March of the Chains: Herding in Restaurant Locations

Nathan Yang

Yale School of Management

Taking the plunge

Retail chains diving into new markets

- What if retailers face market uncertainty that can only be resolved after entry?
- Possible examples:
    - Customer tastes (Bell and Shelman, 2011)
    - Anti-American sentiment (Beamish, Jung, and Kim, 2011)
    - Health consciousness of consumers (Lawrence, Requejo, and Graham, 2011).
    - Employee turnover (Moon, 2003).

Could learning drive retail clustering?

- Retail outlet locations can be seen by all.
  - Nothing to prevent managers from taking advantage of information revealed through past decisions of rivals.
  - Herd into markets others appear to have done well in (i.e., survived).
- Clustering among rivals is well-documented in recent research:
  - Fast food chains are more likely to open new outlets in markets with existing rival outlets (Toivanen and Waterson, 2005; Shen and Xiao, 2011).
  - Smaller rivals follow larger retail banks into the same rural markets (Damar, 2009; Feinberg, 2008).
  - Rival anchor stores have a tendency to locate in the same shopping centers (Vitorino, 2008).
- Alternative explanations:
  - Unobserved heterogeneity.
  - Demand externalities.

Research objective

- Introduce and estimate new dynamic model of entry/exit with:
  - Strategic interactions.
  - Forward looking retailers.
  - Unobserved heterogeneity.
  - **Common uncertainty.**
  - **Learning through entry.**
  - **Learning from others.**
- Can derive simple differences-in-differences test for learning based on the model.
- Counterfactual analysis to determine whether presence of learning induces retail clustering.

Setting

- Canadian hamburger fast food industry from 1970 to 2005.
  - Entry and exit decisions across *small* geographic markets by the 5 major retail chains: A & W, Burger King, Harvey's, McDonald's, and Wendy's.
- Fast food industry popular setting for studying retail clustering.
  - Thomadsen (2007, 2010), Toivanen and Waterson (2005), and Shen and Xiao (2011).

Related literature

- **Retail clustering**.
  - Demand externalities (Datta and Sudhir, 2011; Konishi, 2005; Eppli and Benjamin, 1994; Thomadsen, 2010; Zhu, Singh, and Dukes, 2011), and unobserved heterogeneity (Thomadsen, 2007).
- **Retail competition**.
  - Convenience stores (Nishida, 2008), discount retailers (Ellickson, Houghton and Timmons, 2010; Jia, 2008), fast food (Toivanen and Waterson, 2005), hotels (Suzuki, 2010), and video rental services (Seim, 2006).
- **Social spillovers**.
  - Book sales (Chevalier and Mayzlin, 2006), economic policies (Buera, Monge-Naranjo, and Primiceri, 2010), farming technology (Conley and Udry, 2010), kidney adoption (Zhang, 2010), movie sales (Moretti, 2010), voting (Knight and Schiff, 2007).

Market definition and observable market characteristics

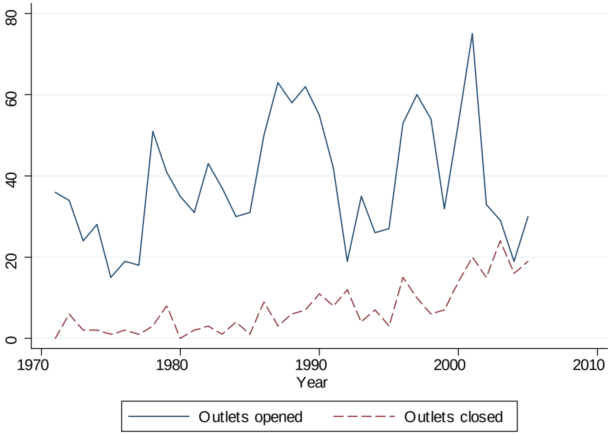
- Forward Sortation Areas (FSA) nested within all Canadian cities.
  - First three digits of Canadian postal code.
  - 608 FSA markets identified in sample.
  - In most cities, average FSA is 1.8 square miles in area.
  - FSAs *smaller* than markets used in previous studies (Ellison, Glaeser, and Kerr, 2010; Shen and Xiao, 2011; Toivanen and Waterson, 2005).
- Match each market with the following characteristics:
  - Population, population density, income, property value, presence of university, proportion of residents working in same FSA, total retail sales, and total number of retail locations.

Entry and exit decisions

- Obtained using archived phone directories from the City of Toronto's Reference Library.
  - 36 annual editions per city $\times$ 30 cities $=$ 1,080 phonebooks searched.
- Track each outlet and get the following:
  - Opening year.
  - Closing year.
  - Exact location.

Aggregate dynamics

Figure: Total number of outlets opened/closed in Canada over time.



Transitions between being active and not active

Table: Tabulation of the lagged active statuses.

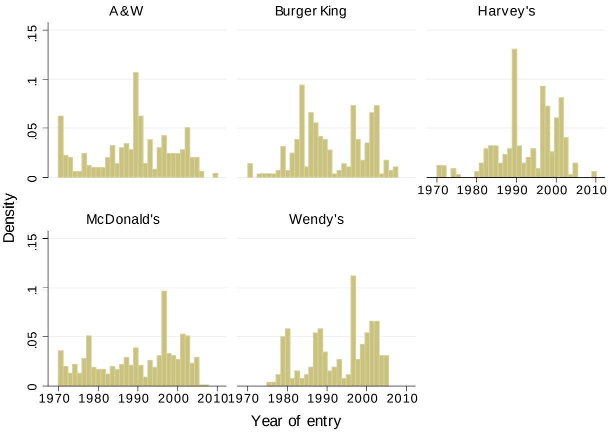
Active two periods ago	0		1	
Active one period ago	0	1	0	1
A & W	16,904	264	96	3,408
Burger King	18,092	200	37	2,343
Harvey's	17,943	228	70	2,431
McDonald's	11,471	449	2	8,750
Wendy's	18,448	177	28	2,019

First movers

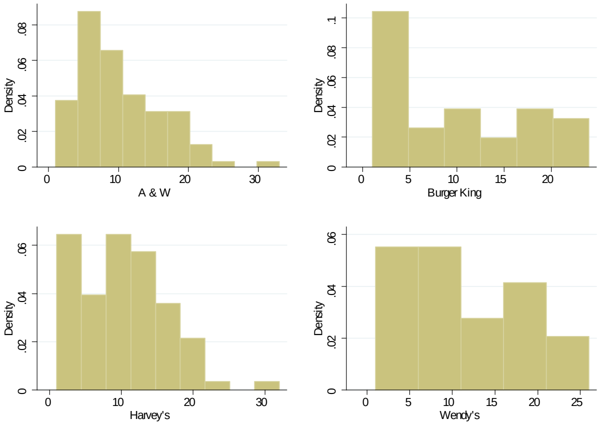
Table: Tabulation of the total number of markets that a chain was the (unique) first entrant.

Chain	First entrant
A & W	100
Burger King	50
Harvey's	65
McDonald's	334
Wendy's	34

Time of entry



Time of exit


Density
0
.02
.04
.06
.08
0
10
20
30
A & W
Density
0
.02
.04
.06
.08
.1
0
5
10
15
20
Burger King
Density
0
.02
.04
.06
0
10
20
30
Harvey's
Density
0
.02
.04
.06
0
5
10
15
20
25
Wendy's

Market characteristics

Table: Summary statistics for markets that were occupied in 1970, and for markets that were occupied after 1970.

	Occupied 1970		Occupied after 1970	
Variable	Mean	Std. Dev	Mean	Std. Dev
Population (persons)	21,144	7,433	23,895	12,809
Population density (persons per sq km)	2,892.93	3,276.488	1,615.26	2,271.38
Total sales (billion CDN)	1.410	1.160	2.330	1.170
Total retail locations	483	364	850	408
Income (dollars)	57,579	14,082.81	55,518.77	18,571.69
Property value (million CDN)	0.322	0.168	0.259	0.161

Specification

$$
\begin{aligned}
\Pr(a_{imt} &= 1|\mathbf{a}_{mt-1}, \mathbf{Z}_{mt}) = \Phi(\alpha_i + \mathbf{Z}_{mt}\boldsymbol{\beta}_i + \sum_{j \neq i} \gamma_{ij} a_{jmt-1} \\
&+ \rho_i t + \eta_m + \varsigma_i t \cdot \eta_m)
\end{aligned}
$$

- a_{imt} is chain i's decision to be active in market m at time t.
- $\mathbf{Z}_{mt}$ are market characteristics.
- a_{jmt-1} is rival j's decision to be active in market m at time $t-1$.
- $\rho_i t$ is time trend.
- η_m is unobserved heterogeneity.

Results

Table: Evidence of clustering based on the chains' decision to be active in market.

	(1) A & W	(2) Burger King	(3) Harvey's	(4) McDonald's	(5) Wendy's
A & W incumbent	3.952***	0.0712	0.0946	0.0541	0.305***
	(0.0709)	(0.0897)	(0.0894)	(0.0875)	(0.0910)
Burger King incumbent	0.363***	4.443***	0.247*	0.214	0.0169
	(0.0990)	(0.119)	(0.108)	(0.137)	(0.124)
Harvey's incumbent	0.00462	0.186	4.231***	-0.0241	0.294**
	(0.0939)	(0.102)	(0.0916)	(0.122)	(0.109)
McDonald's incumbent	0.0614	0.181*	0.364***	4.621***	0.481***
	(0.0715)	(0.0817)	(0.0745)	(0.328)	(0.0841)
Wendy's incumbent	0.385***	0.273*	0.0558	0.0851	4.617***
	(0.102)	(0.114)	(0.109)	(0.168)	(0.137)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	20930	20930	20930	20930	20930

Clustered standard errors (by FSA) in parentheses

Basic setting

- Each chain simultaneously decides whether or not to be active (a_{imt}) in market m at the beginning of time t.
- Chains maximize discounted payoffs:

$$\sum_{s}^{\infty} \beta^{t+s} \Pi_{imt+s}$$

- Π_{imt+s} is the one-shot payoff.
- β is inter-temporal discount rate.

One-shot payoff

$$\begin{aligned}\Pi_{imt}(a_{imt} &= 1) = S_{mt}\theta_{1i} + \sum_{j\neq i}\theta_{2ij}a_{jmt} - FC_i - (1-a_{imt-1})EC_i \\ &+\omega_m - \varepsilon_{imt}.\end{aligned}$$

- $S_{mt} = \mathbf{Z}_{mt}\boldsymbol{\beta}$ is market size.
- θ_{1i} is firm specific fixed effect.
- θ_{2ij} is competitive/complementary effect that rival j has on chain i.
- FC_i is chain i's fixed cost.
- EC_i is chain i's entry cost.
- ω_{im} is permanent market characteristic unknown to empiricists.
- ε_{imt} is chain i's privately known and idiosyncratic shock.

Beliefs about market unobserved heterogeneity

$$\omega_m = \begin{cases} \eta_m(1+\sigma_i) & \text{w.p.} \quad \lambda_{imt} \\ \\ \eta_m & \text{w.p.} \quad 1-\lambda_{imt} \end{cases}.$$

- λ_{imt} is posterior probability of being *uninformed* (i.e., face uncertainty).
- The prior is λ_0.
- σ_i is degree of uncertainty for chain i.
- Assume $\eta_m \in \{-\rho, \rho\}$, such that $\eta_m = \rho$ with probability φ_1.
- Therefore, $E_i(\omega_m|\mathbf{\Omega}_{imt}) = \eta_m + \lambda_{imt}\eta_m\sigma_i$.

Two ways to learn

- **Learning through entry:** Within a year of entering a market, a retailer resolves its uncertainty about the size of the market.
  - $\lambda_{imt} = 0$ if the retailer entered at time $t-1$.
  - $\lambda_{imt+s} = 0$ for all $s > 0$ if $\lambda_{imt} = 0$.
- **Learning from others:** A potential entrant who has not previously entered (and left) the market already can learn from the observed past decisions of their informed rivals.
  - Updates the beliefs, λ_{imt}, using Baye's rule and observed past stay/exit decisions among informed rivals.

Bayesian updating notation

- Set of informed retailers who made informed decisions at $t-1$ is J^*_{mt}.
- Vector of informed decisions made at $t-1$ is $\mathbf{a}^*_{mt-1}$.

$$\lambda_{imt} = \frac{\Pr(\mathbf{a}^*_{mt-1}|\boldsymbol{\omega}_m \neq \mathbf{0})\lambda_{imt-1}}{\Pr(\mathbf{a}^*_{mt-1}|\boldsymbol{\omega}_m \neq \mathbf{0})\lambda_{imt-1} + \Pr(\mathbf{a}^*_{mt-1}|\boldsymbol{\omega}_m = \mathbf{0})(1-\lambda_{imt-1})}$$

$$\Pr(\mathbf{a}^*_{mt-1}|\cdot) = \prod_{j \in J^*_{mt}} P_{jm}(\cdot)^{a_{jmt-1}} \cdot (1 - P_{jm}(\cdot))^{(1-a_{jmt-1})}$$

Markov Perfect Equilibrium (MPE)

$$\varrho_i(\mathbf{X}_{mt}, \varepsilon_{imt}, \eta) = \arg \max_{a_{imt} \in \{0,1\}} E\left[\Pi^{\varrho}_{imt} + \beta \; V^{\varrho}_i(\mathbf{X}_{mt+1}, \varepsilon_{imt+1}, \eta_m) | \cdot\right]$$

- Strategies $\{\varrho_i(\mathbf{X}_{mt}, \varepsilon_{imt}, \eta_m)\}_i$ assumed to depend on state variables, $(\mathbf{X}_{mt}, \varepsilon_{imt}, \eta_m)$ where

$$\mathbf{X}_{mt} = \{\mathbf{a}_{mt-2}, \mathbf{a}_{mt-1}, \boldsymbol{\lambda}_{mt-1}, \mathbf{Z}_{mt}\}.$$

- $V^{\varrho}_i(\mathbf{X}_{mt+1}, \varepsilon_{imt+1}, \eta_m)$ is the continuation value.
- Π^{ϱ}_{imt} is one shot payoff evaluated at strategies $\{\varrho_i(\mathbf{X}_{mt}, \varepsilon_{imt}, \eta_m)\}_i$.
- Integrating strategy function with respect to ε_{imt} yields best response function $P_i(\mathbf{X}_{mt}, \eta_m)$.
- MPE obtained as fixed point.

Identification of structural model

- Strategic interactions (θ_{2ij}).
    - Chain's incumbency status has direct impact on its own flow profits through entry costs, but will only affect rival through best response probability.
    - This is true if chain was not already active 2 periods earlier, or if rival no longer faces uncertainty.
    - Need sufficient variation in a_{imt-2} and a_{imt-1}.
- Learning (λ_0, σ_i).
    - 40 out of 608 markets for which chain re-enters a market.
    - The first time it entered, most likely faced uncertainty, but second time, chain no longer faces uncertainty.
    - In both cases, η_m the same, but $(\lambda_{imt}, \sigma_i)$ enters through the payoff only in first case.
    - Timing of first entry helps identify λ_0.

Simple DID test for learning

$$\delta = [\Pi_i^{\mathbf{P}}(0,1) - \Pi_i^{\mathbf{P}}(0,0)] - [\Pi_i^{\mathbf{P}}(1,1) - \Pi_i^{\mathbf{P}}(1,0)]$$

- Lets focus on two chains.
- Chain i is either a potential entrant or incumbent, while its rival j either stayed or exited at $t-1$.
- Set $\beta = 0$ and let ε_{imt} be uniformly distributed.
- $\Pi_i^{\mathbf{P}}(a_{imt-1}, a_{jmt-1})$ is one-shot payoff given state (a_{imt-1}, a_{jmt-1}).

Simple DID test for learning

- Under null hypothesis of no learning ($\sigma_i = 0$):

$$\delta_0 = \theta_{2ij}\theta_{2ji}\{[P_i(0,1) - P_i(0,0)] - [P_i(1,1) - P_i(1,0)]\}.$$

- Also possible to write δ_0, based on assumptions above, as:

$$\delta_0 = [P_i(0,1) - P_i(0,0)] - [P_i(1,1) - P_i(1,0)].$$

- Therefore learning holds *iff* DID is zero:

$$\sigma_i = 0 \Leftrightarrow \delta_0 = [P_i(0,1) - P_i(0,0)] - [P_i(1,1) - P_i(1,0)] = 0.$$

DID regression

- Let $\Pi_i^{\mathbf{P}}(0,1) = \Pi_i^{01}$, $\Pi_i^{\mathbf{P}}(0,0) = \Pi_i^{00}$, $\Pi_i^{\mathbf{P}}(1,1) = \Pi_i^{11}$, and $\Pi_i^{\mathbf{P}}(1,0) = \Pi_i^{10}$.
- Based on the assumptions above, and some algebra, regression can be written as:

$$\begin{aligned} E(a_{imt}|\mathbf{a}_{mt-1}, \mathbf{Z}_{mt}) &= \Pi_i^{00} + (a_{imt-1} + a_{jmt-1})(\Pi_i^{10} - \Pi_i^{00}) \\ &\quad - a_{imt-1}a_{jmt-1}\delta + \mathbf{Z}_{mt}\boldsymbol{\beta}_i + \eta_m \end{aligned}$$

- $\mathbf{Z}_{mt}$ are observable market characteristics, and η_m is market fixed effect.
- DID test is $H_0 : \delta = 0$, where rejection of null provides evidence in favor of learning.
- With data, get estimates of δ significant at 10-25% level.

Estimation and counterfactuals

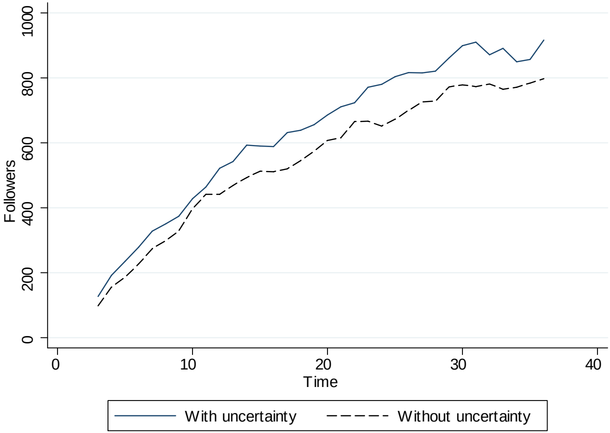
- Use MPE best response probabilities in likelihood-based estimation procedure.
- I embed the following methods into one algorithm:
  - Grid search for prior λ_0.
  - Aguirregabiria and Mira's (2007) NPL to calibrate CCPs.
  - Arcidiacono and Miller's (2011) iterative Expectation-Maximization method for the permanent unobserved heterogeneity.
- Consider counterfactual scenario in which there is **no learning**.
  - Set $\sigma_i = 0$, solve model, and then simulate dynamics.
  - Compare counterfactual with equilibrium in data in which there is **learning**.
  - Is there more/less herding behavior with learning?

Structural estimates

Table: Structural estimation of dynamic entry/exit model.

	A & W	Burger King	Harvey's	McDonald's	Wendy's
Brand value (θ_{1ii})	0.08 (0.01)	0.01 (0.01)	0.01 (0.01)	0.03 (0.01)	0.02 (0.01)
vs A & W (θ_{2iAW})		0.05 (0.03)	0.1 (0.04)	-0.2 (0.03)	0.1 (0.03)
vs Burger King (θ_{2iBK})	-0.2 (0.04)		-0.04 (0.04)	-0.4 (0.03)	-0.05 (0.04)
vs Harvey's (θ_{2iHARV})	0.2 (0.03)	0.04 (0.03)		-0.2 (0.03)	0.09 (0.04)
vs McDonald's (θ_{2iMCD})	-0.3 (0.04)	-0.003 (0.04)	-0.09 (0.04)		-0.01 (0.03)
vs Wendy's (θ_{2iWEND})	0.05 (0.03)	0.03 (0.03)	0.07 (0.03)	0.07 (0.03)	
Fixed costs (FC_i)	-0.04 (0.04)	0.07 (0.04)	0.004 (0.05)	-0.4 (0.05)	0.07 (0.04)
Entry costs (EC_i)	0.1 (0.01)	0.03 (0.02)	0.08 (0.01)	0.02 (0.008)	-0.04 (0.01)
Degree of uncertainty (σ_i)	-0.2 (0.02)	0.03 (0.02)	-0.2 (0.02)	-0.3 (0.01)	-0.01 (0.02)
Prob. of uncertainty (λ_0)	0.2 (0.01)				
Good state parameter (ρ)	0.98 (0.14)				
Prob. of good state (φ_1)	0.45				

Implications of learning: Herding to good markets

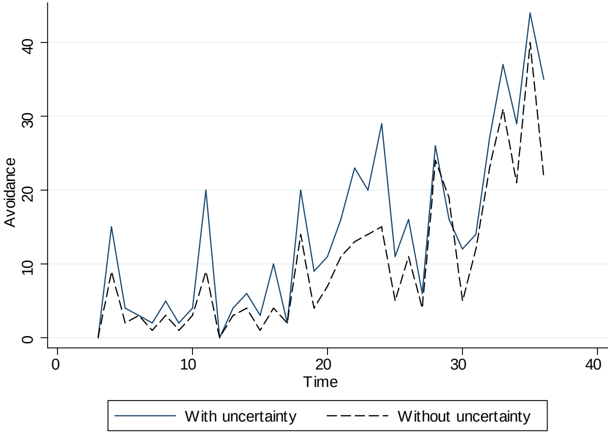


Implications of learning: Strategic delay

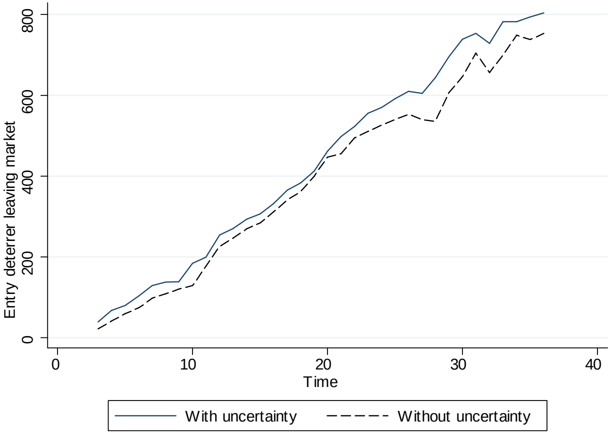
Table: Average number of years before first entering a market.

	With uncertainty	Without uncertainty
A & W	5.0	4.0
Burger King	3.3	4.5
Harvey's	3.3	8.2
McDonald's	7.7	5.8
Wendy's	11.7	11.9

Implications of learning: Avoiding bad markets



Implications of learning: Potential cost of entry deterrence



Future directions

- What are forces that work against/opposite of learning from others?
  - Learning from experience within the same-brand network.
    - Learning that low income markets are better than high income markets.
  - Reputation effects from spatial predation.
    - Understanding that McDonald's will punish those that locate too close to it.

Thank you!