ECO 310: Empirical Industrial Organization

Lecture 12: Models of Market Entry: Spatial Location

Victor Aguirregabiria (University of Toronto)

December 6th, 2018

Models of Market Entry [Cont.]: Outline

1. **Bresnahan & Reiss [Empirical Results]**
2. **Models of Firms' Spatial Location (Seim, 2006)**

1. Bresnahan & Reiss (JPE, 1991)

Market entry with homogeneous firms

- We start with an empirical model of entry in an homogeneous product industry and where all the firms have the same costs.
- There are several reasons why we start with this case.
- 1. This is the simpler empirical model of entry, and where this literature started with the seminal work by Bresnahan & Reiss (JPE, 1990).
- 2. The model with heterogeneous firms typically has multiple equilibria, and this makes the estimation more complicated.
- 3. Sometimes we have very limited information about firms' heterogeneous characteristics.

Market entry with homogeneous firms: Data

- Suppose the researcher has data from M markets in the same industry.
- For instance, the supermarket industry. The M markets are M neighborhoods from different Canadian cities.
- Markets are indexed by m.
- The dataset consists of:

$$\text{Data} = \{\ n_m,\ S_m, X_m : m = 1, 2, ..., M\}$$

n_m = number of active firms;
S_m = market size;
X_m = other exogenous market characteristics affecting demand or costs.

Market entry with homogeneous firms: Model

- All the potential entrants in a market have the same profit function:
  - Same costs, and same demand (homogenous product).
- The profit function of a firm in market m is:

$$V_m(n) - F_m$$

  where $V_m(n)$ is the variables profit, F_m is the fixed cost, and n is the number of active firms in the market.

- We describe below the specification of $V_m(n)$ and F_m in terms of observable variables and unobservables.
- A key feature is that $V_m(n)$ is a strictly decreasing function of n.

Market entry with homogeneous firms: Model [2]

- Under Nash-equilibrium, we have the following conditions:

$$V_m\left(1+\sum_{j\neq i} a_{jm}\right) - F_m \geq 0 \qquad \text{for firms with } a_{im} = 1$$

$$V_m\left(1+\sum_{j\neq i} a_{jm}\right) - F_m < 0 \qquad \text{for firms with } a_{im} = 0$$

- Then, n_m is an equilibrium iff:

$$V_m\left(n_m\right) - F_m \geq 0 \qquad \text{Active firms are in their best response}$$

$$V_m\left(1+n_m\right) - F_m < 0 \qquad \text{Inactive firms are in their best response}$$

Market entry with homogeneous firms: Model [3]

- We can write the Nash-equilibrium conditions also as:

$$V_m\left(1+n_m\right) < F_m \leq V_m\left(n_m\right)$$

- The equilibrium conditions imply restrictions on fixed costs and more generally on the parameters in the profit function.
- Using these restrictions and the data, we estimate the parameters in the profit function.

Specification of the variable profit function

- Bresnahan and Reiss (JPE, 1990) do not model explicitly the form of price/quantity competition and consider a flexible model for the variable profit.

$$V_m(n) = S_m \ \left[X_m^v \ \beta^v - \alpha(n)\right]$$

- S_m represents market size.
- X_m^v is a vector is observable market characteristics affecting variable profits, e.g., income, prices of variable inputs, and β^v is a vector of parameters.
- The parameters $\alpha(1)$, $\alpha(2)$, ... capture the competitive effect. We expect:

$$\alpha(1) < \alpha(2) < \alpha(3) \ \ldots \ < \alpha(N)$$

Specification of the fixed cost

- The specification of fixed cost is:

$$F_m = X_m^f \beta^f + \delta(n) + \varepsilon_m$$

- X_m^f is a vector is observable market characteristics affecting fixed costs, e.g., prices of fixed inputs, and β^f is a vector of parameters.

- ε_m is unobservable of the researcher; and error term.

- The parameters $\delta(1)$, $\delta(2)$, ... capture possible competition effects in fixed costs, as well as potential collusive motives.

$$\delta(1) < \delta(2) < \delta(3) \ ... \ < \delta(N)$$

Equilibrium conditions

- The total profit function is:

$$V_m(n) - F_m = (S_m\ X_m^v)\beta^v - X_m^f\ \beta^f - S_m\ \alpha(n) - \delta(n) - \varepsilon_m$$

- Equilibrium conditions: $n_m = n$ is an equilibrium:

$$V_m\,(1+n) < F_m \leq V_m\,(n)$$

- or equivalently:

$$\begin{aligned}(S_m\ X_m^v)\beta^v - X_m^f\ \beta^f - S_m\ \alpha(n+1) - \delta(n+1) \\ < \varepsilon_m \leq \\ (S_m\ X_m^v)\beta^v - X_m^f\ \beta^f - S_m\ \alpha(n) - \delta(n)\end{aligned}$$

Equilibrium conditions [2]

- Suppose that ε_m is independent of (S_m, X_m) and *iid* $N(0,1)$.
- Let $P_m(n)$ represent the probability $\Pr(n_m = n \mid S_m, X_m)$:

$$
\begin{aligned}
P_m(n) &= \Phi\left(S_m\left[X_m^v \beta^v - \alpha(n+1)\right] - X_m^f \beta^f - \delta(n+1)\right) \\
&- \Phi\left(S_m\left[X_m^v \beta^v - \alpha(n)\right] - X_m^f \beta^f - \delta(n)\right)
\end{aligned}
$$

Estimation of the model parameters

- Let θ be the vector of the parameters of the model. $\theta = \left\{ \beta^{v},\ \beta^{f},\ \alpha(1),\ ...,\ \alpha(N), \delta(1), ..., \delta(N) \right\}$.

- We estimate these parameters using a Maximum Likelihood estimator (MLE).

- The likelihood function of this model and data is:

$$
\begin{aligned}
\mathcal{L}(\theta) &= \prod_{m=1}^{M} \Pr(n_m \mid S_m, X_m; \theta) \\
&= \prod_{m=1}^{M} \left[\begin{array}{c} \Phi\left(S_m \left[X_m^{v} \beta^{v} - \alpha(n+1) \right] - X_m^{f} \beta^{f} - \delta(n+1) \right) \\ - \\ \Phi\left(S_m \left[X_m^{v} \beta^{v} - \alpha(n) \right] - X_m^{f} \beta^{f} - \delta(n) \right) \end{array} \right]
\end{aligned}
$$

- The MLE is the value of θ that maximizes $\mathcal{L}(\theta)$.

Answering empirical questions using estimated model

- **[1] Ratio of Entry costs to Variable profits.**
- We can construct the ration: $\frac{F_m}{V_m(1)}$, e.g., in market m, the entry cost is 46% of the variable profit of a monopolist in this market.
- **[2] How strong is competition? How quickly profits decline with n?**
- $\Pi(n) = S\ [p(n) - AVC(q(n))]\ q(n) - F(n)$
- $\Pi(n) \geq 0$ implies $S \geq S^*(n)$ where the threshold market size $S^*(n)$ is:

$$S^*(n) \equiv \frac{F(n)}{[p(n) - AVC(q(n))]\ \ q(n)}$$

- How does $S^*(n)$ depends on n? e.g., under the hypothesis of **contestable markets** $\frac{S^*(n)}{n}$ becomes constant for $n \geq n^*$ with a small value for n^*.

Bresnahan & Reiss (JPE, 1990): Empirical results

- $M = 202$ local markets (small towns)
- Five industries: dentists, doctors, drug stores, plumbers and tire dealers.
- Main Findings:
  - Entry thresholds converge quite fast after the second entrant.
  - After three or four firms, an additional entrant doesn't affect much competition.

Bresnahan Reiss (JPE 1990)

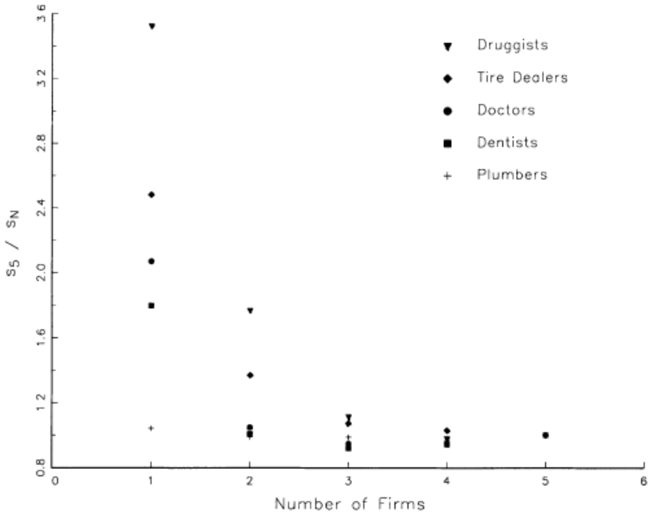


FIG. 4.—Industry ratios of s_5 to s_N by N

Bresnahan & Reiss (JPE, 1990)

- **[Question]** What can we learn about the "nature of competition" in an industry from the empirical relationship between market size (S_m) and market structure/concentration (n_m)?
- $\Pi(n) = S\ [p(n) - AVC(q(n))]\ q(n) - F(n)$
- $\Pi(n) \geq 0$ implies $S \geq S^*(n)$ where the threshold market size $S^*(n)$ is:
$$S^*(n) \equiv \frac{F(n)}{[p(n) - AVC(q(n))]\ \ q(n)}$$
- How does $S^*(n)$ depends on n? e.g., under the hypothesis of **contestable markets** $\dfrac{S^*(n)}{n}$ becomes constant for $n \geq n^*$ with a small value for n^*.

Bresnahan & Reiss (JPE, 1990)

TABLE 5

A. Entry Threshold Estimates

Profession	Entry Thresholds (000's)					Per Firm Entry Threshold Ratios			
	S_1	S_2	S_3	S_4	S_5	s_2/s_1	s_3/s_2	s_4/s_3	s_5/s_4
Doctors	.88	3.49	5.78	7.72	9.14	1.98	1.10	1.00	.95
Dentists	.71	2.54	4.18	5.43	6.41	1.78	.79	.97	.94
Druggists	.53	2.12	5.04	7.67	9.39	1.99	1.58	1.14	.98
Plumbers	1.43	3.02	4.53	6.20	7.47	1.06	1.00	1.02	.96
Tire dealers	.49	1.78	3.41	4.74	6.10	1.81	1.28	1.04	1.03

B. Likelihood Ratio Tests for Threshold Proportionality

Profession	Test for $s_4 = s_5$	Test for $s_3 = s_4 = s_5$	Test for $s_2 = s_3 = s_4 = s_5$	Test for $s_1 = s_2 = s_3 = s_4 = s_5$
Doctors	1.12 (1)	6.20 (3)	8.33 (4)	45.06* (6)
Dentists	1.59 (1)	12.30* (2)	19.13* (4)	36.67* (5)
Druggists	.43 (2)	7.13 (4)	65.28* (6)	113.92* (8)
Plumbers	1.99 (2)	4.01 (4)	12.07 (6)	15.62* (7)
Tire dealers	3.59 (2)	4.24 (3)	14.52* (5)	20.89* (7)

2. Models of Firms' Spatial Location

Models of Firms' Spatial Location

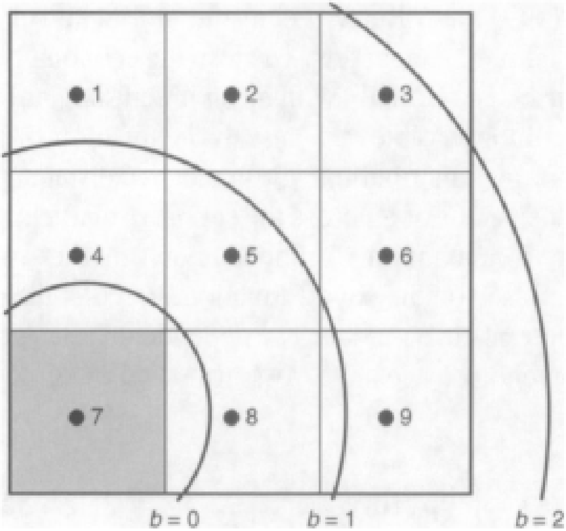
- Consider the decision of retail firm of where to open a new store within a city, e.g., a coffe shop, restaurant, supermarket, department store, etc.
- Different factors can play an important role:
  - Demand: what is the consumer traffic at different locations;
  - Rental prices
  - Location of competitors
- Geographic distance can be an important source of product differentiation. Ceteris paribus, a firm's profit increases with its distance to competitors.
- How profits decline when stores get closer?

Model: The city (1)

- From a geographical point of view, a market (city) is a set, for instance **a rectangle**, in the space $\mathbb{R}^2$.
- Suppose that we divide this city/rectangle into L small squares, each one with its center.
- We can call each of these squares a submarket, or neighborhood, or location.
- A market/city can have hundreds of these submarkets/locations, e.g., $L = 200$.
- We index these locations by $\ell = \in \{1, 2, ..., L\}$

FIGURE 1

IMPACT ON PROFITS OF COMPETITORS' LOCATIONS: ILLUSTRATION

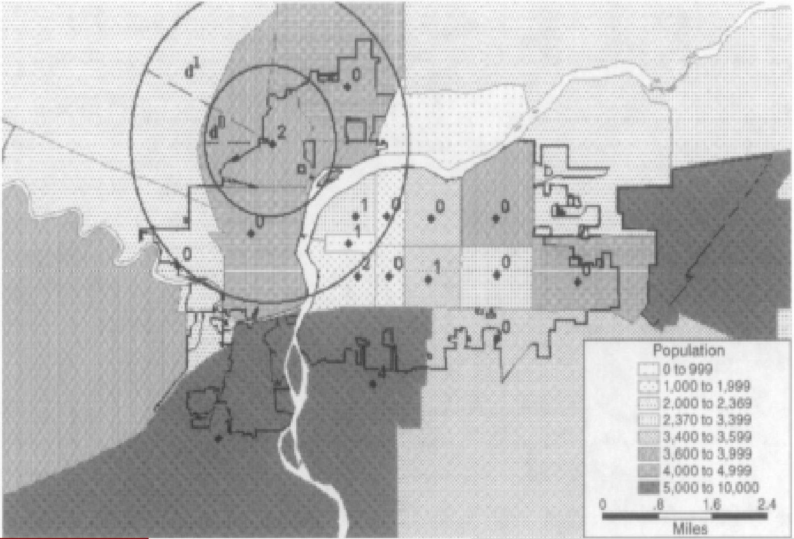


Model: The city (2)

- Each location has some characteristics that can affect demand and costs of a firm in that location:
  - Population; demographic characteristics of the population; rental prices.
- We represent the exogenous characteristics of location ℓ using the vector $\mathbf{x}_{\ell}$.
- Therefore, we can see a city as a landscape of the characteristics $\mathbf{x}_{\ell}$ over the L locations.

FIGURE 2

SAMPLE MARKET: GREAT FALLS, MONTANA



Model: Firms

- There are N potential entrants in this industry (e.g., supermarkets) and city (Toronto).
- In the simpler version of the model, each potential entrant has only one possible store: no multi-store firms (chains).
- We consider this simpler version.
- Let a_i represent the entry / location decision of firm i.

$$a_i \in \{0, 1, ..., L_m\}$$

- $a_{im} = 0$ represents "no entry";
- $a_{im} = \ell > 0$ represents entry in location ℓ.

Model: Profit function

- What is the profit of firm i if it opens a store in location ℓ?
- In principle, we could consider a model of consumer choice of where to purchase (e.g., logit), a model of price competition between active firms; obtain the Bertrand equilibrium of that game, and the corresponding equilibrium profits.
- This approach has several important complications, and it requires having data on prices and quantities at every location.
- Instead, Seim (2006) considers a convenient shortcut.
- Her model does not explicitly specifies consumer choices and price competition, but it incorporates the idea that consumers face transportation costs and this implies that geographic distance with competitors (spatial differentiation) can increase a firm's profit.

Model: Profit function [2]

- Suppose that we draw a circle of radius d around the center point of location ℓ, e.g., a radius of 1km.
- From the point of view of a store located at ℓ, we can divide its competitors in two groups:
  - Close competitors: within the circle of radius d.
  - Far away competitors: outside the circle of radius d.
- Let $N_\ell(close)$ and $N_\ell(far)$ be the number of close and far away competitors relative to location ℓ.
- We can consider a profit function that depends on γ_{close} $N_\ell(close) + \gamma_{far}$ $N_\ell(far)$, where γ_{close} and γ_{far} are parameters to estimate.
- We expect $\gamma_{close} < \gamma_{close} < 0$. The difference between γ_{close} and γ_{far} tell us how important is geographic distance as a form of differentiation to increase profits.

Model: Profit function [3]

- We can generalized this idea to allow for multiple circles, with different radius, around a location the center point of a location ℓ.
- Let $d_1 < d_2 < ... < d_B$ be B different radius of increasing magnitude, e.g., $d_1 = 0.2\ km$, $d_2 = 0.5\ km$, $d_3 = 1\ km$, $d_4 = 2\ km$, ..., $d_{10} = 20\ km$.
- Given these radiuses, we can construct the number of firms with each of the bands defined by these radiuses:

  $N_\ell(1) =$ Number of firms within the circle of radius d_1;

  $N_\ell(2) =$ Number of firms within the band defined by the circles with radiuses d_1 and d_2;

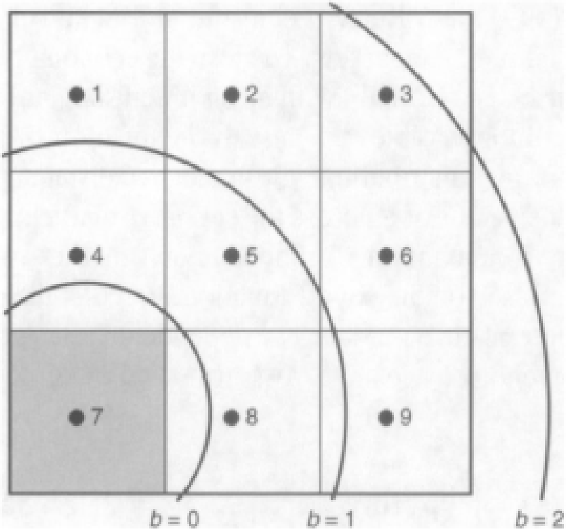
  $N_\ell(3) =$ Number of firms within the band defined by the circles with radiuses d_2 and d_3;

  ...

  $N_\ell(B+1) =$ Number of firms outside the circle with radius d_B.

FIGURE 1

IMPACT ON PROFITS OF COMPETITORS' LOCATIONS: ILLUSTRATION



Model: Profit function [4]

- Profit of an active firm at location ℓ is:

$$\Pi_i(\ell) = x_\ell \ \beta + \xi_\ell + \sum_{b=1}^{B} \gamma_b \ N_\ell(b) + \varepsilon_{i\ell}$$

- We expect:

$$\gamma_1 < \gamma_2 < ... < \gamma_B < 0$$

- ξ_ℓ represents attributes of location ℓ which are known to firms bur unobserved to the researcher.

- ε_{i1}, ε_{i2}, ..., ε_{iL} are assumed iid over firms and locations with extreme value distribution.

Model: Equilibrium (1)

- Suppose that firm i knows the actions of the other firms such that she knows, the landscape of firms: $\varepsilon_{i\ell}$, $N_\ell(1)$, ..., $N_\ell(B)$ for every location $\ell = 1, 2, ..., L$.

- Best response of firm i is to choose location ℓ that maximizes:

$$\Pi_i(\ell) = x_\ell\ \beta + \xi_\ell + \sum_{b=1}^{B} \gamma_b\ N_\ell(b) + \varepsilon_{i\ell}$$

  including the possibility of no entry, $a_i = 0$ with $\Pi_i(0) = 0$.

- Given the logit assumption on $\varepsilon_i(\ell)'s$ the proportion or share of firms with a **best response** of locating in ℓ is:

$$s_\ell = \frac{n_\ell}{N} = \frac{\exp\left\{x_\ell\ \beta + \xi_\ell + \sum_{b=1}^{B} \gamma_b\ N_\ell(b)\right\}}{1 + \sum_{j=1}^{L} \exp\left\{x_j\ \beta + \xi_j + \sum_{b=1}^{B} \gamma_b\ N_j(b)\right\}}$$

Model: Equilibrium (2)

- The equilibrium of the model is described by L simultaneous equations, one for the share of each location ℓ.
- The L are simultaneously determined: Note that $N_\ell(b)$ is just equal to N times the sum of shares s_j in locations j within the band b around location ℓ.
- In equilibrium, a change in x_ℓ in a single location affects the shares s_j at every location in the city.
- Example: Policy that encourages entry in location 1.
  - Direct effect of substitution from other locations.
  - Indirect equilibrium effect: that has the form of the waves generated by a water drop.

Data and Estimation

- Suppose that we have data from an industry (e.g., supermarkets) in a city (Toronto). We observe:

$$\text{Data} = \{x_\ell,\ n_\ell : \ell = 1, 2, ..., L\}$$

  We also know the potential number of entrants, N.

- Given these data, we can construct the shares: $s_\ell : \ell = 1, 2, ..., L$, with:

$$s_\ell = \frac{n_\ell}{N} \quad \text{and} \quad s_0 = \frac{N - n_1 - ... - n_L}{N}$$

- The logit model implies that, for locations with $n_\ell > 0$:

$$\ln\left(\frac{s_\ell}{s_0}\right) = x_\ell\ \beta + \sum_{b=1}^{B} \gamma_b\ N_\ell(b) + \xi_\ell$$

- This is a linear regression model with regressors x_ℓ, $N_\ell(1)$, ..., $N_\ell(B)$, and error term ξ_ℓ.

Inconsistency of OLS

$$\ln\left(\frac{s_\ell}{s_0}\right) = x_\ell\ \beta + \sum_{b=1}^{B} \gamma_b\ N_\ell(b) + \xi_\ell$$

- Regressors $N_b(\ell)$ are endogenous: they are determined in the equilibrium of the model.
- $N_\ell(b)$ is correlated with ξ_ℓ. OLS estimator is biased.
- We expect: $cov(N_1(\ell), \xi_\ell) > 0$ and $cov(N_1(\ell), \xi_\ell) > cov(N_2(\ell), \xi_\ell) > ... > cov(N_B(\ell), \xi_\ell)$
- This implies that OLS estimator of γ_1 is upward biased, and $bias(\gamma_1) > bias(\gamma_2) > ... > bias(\gamma_B)$
- We might wrongly conclude that distance does not affect competition. Example: True γ's: $\gamma_1 = -2$, $\gamma_2 = -1$, $\gamma_3 = -0.5$, and OLS estimates: $\gamma_1^{OLS} = -0.5$, $\gamma_2 = -0.5$, $\gamma_3 = -0.5$.

Instrumental variables estimation

$$\ln\left(\frac{s_\ell}{s_0}\right) = x_\ell\ \beta + \sum_{b=1}^{B} \gamma_b\ N_\ell(b) + \xi_\ell$$

- The model implies instruments for the endogenous regressors $N_\ell(b)$.
- Market characteristics x_j in locations other than ℓ do not enter in the equation for location ℓ but affect the equilibrium values $N_\ell(b)$.
- Let $\overline{x}_\ell(b)$ be the mean value of x_j in the those locations that belong to the band b around location ℓ:

$$\overline{x}_\ell(b) = \frac{\sum_{j=1}^{L} 1\{\text{location } j \text{ belongs to band } b \text{ around } \ell\}\ x_j}{\sum_{j=1}^{L} 1\{\text{location } j \text{ belongs to band } b \text{ around } \ell\}}$$

- We can use $\overline{x}_\ell(b)$ as an instrument for $N_\ell(b)$.

Entry and store location: Results

- Seim (2006) finds very significant results of spatial differentiation (γ parameters decline very significantly with distance)
- Market structure and spatial structure of stores under two different scenarios of city growth.

  - Growth in population but keeping city boundaries.
  - Growth in population and in city boundaries

- The model can be used to study how changes in the exogenous characteristics x_ℓ of a single location (e.g., new amenities, schools, new local regulations, transportation, developments) can affect the landscape of firms in the whole city.