Empirical Industrial Organization

Seim (2006), "An Empirical Model of Firm Entry with Endogenous Product-type Choices"

Sergey Alexeev

University of Technology Sydney

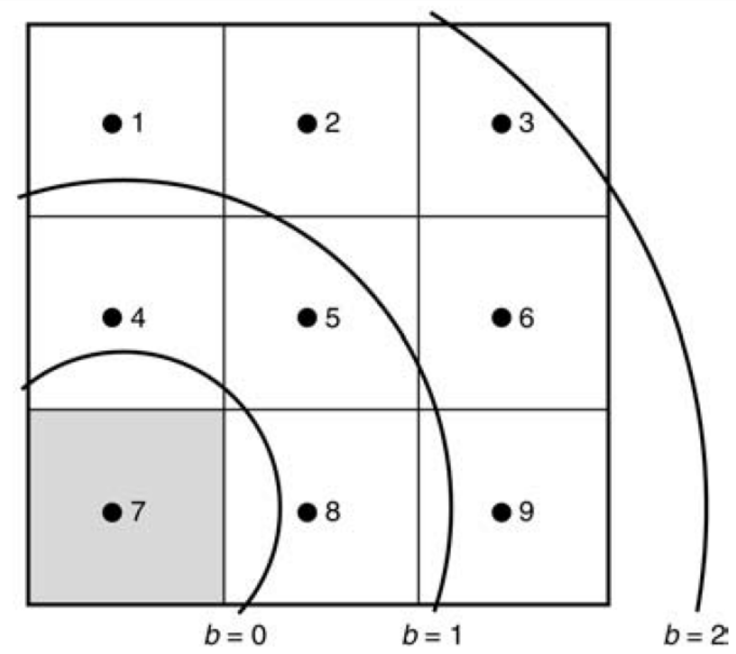
Table of contents

Introduction

- Firms choose a location in abstract characteristics space;
- Product positioning within a market is one of the choices;
- Paper presents empirically tractable equilibrium model to analyze the determinants of firms' product positions:
  - Incomplete-information framework with idiosyncratic sources of profitability (not observed by rivals);
  - E.g. managerial talent, customer service, inventory maintenance;
- With application on a sample of video retailers
- The results support that firms use spatial differentiation to shield themselves from competition.
- The effect is illustrated with counter-factual exercise:
  - Growing market gives firms more local market power (more scope for spatial differentiation);
  - However, payoffs from differentiation are lower as the population is more dispersed (demand falls).

Model

Setup and payoffs


1
2
3
4
5
6
7
8
9
b = 0
b = 1
b = 2

- f in $\mathcal{F}$ simultaneously and independently chooses whether and where enter m;
- The number of actual entrants is $\mathcal{E}$;
- The set of possible locations in m is $\ell = 0, 1, ...\mathcal{L}^m$
- $\mathbf{d}_f$ is f's location decision, where $d_{f_\ell} = 1$ if ℓ is chosen and 0 otherwise

$$\Pi^m_{f_\ell} = \mathbf{X}^m_\ell \beta + \xi^m + h(\Gamma^m_{\cdot\ell}, \mathbf{n}^m) + \varepsilon^m_{f_\ell} \tag{1}$$

- ξ^m and $\mathbf{X}^m_\ell$ are (un)observed demand shifter of m in ℓ;
- $\Gamma = \mathcal{L}^m \times \mathcal{L}^m$ and $\mathbf{n}^m$ is a the number of firms in $\mathcal{L}$;
- $\varepsilon^m_{f_\ell}$ is idiosyncraticity of f in ℓ with private realization and common density.

Assumption 1: *Independent symmetric private values*
$\varepsilon_1^m, ..., \varepsilon_{\mathcal{F}}^m$ are *i.i.d.* with $G(\cdot)$ and private

Assumption 2: *Additively separable marginal competititors affects*
$h(\Gamma_{\cdot \ell}^m, \mathbf{n}^m) = \sum_{k=1}^{\mathcal{L}^m} \gamma_{k\ell} n_k^m$

- Indexing $b = 0, 1, ..., B$ and omitting m gives:

$$\Pi_{f_\ell} = \xi + \mathbf{X}_\ell \beta + \sum_b \gamma_b N_{b\ell} + \varepsilon_{f_\ell} \tag{2}$$

- γ_b competition impact in b:
    - γ_0 for 0 and $\mathcal{D}_1$, γ_1 for $\mathcal{D}_1$ and $\mathcal{D}_2$ etc.;
- Total number of firms in b:
    - $N_{b\ell} = \sum_k \mathbb{I}_{k\ell}^b n_k$, where $\mathbb{I}_{k\ell}^b = 1$ if $\mathcal{D}_b \leq d_{k\ell} < \mathcal{D}_{b+1}$
- Note the summing of $N_{b\ell}$ across bands:
    - $\sum_b N_{b\ell} = \mathcal{E}$

Model

Conjectures and equilibrium

- With imperfect knowledge the expected profit in ℓ:

$$\begin{aligned} \mathbb{E}[\Pi_{f_\ell}] &= \xi + \mathbf{X}_\ell \beta + \sum_b \gamma_b \mathbb{E}[N_{b\ell}] + \varepsilon_{f_\ell} \\ &= \mathbb{E}[\bar{\Pi}_{f_\ell}] + \varepsilon_{f_\ell} \end{aligned} \tag{3}$$

where $\mathbb{E}[N_{b\ell}] = \sum_k \mathbb{I}^b_{k\ell} \mathbb{E}[n_k]$ is expected number of f's in b

- Due to symmetry f's perception of g's location strategy ℓ:

$$\begin{aligned} p_{g\ell}(\, d_{g\ell} = 1 |\, \xi, \mathbf{X}, \mathbb{E}, \theta_1 \equiv (\beta, \gamma)) = \\ \Pr(\mathbb{E}[\bar{\Pi}_{g_\ell}(\cdot)] + \varepsilon_{g_\ell} \geq \mathbb{E}[\bar{\Pi}_{g_k}(\cdot)] + \varepsilon_{g_k}), \\ \forall k \neq \ell, \forall g \neq f \end{aligned} \tag{4}$$

- Then f's expected number of competitors in ℓ is $(\mathcal{E} - 1)p_{g\ell}$;
- It collapses expect number of firms in b to:

$$\mathbb{E}[N_{b\ell}] = \sum_k \mathbb{I}^b_{k\ell} \mathbb{E}[n_k] = \sum_k \mathbb{I}^b_{k\ell} (\mathcal{E} - 1) p_{gk} + \mathbb{I}_{b=0} \tag{5}$$

- Types: $\varepsilon \sim GEV(\mu, \delta, \xi)$;
- This results in multinomial Logit with unidentified δ:

$$p_{g\ell} = \frac{\exp(\mathbb{E}[\bar{\Pi}_{g_\ell}])}{\sum_{k=1}^{\mathcal{L}} \exp(\mathbb{E}[\bar{\Pi}_{g_k}])} \tag{6}$$

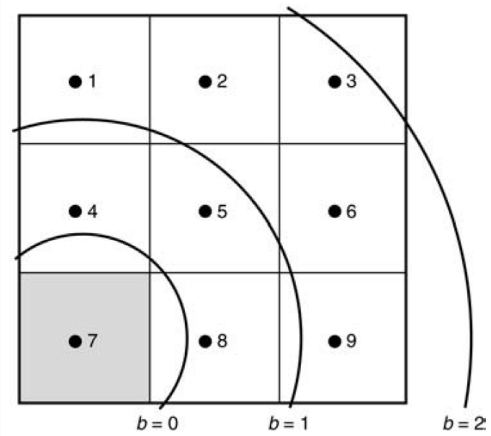
- A1 imply $\mathbf{p_g} = \mathbf{p_f} = \mathbf{p}^*$ and plugging in (3) and (5) into (6) gives a firm's vector of equilibrium conjectures over all ℓ (BNE):

$$\begin{aligned} p_\ell^* &= \frac{\exp(\bar{\Pi}_\ell(\mathbf{X},\mathbf{p}^*,\mathcal{E},\theta_1))}{\sum_{k=1}^{\mathcal{L}} \exp(\bar{\Pi}_k(\mathbf{X},\mathbf{p}^*,\mathcal{E},\theta_1))} \\ &= \frac{\exp(\mathbf{X}_\ell\beta+\gamma_0+(\mathcal{E}-1)\sum_b \gamma_b \sum_j \mathbb{I}_{j\ell}^b p_j^*)}{\sum_{k=1}^{\mathcal{L}} \exp(\mathbf{X}_k\beta+\gamma_0+(\mathcal{E}-1)\sum_b \gamma_b \sum_j \mathbb{I}_{jk}^b p_j^*)} \quad \forall \ell = 1, ..., \mathcal{L}, \end{aligned} \tag{7}$$

- System of $\mathcal{L}$ equations define a fixed point:
    - Exist by Brouwer's FPT and unique under a reasonable assumption (see Appendix).

Impact on profits of competitors' locations: illustration

- Using (3) with (4) gives:
$\mathbb{E}[\bar{\Pi}_7] =$
$\xi + \mathbf{X}_7\beta + \gamma_0 + (\mathcal{E} - 1)(\gamma_0 p_7^* + \gamma_1(p_4^* + p_5^* + p_8^*) + \gamma_2(p_1^* + p_2^* + p_3^* + p_6^* + p_9^*))$



- f's expected number of competitors in a particular distance band is a function of the number of entrants into a market:

$$\mathbb{E}[N_{b\ell}] = \sum_k \mathbb{I}^b_{k\ell}\mathbb{E}[n_k] = \sum_k \mathbb{I}^b_{k\ell}(\mathcal{E}-1)p_{gk} + \mathbb{I}_{b=0} \qquad \text{(repeated 5)}$$

- In equilibrium, the probability of entry involves a comparison of a weighted average of payoffs across locations to the normalized payoff of not entering;
- Note the role of ξ;

$$\Pr(\text{entry}) = \frac{\exp(\xi)\left[\sum_{\ell=1}^{\mathcal{L}} \exp(\bar{\Pi}_\ell(\mathbf{X}, \mathbf{p}^*, \mathcal{E}, \theta_1))\right]}{1 + \exp(\xi)\left[\sum_{\ell=1}^{\mathcal{L}} \exp(\bar{\Pi}_\ell(\mathbf{X}, \mathbf{p}^*, \mathcal{E}, \theta_1))\right]} \qquad (9)$$

- Since probability of entry is identical across competitors:

$$\mathcal{E} = \mathcal{F} \cdot \Pr(\text{entry}) \qquad (10)$$

- Note that (10) through (9) depends on $\mathcal{F}$;

Model

Estimation

- To bypass non-linearity of (7) and (10) an approach close BLP is used;
- Expected number of entrants predicted by (10) assumed to match the data:
  - Done by adjusting ξ, market level effect.
- For observed $\mathcal{E}$ and assumed $\mathcal{F}$ equation (10) with (9) define ξ:

$$\xi = \ln(\mathcal{E}) - \ln(\mathcal{F} - \mathcal{E}) - \ln\left(\sum_{\ell=1}^{\mathcal{L}} \exp(\bar{\Pi}_\ell(\mathbf{X}, \mathbf{p}^*, \mathcal{E}, \theta_1))\right) \tag{11}$$

- In (11) $\xi \sim N(\mu, \sigma)$ estimated on the vector of ξ across the set of M markets.

$$p_\ell^* = \frac{\exp(\mathbf{X}_\ell \beta + \gamma_0 + (\mathcal{E} - 1)\sum_b \gamma_b \sum_j \mathbb{I}_{j\ell}^b p_j^*)}{\sum_{k=1}^{\mathcal{L}} \exp(\mathbf{X}_k \beta + \gamma_0 + (\mathcal{E} - 1)\sum_b \gamma_b \sum_j \mathbb{I}_{jk}^b p_j^*)} \qquad \text{(repeated 7)}$$

$$\Pr(\text{entry}) = \frac{\exp(\xi)\left[\sum_{\ell=1}^{\mathcal{L}} \exp(\bar{\Pi}_\ell(\mathbf{X}, \mathbf{p}^*, \mathcal{E}, \theta_1))\right]}{1 + \exp(\xi)\left[\sum_{\ell=1}^{\mathcal{L}} \exp(\bar{\Pi}_\ell(\mathbf{X}, \mathbf{p}^*, \mathcal{E}, \theta_1))\right]} \qquad \text{(repeated 9)}$$

$$\mathcal{E} = \mathcal{F} \cdot \Pr(\text{entry}) \qquad \text{(repeated 10)}$$

- Each market is treated as an independent $\mathcal{F}^m$;
- The dependent variable consists of a vector of each firm f's observed location choice, stacked across firms and markets.

$$L(\theta_1, \theta_2) = \prod_{m=1}^{M} \mathbf{p}_{\theta_1}(\mathbf{d}^m | \xi^m, \mathbf{X}^m, \mathcal{E}^m) g_{\theta_2}(\xi^m | \mathbf{X}^m, \mathcal{E}^m, \mathcal{F}^m) \quad (12)$$

where $\mathbf{d}^m = (d_1^m, d_2^m, ..., d_{\mathcal{F}}^m)$, g_{θ_2} is density of ξ^m and $\theta_2 = (\mu, \sigma)$

- Computes the likelihood of observing entrants location choices conditional on the market-level effect (i.e. Logit);
- Then multiplying by the probability of observing the particular ξ realization (that equates predicted and actual entrants) gives unconditional likelihood;
- (θ_1, θ_2), $\mathbf{X}$, $\mathcal{F}^m$ and $\mathcal{E}^m$ equation (7) gives approx. to a FP;
- Then $\mathbf{p}^*$, $\mathcal{F}$ and $\mathcal{E}$ and equation (11) gives equilibrium ξ for each m;
- Parameter estimates are obtained by maximizing (12) using a Nelder-Meade optimization algorithm

Data

- The video retail industry:
  - homogeneous and inexpensive good;
  - stores differentiate themselves in the variety and depth of inventory carried, rental contract terms, and drop-off convenience;
  - the main differentiation is spatial location because customers unwilling to travel a long distance.

TABLE 1 **Descriptive Statistics: Markets and Locations**

	151 Sample Markets		
	Mean	Minimum	Maximum
Market Level			
Population, market	74,367	41,352	142,303
Population, main city	59,428	40,495	140,949
Population, all tracts in market	92,563	41,614	193,322
Largest incorporated place within 10 miles	2,618	—	9,972
Largest incorporated place within 20 miles	7,916	—	24,725
Tract Level			
Number of tracts	21.13	8.00	49.00
Number of store locations	18.72	7.00	44.00
Tract population	4,380	247.00	32,468
Area (square miles)	10.10	.10	181.50
Average distance (miles) to other locations in market	3.49	1.08	8.05

Note: The largest incorporated place within 10 and 20 miles is relative to the centroid of the market's main city. The distance between locations within a market is computed as the distance between the tracts' population-weighted centroids. Demographic data are as of 1999.

TABLE 2 **Tract-Level Demographic Characteristics**

	Mean	Minimum	Maximum
Demographic Characteristics			
Population	4,417	247	20,163
Population, within .5 miles of tract	4,952	247	23,676
Population, .5–3 miles of tract	42,281	0	145,499
Population, 3–10 miles of tract	54,817	0	169,271
Per capita income, within .5 miles of tract	17,807	3,484	60,347
Per capita income, .5–3 miles of tract	17,413	0	38,934
Per capita income, 3–10 miles of tract	19,417	0	38,452
Business Characteristics			
Establishment density per square mile	177.86	.15	5,239.48

Note: The tract's total population is placed at the population-weighted centroid. Population within different distance bands to the tract under consideration is computed as the sum of the population in tracts for which the distance to the considered tract's centroid falls within the specified range. Demographic data are as of 1999.

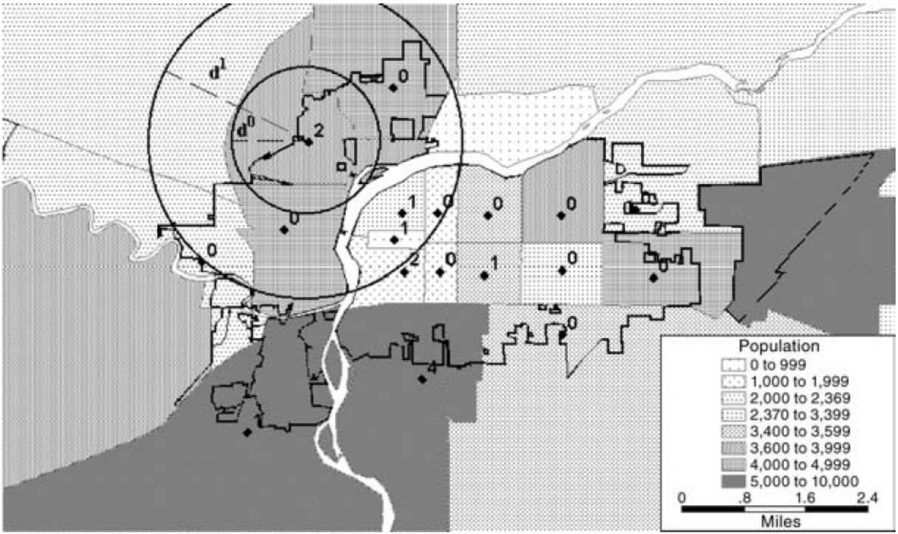
TABLE 3 **Store Location Patterns, Sample Markets**

	Mean	Minimum	Maximum
Firms, market	13.68	4.00	33.00
Store Clustering			
Firms, tract	.73	.00	9.00
Firms, within .5 miles of tract	.80	.00	10.00
Firms, within .5–3 miles of tract	6.12	.00	27.00
Firms, within 3–10 miles of tract	7.94	.00	33.00
Location Patterns within City's Area			
Distance to city center (miles)[a]	3.02	.02	14.96

Note: All stores are placed at the tract's population-weighted centroid. Competitors within different distance bands to a firm's location are computed as the number of firms in tracts for which the distance to the firm's tract falls in the specified range.

[a] The city center is taken to be the population-weighted centroid of the market's main city.

Sample market: Great Falls, Montana



Results

Parameter estimates

Parameter estimates, entry and location-choice model

TABLE 4 **Parameter Estimates, Entry and Location-Choice Model**

	Potential Entrant Pool			
	2 × Total Entrants		50 Firms	
Variable	Coefficient (Standard Error)	Marginal Effect	Coefficient (Standard Error)	Marginal Effect
Population$_0$ (000)	1.8191	.0333	2.1258	.0393
	(.1534)		(.1764)	
Population$_1$ (000)	1.3109	.0236	1.7349	.0314
	(.1200)		(.1498)	
Population$_2$ (000)	.6070	.0121	1.1348	.0227
	(.1192)		(.1486)	
Business Density	−.8077	−.0155	−.8889	−.0173
	(.1458)		(.1477)	
Average Per Capita Income$_0$ (0000)	.9309	.0180	1.0380	.0204
	(.1136)		(.1233)	
Average Per Capita Income$_1$ (0000)	1.0081	.0193	.9188	.0178
	(.2081)		(.2043)	
Average Per Capita Income$_2$ (0000)	.4851	.0092	.4884	.0094
	(.2512)		(.2601)	
γ_0	−3.4520		−3.3853	
	(.3111)		(.3266)	
γ_1	−1.0103		−1.0087	
	(.0745)		(.0923)	
γ_2	−.3448		−.4870	
	(.0738)		(.0934)	
σ	3.5829		4.6760	
	(.3110)		(.4316)	
μ	−2.8764		−7.0364	
	(1.3425)		(1.5801)	

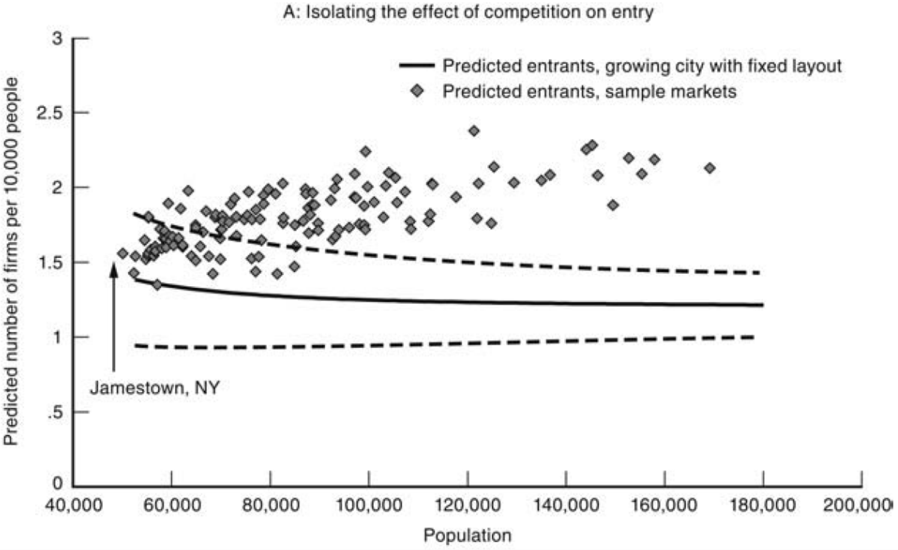
Note: Results based on 1999 demographic and firm data. Subscript 0 denotes the immediately adjacent locations to the chosen tract, within .5 miles in distance; subscript 1 denotes tracts at .5 to 3 miles in distance from the chosen tract; and subscript 2 denotes tracts at more than 3 miles distance from the chosen tract. Tract-level business density is defined as the number of establishments (0000) per square mile. γ denotes competitive effects, and σ and μ are the estimates of the parameters of the distribution of ξ.

Results

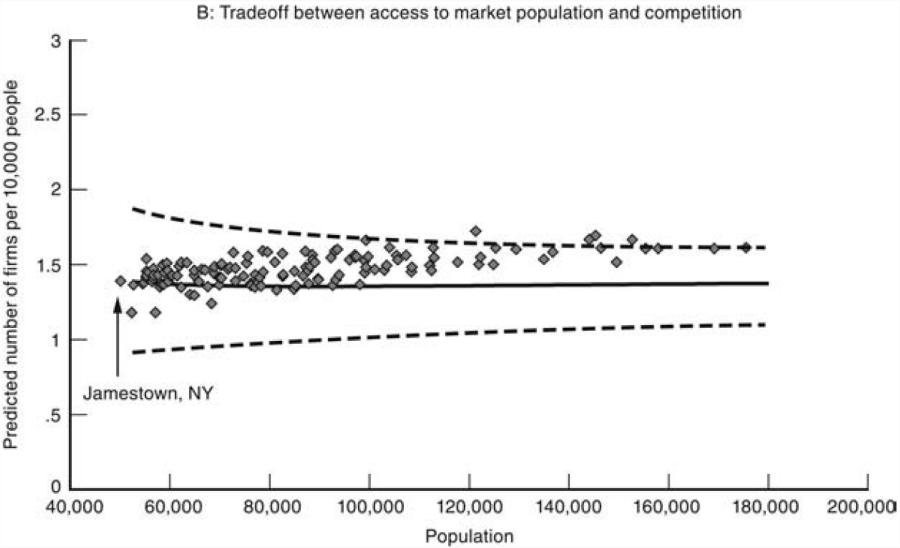
Illustration of results

- The lessening of competitive effects imply that geographic dispersion in demand is used to avoid competition;
- That means that more stores enter as the market area and scope for differentiation grow;
- A counter-factual increase of size of the characteristic space can be done;
    - Note that as city grows it spreads out and population increases as well;
    - So exercise is done in two steps:
        1. Allows a city to grow in population only, holding its geographic layout;
        2. Predicted entry under the expansion path is then contrasted with entry that would occur were the city to grow both in population and area.

The role of spatial dispersion on entry



The role of spatial dispersion on entry



References

Seim, K. (2006). “An Empirical Model of Firm Entry with Endogenous Product-type Choices”. In: *The RAND Journal of Economics* 37.3, pp. 619–640.