t^{(-i)}} P\left(d_t^{(-i)} \left| x_t^{(i)}\right.\right) F_j\left(x_{t+1} \left| x_t, d_t^{(-i)}\right.\right) \tag{1}$$

- The setup for player i is now identical to the optimization problem described in the second lecture for a stationary environment.

Adapting Dynamic Games to the CCP Framework

Inversion and representation theorems

- The inversion and representation theorems of the previous lecture apply to this multiagent setting with two critical differences.
- The first difference is a straightforward extension but the second complicates identification and predicting counterfactuals:
  1. $f_{jt}\left(x_{t+1}\left|x_t\right.\right)$ is a primitive in single agent optimization problems, but $f_{jt}^{(i)}\left(x_{t+1}\left|x_t\right.\right)$ depends on CCPs of the other players, $P_t\left(d_t^{(\sim i)}\left|x_t\right.\right)$, as well as the primitive $F_{jt}\left(x_{t+1}\left|x_t, d_t^{(\sim i)}\right.\right)$. However both $P_t\left(d_t^{(\sim i)}\left|x_t\right.\right)$ and $F_j\left(x_{t+1}\left|x_t, d_t^{(-i)}\right.\right)$ are identified so it is easy to place restrictions on $f_{jt}\left(x_{t+1}\left|x_t\right.\right)$ using (1).
  2. $u_{jt}(x_t)$ is a primitive in single agent optimization problems, but $u_{jt}^{(i)}(x_t)$ is a reduced form parameter found by integrating $U_{jt}^{(i)}\left(x_t, d_t^{(\sim i)}\right)$ over the joint probability distribution $P_t\left(d_t^{(\sim i)}\left|x_t\right.\right)$.

- Note that:
  1. there might be multiple equilibria, but we assume:
     - either every firm plays in the same market
     - or every market plays the same equilibrium.
  2. in contrast to ML we do not solve for the equilibrium.
  3. estimation is based on conditions that are satisfied by every MPE.
  4. the estimation approach is identical to the approach we described in the individual optimization problem.
- The basic difference between estimating this dynamic game and an individual optimization problem using a CCP estimator revolves around how much the payoffs of each player are affected by state variables partially determined by other players through their conditional choice probabilities.

- Suppose there is a finite maximum number of firms in a market at any one time denoted by I.
- If a firm exits, the next period an opening occurs to a potential entrant, who may decide to exercise this one time option, or stay out.
- At the beginning of each period every incumbent firm has the option of quitting the market or staying one more period.
- Let $d_t^{(i)} \equiv \left(d_{t1}^{(i)}, d_{t2}^{(i)}\right)$, where $d_{t1}^{(i)} = 1$ means i exits or stays out of the market in period t, and $d_{t2}^{(i)} = 1$ means i enters or does not exit.
- If $d_{t2}^{(i)} = 1$ and $d_{t-1,1}^{(i)} = 1$ then the firm in spot i at time t is an entrant, and if $d_{t-1,2}^{(i)} = 1$ the spot i at time t is an incumbent.

Entry Exit Game

State variables

- In this application there are three components to the state variables and $x_t = (x_1, x_{2t}, s_t)$.
- The first is a permanent market characteristic, denoted by x_1, and is common across firms in the market. Each market faces an equal probability of drawing any of the possible values of x_1 where $x_1 \in \{1, 2, \ldots, 10\}$.
- The second, x_{2t}, is whether or not each firm is an incumbent, $x_{2t} \equiv \{d_{t-1,2}^{(1)}, \ldots, d_{t-1,2}^{(I)}\}$. Entrants pay a start up cost, making it more likely that stayers choose to fill a slot than an entrant.
- A demand shock $s_t \in \{1, \ldots, 5\}$ follows a first order Markov chain.
- In particular, the probability that $s_{t+1} = s_t$ is fixed at $\pi \in (0, 1)$, and probability of any other state occurring is equally likely:

$$\Pr\{s_{t+1} | s_t\} = \begin{cases} \pi \text{ if } s_{t+1} = s_t \\ (1 - \pi)/4 \text{ if } s_{t+1} \neq s_t \end{cases}$$

- Each active firm produces one unit so revenue, denoted by y_t, is just price.
- Price is determined by:
  1. the supply of active firms in the market, $\sum_{i=1}^{I} d_{t2}^{(i)}$
  2. a permanent market characteristic, x_1
  3. the Markov demand shock s_t
  4. another temporary shock, denoted by η_t, distributed *iId* standard normal distribution, revealed to each market after the entry and exit decisions are made.
- The price equation is:

$$y_t = \alpha_0 + \alpha_1 x_1 + \alpha_2 s_t + \alpha_3 \sum_{i=1}^{I} d_{t2}^{(i)} + \eta_t$$

Entry Exit Game

Expected profits conditional on competition

- We assume costs comprise a choice specific disturbance $\epsilon_{tj}^{(i)}$ that is privately observed, plus a linear function of $\left(x_t^{(i)}, s_t^{(i)}, d_t^{(-i)}\right)$.
- Net current profits for exiting incumbent firms, and potential entrants who do not enter, are $\epsilon_{1t}^{(i)}$. Thus $U_1^{(i)}\left(x_t^{(i)}, s_t^{(i)}, d_t^{(-i)}\right) \equiv 0$.
- Current profits from being active are the sum of $\left(\epsilon_{2t}^{(i)} + \eta_t\right)$ and:

$$U_2^{(i)}\left(x_t^{(i)}, s_t^{(i)}, d_t^{(-i)}\right) \equiv \theta_0 + \theta_1 x_1 + \theta_2 s_t + \theta_3 \sum_{\substack{i'=1 \\ i' \neq i}}^{I} d_{2t}^{(i')} + \theta_4 d_{1,t-1}^{(i)}$$

  where θ_4 is the startup cost that only entrants pay.
- In equilibrium $E\left(\eta_t\right) = 0$ so:

$$u_j^{(i)}\left(x_t, s_t\right) = \theta_0 + \theta_1 x_1 + \theta_2 s_t + \theta_3 \sum_{\substack{i'=1 \\ i' \neq i}}^{I} p_2^{(i')}\left(x_t, s_t\right) + \theta_4 d_{1,t-1}^{(i)}$$

Entry Exit Game

Terminal choice property

- We assume the firm's private information, $\epsilon_{jt}^{(i)}$, is distributed T1EV.
- Later we show that since exiting is a terminal choice, with a payoff normalized to zero, given T1EV, the conditional value function for being active is:

$$\begin{aligned} v_2^{(i)}(x_t, s_t) &= u_2^{(i)}(x_t, s_t) \\ &\quad -\beta \sum_{x \in X} \sum_{s \in S} \left(\ln \left[p_1^{(i)}(x, s) \right] \right) f_2^{(i)}(x, s | x_t, s_t) \end{aligned}$$

- The future value term is then expressed as a function solely of the one-period-ahead probabilities of exiting and the transition probabilities of the state variables.

Entry Exit Game

Monte Carlo

- The number of firms in each market is set to six and we simulated data for 3,000 markets.
- The discount factor is set to $\beta = 0.9$.
- Starting at an initial date with six potential entrants in the market, we solved the model, ran the simulations forward for twenty periods, and used the last ten periods to estimate the model.
- The key difference between this Monte Carlo and the renewal Monte Carlo is that the conditional choice probabilities have an additional effect on both current utility and the transitions on the state variables due to the effect of the choices of the firm's competitors on profits.

Entry Exit Game

Results from Monte Carlo simulations (Arcidiacono and Miller, 2011)

	DGP (1)	s_t Observed (2)	Ignore s_t (3)	CCP Model (4)	CCP Data (5)	Two-Stage (6)	No Prices (7)
Profit parameters							
θ_0 (intercept)	0	0.0207	−0.8627	0.0073	0.0126	−0.0251	−0.0086
		(0.0779)	(0.0511)	(0.0812)	(0.0997)	(0.1013)	(0.1083)
θ_1 (obs. state)	0.05	−0.0505	−0.0118	−0.0500	−0.0502	−0.0487	−0.0495
		(0.0028)	(0.0014)	(0.0029)	(0.0041)	(0.0039)	(0.0038)
θ_2 (unobs. state)	0.25	0.2529		0.2502	0.2503	0.2456	0.2477
		(0.0080)		(0.0123)	(0.0148)	(0.0148)	(0.0158)
θ_3 (no. of competitors)	−0.2	−0.2061	0.1081	−0.2019	−0.2029	−0.1926	−0.1971
		(0.0207)	(0.0115)	(0.0218)	(0.0278)	(0.0270)	(0.0294)
θ_4 (entry cost)	−1.5	−1.4992	−1.5715	−1.5014	−1.4992	−1.4995	−1.5007
		(0.0131)	(0.0133)	(0.0116)	(0.0133)	(0.0133)	(0.0139)
Price parameters							
α_0 (intercept)	7	6.9973	6.6571	6.9991	6.9952	6.9946	
		(0.0296)	(0.0281)	(0.0369)	(0.0333)	(0.0335)	
α_1 (obs. state)	−0.1	−0.0998	−0.0754	−0.0995	−0.0996	−0.0996	
		(0.0023)	(0.0025)	(0.0028)	(0.0028)	(0.0028)	
α_2 (unobs. state)	0.3	0.2996		0.2982	0.2993	0.2987	
		(0.0045)		(0.0119)	(0.0117)	(0.0116)	
α_3 (no. of competitors)	−0.4	−0.3995	−0.2211	−0.3994	−0.3989	−0.3984	
		(0.0061)	(0.0051)	(0.0087)	(0.0088)	(0.0089)	
π (persistence of unobs. state)	0.7			0.7002	0.7030	0.7032	0.7007
				(0.0122)	(0.0146)	(0.0146)	(0.0184)
Time (minutes)		0.1354	0.1078	21.54	27.30	15.37	16.92
		(0.0047)	(0.0010)	(1.5278)	(1.9160)	(0.8003)	(1.6467)

Mean and standard deviations for 100 simulations. Observed data consist of 3000 markets for 10 periods with 6 firms in each market. In column 7, the CCP's are updated with the model.

Environmental Regulation Costs (Ryan, 2012)

Overview

- Background
  - Environmental Protection Agency (EPA) regulates the emissions of airborne pollutants such as ozone, sulfur dioxide, and nitrogen oxides.
  - 1962 Rachel Carson publishes *Silent Spring*.
  - 1970 Clean Air Act (CAA) passed and EPA established.
  - 1990 Amendments to CAA significantly strengthens it.
- Study of effects of amendments on the U.S. cement industry:
  - Data on about 2,000 observations covers period 1980 through 1998
  - 517 observations on 27 regional market segments.
  - Each market contains between 1 and 20 plants, 5 on average.
- The main findings of the study are that:
  - Entry costs increased but incumbents now face less competition.
  - Overall welfare decreased by at least $810M.
  - Focusing on certification process alone dramatically underpredicts welfare costs of new regulations

- Each period in this dynamic game firms choose:
  - *quantity produced* for the cement market.
  - *capacity adjustments* to firm size
  - *entry* by new firms or *exit* by incumbents.
- Solving for behavior in the model:
  - Entry and exit are determined in an MPE for the dynamic game.
  - Capacity adjustments follow an *exogenous (S,s) rule*.
  - Current quantity produced is embedded in a *static Cournot game* between incumbents as part of the MPNE.

- Estimation proceeds sequentially:
  - demand curve for each market
  - production costs estimated for each firm from the first order conditions of an interiro solution to a static Cournot equilibrium model
  - We review the estimation of the static game below.
  - a CCP estimator gives the adjustment and entry/exit costs.
- Policy experiment computes MPNE:
  - *before* and *after* the Amendment was introduced.
  - on a *representative* market (Alabama).
  - for two initial conditions, either no firms or two firms in the market.
  - where the regulation is more costly if there are fewer firms in the industry.
  - Extrapolating to the whole country these estimates range between $US million 810 to $US billion 1.3

- Demand in market j at time t is estimated with the equation:

$$\ln Q_{jt} = \alpha_0 + \alpha_1 \ln P_{jt} + \alpha_{2j} + \alpha_{3t}' X_{jt} + \varepsilon_{jt}$$

  where:

  - Q_{jt} is quantity demanded in market j at time t
  - α_1 is the elasticity of demand
  - α_{2j} is a market demand shifter *(including housing permits, time trends, population)*
  - X_{jt} is a vector of covaiates that affect demand
  - ε_{jt} is unobserved

- The instruments are supply side shifters *(including coal prices, gas prices, electricity rates, wage rates)*,

- The cost of production for firm i is given by:

$$C_i(q_i, s_i) = \delta_0 + \delta_1 q_i + \delta_2 \mathbf{1}\{q_i - vs_i\}(q_i - vs_i)^2$$

where:

  - s_i is current *(soft)* capacity of firm i
  - q_i is quantity produced by firm i
  - $C_i(q_i, s_i)$ denotes total costs of firm i producing q_i with capacity s_i
  - δ_0 are fixed costs *(of keeping the firm open)*
  - δ_1 is marginal cost of production up to capacity s_i
  - $\mathbf{1}\{q_i - vs_i\}(q_i - vs_i)^2$ captures effects of overtime and overutilization

A Static Cournot Model of Supply and Demand

Estimating production costs

- Abbreviate the specifications above, and denote by:
  - $Q \equiv \alpha_0 P^{\alpha_1}$ demand for concrete Q as a function of price P.
  - $P(Q, \alpha) \equiv (Q/\alpha_0)^{1/\alpha_1}$ the inverse demand curve.
  - $q_i^o \equiv q_i^o(\alpha, \delta)$ the equilibrium quantity produced by firm i.
  - $Q^{(-i)} = \sum_{k=1, k\neq i}^{I} q_k^o$ equilibrium production by the other firms.
- Given any fixed positive number $Q^{(-i)}$, firm i chooses q_i to maximize:

$$q_i \left[\alpha_0^{-1} \left(Q^{(-i)} + q_i \right) \right]^{1/\alpha_1} - C_i(q_i, \delta)$$

- In an interior solution, the policy function for firm i satisfies:

$$q_i^o P' \left(\sum_{k=1}^{I} q_k^o, \alpha \right) + P \left(\sum_{k=1}^{I} q_k^o, \alpha \right) = C_i'(q_i^o, \delta)$$

- Estimate δ by solving *(for all the markets and time periods)*:

$$\widehat{\delta} = \arg\min_{\delta} \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{i=1}^{I_{jt}} \left[q_{it} - q_{ijt}^o(, \widehat{\alpha}, \delta) \right]^2$$