March of the Chains: Herding in Restaurant Locations

Nathan Yang[†]
Yale School of Management

October 22, 2012

Abstract

Does learning from others generate retail clusters? Uncertainty about new markets provides an opportunity for learning from others, where incumbents' past stay/exit decisions are informative to potential entrants. The setting is Canada's fast food industry from 1970 to 2005, where I present a new estimable dynamic oligopoly model of entry/exit with unobserved heterogeneity, common uncertainty about pro…tability, learning through entry, and learning from others. With the estimated model, I …nd that learning induces retailers to herd into markets that others have previously done well in, avoid entering markets that others have previously failed in, and for some, strategically delay entry. Finally, I show that entry deterrence may come at a cost, in the form of added risk from entering early.

Keywords: Agglomeration, dynamic discrete choice game, market structure, retail industry.

I am thankful for supervisvifr u8423(r)-112(u)-flen rb.372cm82ra, andu.97 3(a)-14(r)-112eful
gr-flefully acknowedge fundu.72cmc(n)-16(g)-284(-t23(r)-112(u)--11.293(S)-16(o)-43(112(u)cm82a)-14(l)-277(S)-16(112(3cmc(e10(v)-1416(
y

1 Introduction

Retail managers are often faced with the di¢ cult decision of where to place their stores.[1] Such decisions are challenging because of the uncertainty retailers face; especially so if this uncertainty cannot be fully resolved via market research. For instance, American retailers may be uncertain about a market's tastes (Bell and Shelman, 2011), anti-American sentiment (Beamish, Jung, and Kim, 2011), and health consciousness (Lawrence, Requejo, and Graham, 2011). In some cases, it is only by diving into a market that such uncertainty would be resolved (i.e., learning through entry). But upon entering a market, subsequent stay/exit decisions are publicly seen, and thus, prospective entrants can infer market pro…tability based on such observations (i.e., learning from others). In fact, it has been conjectured by Toivanen and Waterson (2005) in their study of Burger King and McDonald's in the United Kingdom, as well as Shen and Xiao (2012) in their study of Kentucky Fried Chicken and McDonald's in China, that learning from others may explain the commonly observed clustering of seemingly rival retail chains.[2] Similar patterns have also been documented for the retail banking industry (Damar, 2009; Feinberg, 2008), as well as department stores (Vitorino, 2008).

In past literature about retail, researchers have posited unobserved heterogeneity and demand externalities as typical explanations for retail clustering. A nearby mall, local attraction, or highway exit can easily generate retail agglomeration among rivals (Orhun, 2012; Thomadsen, 2007), as can restrictive retail zoning provisions (Datta and Sudhir, 2011) - both factors pointing towards unobserved heterogeneity. Alternatively, a store may generate demand externalities for neighboring rivals if its presence helps draw in additional consumer tra¢ c (Datta and Sudhir, 2011; Eppli and Benjamin, 1994; Konishi, 2005), or if its close proximity can credibly soften price competition via market segmentation or cannibalization[3] concerns (Thomadsen, 2010; Zhu, Singh and Dukes, 2011).

Despite the well-developed(elop)-28(a)28(w)enattr37(at)1(u)-1(uc)27(91(s)-1(91(if)-240(iattr3ds)lein)-1(s.78(f information that can possibly be revealed when an existing and informed chain decides to stay or exit a market. My objective is to understand how these externalities will a¤ect an industry, and whether they contribute to behavior consistent with clustering. The setting for my analysis is Canada's fast food industry, where I study the entry/exit decisions of the …ve major fast food chains in Canada - A & W, Burger King, McDonald's, and Wendy's, along with the Canadian chain Harvey's - from the industry's beginning[5] around 1970 to 2005 - across small geographic markets nested within all Canadian cities (Section 2).

Section 3 presents a descriptive empirical regularity that shares similarities with previous studies (Shen and Xiao, 2012; Toivanen and Waterson, 2005). In particular, I …nd that the incumbency status of a chain has a positive e¤ect on its rivals' decisions to enter a local market, even when (time-varying) unobserved heterogeneity is accounted for. A consistent theme throughout this empirical analysis is that fast food chains tend to follow their rivals into markets. These patterns are certainly suggestive of clustering. Not surprisingly, the fast food industry has become an increasingly popular laboratory for studying retail agglomeration (Thomadsen, 2007, 2010; Toivanen and Waterson, 2005).

This paper inord5

allowing the retailers in my model to be forward looking, they can react appropriately to information externalities. For instance, a potential entrant may have an incentive to strategically delay entry as dictated by the option value of waiting (Chamley, 2004), while an incumbent may anticipate increased competition in the future as its decision to stay in the market may cause rational herding. Under a static setting, such behavior is restricted.

Identi…cation of the model's parameters is discussed in Section 5, where key issues pertain to how learning can be separately identi…ed from unobserved heterogeneity and strategic interactions. The intuition behind identi…cation of learning is as follows: unlike unobserved heterogeneity and strategic interactions, a retailer will react di¤erently to its rival's past decision to stay/exit depending on whether the retailer is an uninformed potential entrant, or an informed incumbent; therefore, learning, unobserved heterogeneity, and demandrr94(t)1vi(and)-399(wil(and)-2eil(and)-281(i)1(on)-1(dp368(de)-1

McDonald's and Wendy's. In Canada, no other chains with national presence entered the industry but failed as a whole. Hence, the set of ...ve chains I look at is very representative of hamburger fast food chains in Canada. Note that there exist quick-service outlets that do not serve hamburgers, such as Kentucky Fried Chicken, Subway, and Taco Bell, which I leave out from my analysis largely because the products o¤ered by hamburger chains are likely to be more substitutable with one another. Furthermore, these chains are late entrants into Canada relative to the hamburger chains. Although Kentucky Fried Chicken was available as early as 1953, it was primarily served through convenience stores until the 1980s. Subway's ...rst outlet in Canada was opened in 1986, while Taco Bell's ...rst outlet in Canada was opened in 1981.

Since 1970, Canada has become a very important foreign market for American retail chains. Canada provides American chains a real growth option,[7] without the risk associated with more exotic markets overseas (Holmes, 2010). Not surprisingly, American chains tend to launch in Canada ...rst before they expand to other countries (Smith, 2006); this strategy is a general phenomenon seen in the entire retail industry. In fact, McDonald's was largely motivated to expand globally after its success in Canada (Love, 1995). Using Canada as a stepping stone, all four of the American chains are currently active players in the global fast food industry. Today, McDonald's has almost 31,000 outlets around the world, Burger King has 4,000 outlets, then A & W follows with about 700, and 400 for Wendy's internationally. The largest domestic chain, Harvey's, boasts a store count of over 200 outlets in Canada.

Many of these franchises were founded in the United States prior to 1970. A & W in 1956, Burger King in 1952, McDonald's in 1952, and Wendy's in 1969; Canada's chain Harvey's was founded in 1959. The ...rst American chains to set up in Canada were A & W (1956), and McDonald's (1967). Although their relative standings have changed over time, these ...ve chains areivivoutlartisga28(e)-411

Table 1: Coverage of CMAs in sample.

Province	Cities
Alberta	Calgary, Edmonton
British Columbia	Vancouver, Victoria, Kelowna, Abbotsford
Manitoba	Winnipeg
New Brunswick	Moncton, Saint John
Newfoundland	St. John's
Nova Scotia	Halifax
Ontario	Toronto, Ottawa, Hamilton, London, Windsor, Niagra Falls, Peterborough, Guelph, Kitchener, Kingston, Oshawa, Barrie, Brantford, Sudbury, Thunder Bay
Saskatchewan	Saskatoon, Regina

United Kingdom, which are equivalent to cities. Ellison, Glaeser, and Kerr (2010) use Primary Metropolitan Statistical Areas, Counties, and States; all of which are larger than FSAs. Finally, Shen and Xiao (2011) focus on city markets in China. I find 608 FSA markets based on the cities used in my sample. Because this study is focused on understanding retail clustering, we need a market definition that is as small as possible. One nice feature of the FSA market definition is that they were established well before the fast food chains entered Canada, and that all of the FSA market definitions in my sample have not undergone changes over time.

The FSA regions I sample are those nested within Canada's Census Metropolitan Areas (CMAs), or loosely speaP(samp)sas

Table 2: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Population (persons)	20,333	11,206	44	89,686
Population density (persons per sq km)	2,344.034	3,487.339	0.186	144,908.844
Total sales (billion CDN)	1.087	1.100	0.001	9.155

Table 3: 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

2.3 Entry and exit data

I turned to archived phone books at the City of Toronto's Reference Library for information about each outlet's location, time of opening, and if applicable, time of closing. There, I am able to find series of phone books, from 1970 to 2005 for virtually all 33 of the CMAs in Canada. Searches based on CMAs are necessary as the library does have complete series for the smaller Census Areas (CA's). Note that the CMAs of Sherbrooke, Saguenay and Trois-Rivieres are left out because of missing phone directories over certain time intervals. This method allows me to identify:

1. Opening year: The first year in which a particular outlet is listed in the phone directory.
2. Closing year: The last year in which a particular outlet is listed in the phone directory.
3. Location: The exact address of each outlet.

Outlets that first appear in the 1970 phone books may have opened in earlier years. To investigate whether this cut-off is appropriate, I look at the older phone directories (1950-1970) for some cities. With the exception of a few A & W and Harvey's outlets, very few in my sample actually opened before 1970. Each address is later geocoded and assigned a 6-digit postal code using Geocoder.ca. For each relevant FSA, I identify whether or not a chain is active in a particular FSA; a chain is defined to be active if it has at least one active store in the market.

Figure 1 highlights the amount of variation in both entry and exit over time. Furthermore, there is quite a lot of variation in the sequence of entry/exit decisions, as indicated in Table 3. In general, the fast food industry is quite dynamic.

Table 4 shows that each FSA can contain upwards of 9 outlets for a given chain. However, the fast food chains typically operate either 0 or 1 outlet in each market. Fewer than 5% of my

Figure 1: Total number of outlets opened/closed in Canada over time.

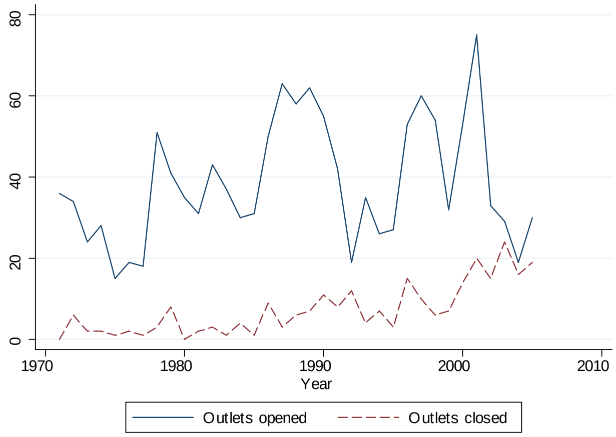


Table 4: Tabulation of market-time observations that contain 0, 1, ..., 9 outlets belonging to each of the chains.

	A & W	Burger King	Harvey's	McDonald's	Wendy's
0	18,018	19,182	19,070	12,192	19,539
1	3,126	2,505	2,536	7,027	2,174
2	508	188	228	1,891	142
3	160	13	46	536	28
4	67	0	6	142	5
5	9	0	2	55	0
6	0	0	0	28	0
7	0	0	0	9	0
8	0	0	0	5	0
9	0	0	0	3	0

market-time observations have a chain operating more than 1 outlet. Note that eventually, all FSAs contain at least one active chain by the end of my sample.

Also, the chains in general differ in terms of their entry timing (Table 5). We see that A & W and McDonald's typically enter first. Burger King, Harvey's, and Wendy's are more often than not followers into markets. In general, there is a lot of variation in terms of the timing of their entry (Figure 2). Furthermore, we get variation in the timing of exit for the retailers, as highlighted in Figure 3; the timing of exit appears to be spread out quite well, suggesting no deterministic patterns in exit due to franchisee contract renegotiations.

There are a handful of markets that were already occupied at the beginning of my sample in 1970. To see whether these markets are inherently different from markets that were occupied after

Figure 2: Histogram of entry years.

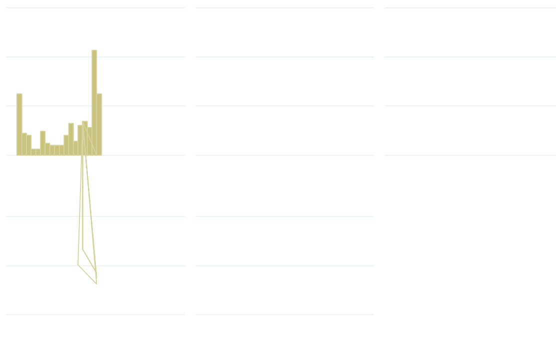


Figure 3: Histogram of the exit years.

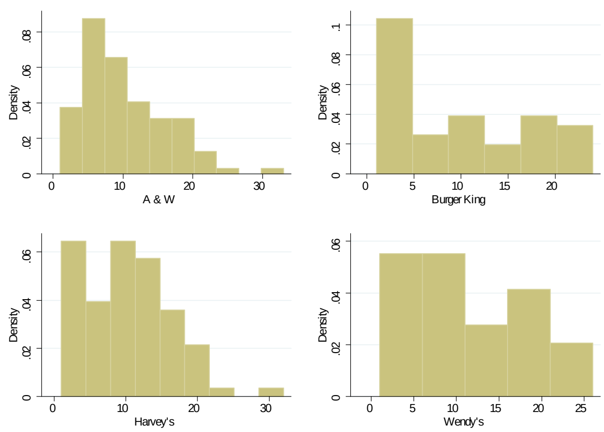


Table 5: Tabulation of the total number of markets that a chain was the (unique) …rst entrant.

Chain	First entrant

1970, I calculate the mean and variance for the main variables for two sub-samples. The first sub-sample is for markets that were occupied in 1970, and the second sub-sample is for markets that were occupied after 1970. Table 6 presents the summary statistics, and in general, there are no obvious differences between these two sub-samples. It is worth noting that the markets that were first occupied in 1970 do not appear to be systematically better than markets that were explored later on.

Table 6: 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

Table 8: Evidence of clustering based on the chains' decision to enter a market.

shock a_{imt}, which one may interpret as some form of manager/franchisee ability. Finally, V_{m} is a

1. Learning through entry: Within a year of entering a market, a retailer becomes informed and resolves its uncertainty about the size of the market. Therefore, $<_{imt} = 0$ if the retailer entered at time 1. Furthermore, the retailer does not forget, so that $<_{imt\ s} = 0$ for all 0 if $<_{imt} = 0$.

2. 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. The way in which the potential entrant updates the beliefs, $<_{imt}$, is described in detail below.

To set up the process by which an uninformed potential entrant can learn from its peers, I ...rst de...ne the set of retailers that made informed decisions at time 1:

$$ {}^{*}_{mt} = f \quad : \quad {}_{kmt-} = 1g \tag{5} $$

Note that each ...rm knows that every member in the set ${}^{*}_{mt}$ no longer faces uncertainty at period 1. The vector of decisions among those that belong in the set of informed retailers ${}^{*}_{mt}$ at time 1 is given by:

$$ {}^{*}_{mt-} = [\ {}_{jmt-} \ : \ 2 \ {}^{*}_{mt}] \tag{6} $$

With this notation in place and using Baye's rule, a potential entrant can then update its beliefs $<_{imt-}$ using the following recursive equation:

$$ <_{imt} = \frac{\Pr({}^{*}_{mt-} \ j \ {}_{m} \neq 0)<_{imt-}}{\Pr({}^{*}_{mt-} \ j \ {}_{m} \neq 0)<_{imt-} + \Pr({}^{*}_{mt-} \ j \ {}_{m} = 0)(1 \quad <_{imt-})} \tag{7} $$

Given the assumption of independent private information shocks, the conditional probability $\Pr({}^{*}_{mt-} \ j \ )$ is then de...ned as

$$ \Pr({}^{*}_{mt-} \ j \ ) = \prod_{j \in J_{mt}} {}_{jm}(\ )^{a_{jmt\ 1}} \ (1 \quad {}_{jm}(\ ))^{-a_{jmt\ 1}} \tag{8} $$

where ${}_{jm}(\ ) = \Pr({}_{jmt} = 1j \ )$. The probability $\Pr({}^{*}_{mt-} \ j \ )$ captures the information content associated with observed ${}^{*}_{mt-}$, which is a vector of actions at period 1 of these ...rms that belong to the set ${}^{*}_{mt}$. With this learning process in place, it becomes clear what the components of the information set are:

$$ \Omega_{imt} = f \ {}_{mt-} \quad {}_{mt-} \quad {}_{mt-} \ g \tag{9} $$

4.4 Markov Perfect Equilibrium (MPE)

The vector of payo¤ relevant state variables for …rm is ($_{mt}$ a_{imt} ς). Here,

$$_{mt} = f \;_{mt-} \;\; _{mt-} \;\; _{mt-} \;\; _{mt} g \tag{10}$$

where $_{mt-}$ = f $_{imt-}$ g$_i$, $_{mt-}$ = f $_{imt-}$ g$_i$, $_{mt-}$ = f$<_{imt}$g$_i$ and $_{mt}$ are exogenous market characteristics. An assumption I make regarding the equilibrium is that the strategy functions, ff_i($_{mt}$ a_{imt} ς_m)g$_i$ depend on the state variables; hence, the equilibrium is Markov Perfect. Given this state, the equilibrium strategies can be written as

$$f_i(\;_{mt}\; a_{imt}\; \varsigma_m) = \arg \max_{a_{imt} \in \{\;;\;\}} [\;^{\%}_{imt} + \sigma \;\; ^{\%}_{i}(\;_{mt}\;\; a_{imt}\;\; \varsigma_m) \,j\; _{mt}\; a_{imt}\; \varsigma_m] \tag{11}$$

where $^{\%}_{i}$($_{mt}$ a_{imt} ς_m) is the continuation value de…ned as

$$^{\%}_{i}(\;_{mt}\; a_{imt}\; \varsigma_m) =$$

I do allow for unobserved heterogeneity by introducing a market fixed effect, ς_m. Most importantly, the introduction of dynamics aides in identification, as it provides an important exclusion restriction. For instance, a retail chain's incumbency status has a direct impact on its flow profits via the entry costs, but will only affect its rival through its best response probability $_i(\ _{mt}\ \varsigma_m)$.

However, the incumbency status only acts as an effective exclusion restriction if the chain is not active two periods earlier ($_{imt-}$ = 0), or if the rival no longer faces any uncertainty about the market size ($_{jmt-}$ = 1 or $<_{jmt-}$ = 0). Otherwise, its decision to stay/exit will have a direct impact on the rival's payoff via the learning mechanism. Consequently, the parameters related to learning are confounded with the strategic interaction parameters. To separate out the parameters related to learning ($<\ J_i$) from the strategic interaction effects, I need sufficient variation in $_{imt-}$ and $_{imt-}$, given the functional form of the learning process as defined in my model.

Furthermore, ($<\ J_i$) are also confounded with the market fixed effect ς_m. In order to separately identify the parameters associated with learning from unobserved heterogeneity, I take advantage of one important source of variation generated by firm re-entry into markets. For example, consider a market in which a retail chain entered, left, and then re-entered. In my data, there are about 40 (out of 608) markets for which we see such behavior. The first time this chain entered, it most likely faced uncertainty. However, the second time it enters, the chain no longer faces uncertainty. In both cases, ς_m is the same, but ($<_{imt}\ J_i$) enters through the payoff only in the first case. Furthermore, timing of its first entry helps identify the prior $<$, as less weight is placed on the prior if the chain had more opportunities to learn from the past decisions of others.

Related to the issue of unobserved heterogeneity, there is likely an initial conditions problem in estimating this model, as some markets already have incumbents in the first year. In such cases, there could be a selection problem. To address this concern, I follow Arciacono and Miller's (2011) suggestion of using the first period observations to estimate the prior probability j of being in the a good market, where this prior probability is initialized using a flexible probit model.

In my model, the retail chains condition their strategies on the state $_{mt}$, which only contains information about the actions of competitors in the last two periods ($_{mt-}$ $_{mt-}$). It would appear as though the retailers were only learning based on these lagged decisions; therefore, my specification for their beliefs may not capture the full extent of their learned knowledge. However, the recursive structure of their learning process suggests otherwise. Note that their beliefs can be represented as a recursive relation $<_{imt} = \ (<_{imt-}\ _{mt-}\ _{mt-})$. If one solves this recursive relation, then their current period beliefs can actually be represented as $<_{imt} = \ (<\ f\ _{mt-s} g_{s>}\ )$.

Therefore, the inclusion of $<_{imt-}$ as a state variable is a compact way of representing knowledge inferred from past decisions $f_{mt-s}g_{s>}$. In other words, $<_{imt-}$ is a sufficient statistic for $f_{mt-s}g_{s>}$.

5.2 A simple DID specification test for learning

Using the framework set forth by my model, I can show the existence of a simple DID that can be adopted as an empirical test for the presence of learning using only the raw data patterns. This test will ultimately inform us as to the appropriateness of including a learning process in the dynamic entry/exit model. It however, will not tell us whether learning is actually causing

Note that we can also write the bes ... ity as

$$_i(\; _{imt-} \quad _{jm} \quad _i(\; ^P_i(\; _{imt-}$$
$$^P_i(\; _{imt-} \quad _{jm}$$

Therefore, the DID test under ... pothesis can also b ... d as:

$$\omega = [\; _i(0 \; 1 \; \varsigma \quad \varsigma_m)] \quad [\; _i(1 \; 1 \; \varsigma_m \quad _m)]$$

The two expressions for the D ... er the null hypoth ... ch other if an

$\omega = 0$.

5.2.1 Is learning present in ... od industry?

I illustrate this test by calculating ... for each chain-to- ... tion based o ... regression; to account for the mar ... also condition o ... market cha ... and market …xed e¤ects. Label ^P_i ... $_i$, $^P_i(0 \; 0 \; \varsigma_m$ ... $(1 \; 1 \; \varsigma_m)$ ... $^P_i(1 \; 0 \; \varsigma_m) = \; _i$. These objects ca ... ed via the foll ... sion:

$$(\; _{imt}| \; _{mt-} \quad _{mt}) = (1$$
$$+ \; _{imt}$$
$$_{imt}| \; \gamma$$

This linear regression is consistent ... my assumption that ... null ... 110.9 ... 1823.9 ... [(...)]

Table 9: DID test for learning using ...xed e¤ects linear regression.

	A & W	Burger King	Harvey's	McDonald's	Wendy's
A & W		-0.001 (0.01)	0.04 (0.03)	-0.03 (0.02)	0.02 (0.01)
Burger King	-0.007 (0.02)		0.005 (0.02)	-0.02 (0.03)	-0.01 (0.02)
Harvey's	-0.03 (0.02)	0.02 (0.02)		-0.03 (0.02)	-0.02 (0.02)
McDonald's	N/A	N/A	N/A		N/A
Wendy's	-0.04 (0.03)	0.01 (0.04)	0.01 (0.03)	0.003 (0.008)	

I do not have enough variation to identify the DID e¤ect McDonald's has on the other retailers. In general, there is also some heterogeneity in the DID across di¤erent retailers. This ...nding suggests that each retailer faces varying levels of ex ante uncertainty, as captured by my structural model. Ultimately, this model speci...cation test provides reduced form evidence in favor of the presence of learning, and justi...es the inclusion of uncertainty and learning in the structural model I estimate.

5.3 Estimation strategy

The parameters in my model are $= f \quad _i \quad _i \varphi_i \varphi_{ij} \quad J_i g_{\forall i}, < G$, and j . Therefore, conditional on $_{mt}$, and $\Upsilon = f \; < G j \; g$, the best response probability function $_i( )$ is used to construct the pseudo-likelihood equation. To estimate the speci...cation that incorporates a mixture distribution, I embed Arcidiacono and Miller's (2011) iterative Expectation-Maximization (EM) method with Aguirregabiria and Mira's (2007) Nested Pseudo Likelihood (NPL) procedure. A few additional steps are needed, as I outline in the Appendix. For notational simplicity, I use a subscript ς to indicate the CCP associated with the unobserved state $\varsigma_m = \varsigma$. The criterion for optimization is:

$$(\Upsilon \quad) = \sum_{i;m;t;} j \quad [\; _i(\; _{-i} (\; _{mt}) \quad _{mt} j \Upsilon)] \tag{26}$$

$$[\; _i(\; _{-i} (\; _{mt}) \quad _{mt} j \Upsilon)] \; = \quad _{imt}$$

for the initial CCPs for consistency, while at the same time, being tractable. Moreover, the NPL estimates are more efficient than alternative two-step methods.[10]

Multiple equilibria would be a particular concern if I instead adopted a nested fixed point algorithm to estimate the game, as doing so would require explicitly solving the model for each maximum likelihood iteration. When using the NPL, the main concern are multiple NPL fixed points. One way to test whether the pseudo-likelihood yields multiple NPL fixed points is to initialize the NPL at randomly drawn first-stage CCPs. If the NPL fixed point and estimated parameters are the same for each initialization, then multiple NPL fixed points are unlikely to be an issue.

6 Main results

6.1 Summary of estimates

My structural estimates are summarized in Table 10. There is some heterogeneity in terms of each chain's cost structure. It is noteworthy is that McDonald's enjoys the highest brand value, as reflected in φ_{MCD}. McDonald's high brand value in Canada should not be surprising, as it has

Table 10: Structural estimation of dynamic entry/exit model.

	A & W	Burger King	Harvey's	McDonald's	Wendy's
Brand value (φ_{ii})	0.08 (0.01)	0.01 (0.01)	0.01 (0.01)	0.03 (0.01)	0.02 (0.01)
vs A & W (φ_{iAW})		0.05 (0.03)	0.1 (0.04)	-0.2 (0.03)	0.1 (0.03)
vs Burger King (φ_{iBK})	-0.2 (0.04)		-0.04 (0.04)	-0.4 (0.03)	-0.05 (0.04)
vs Harvey's (φ_{iHARV})	0.2 (0.03)	0.04 (0.03)		-0.2 (0.03)	0.09 (0.04)
vs McDonald's (φ_{iMCD})	-0.3 (0.04)	-0.003 (0.04)	-0.09 (0.04)		-0.01 (0.03)
vs Wendy's (φ_{iWEND})	0.05 (0.03)	0.03 (0.03)	0.07 (0.03)	0.07 (0.03)	
Fixed costs ($_i$)	-0.04 (0.04)	0.07 (0.04)	0.004 (0.05)	-0.4 (0.05)	0.07 (0.04)
Entry costs ($_i$)	0.1 (0.01)	0.03 (0.02)	0.08 (0.01)	0.02 (0.008)	-0.04 (0.01)
Degree of uncertainty (J_i)	-0.2 (0.02)	0.03 (0.02)	-0.2 (0.02)	-0.3 (0.01)	-0.01 (0.02)
Prob. of uncertainty ($<$)	0.20 (0.01)				
Good state parameter (G)	0.98 (0.14)				
Prob. of good state (j)	0.45				

has a negative sign. What this means is that the retailers tend to bias downwards their beliefs about pro…tability in markets that are inherently good, and bias upwards their beliefs in those that are inherently bad. Also notice that the degree of uncertainty is di¤erent across the retailers; in particular, we see that J_i is largest for A & W, Harvey's, and McDonald's, these retailers appear to be the most sensitive to uncertainty.

6.2 Can learning induce clustering?

The estimated structural model provides us an opportunity to look explicitly at the role of uncertainty in retail agglomeration. To investigate the impact of uncertainty on market outcomes, I compare the entry/exit decisions when uncertainty is present to when uncertainty is not present. One may interpret a counterfactual reduction of uncertainty as the hypothetical event where the Canadian government releases to the public its (initially) con…dential detailed data on restaurant sales (by category) from tax returns, or detailed information about market characteristics such as tra¢ c lights. Such a policy is realistic, as many municipalities in Canada have adopted an open data initiative.

The objective of this analysis is to establish a link between uncertainty and retail clustering. Since uncertainty and learning are closely intertwined, such a link implies a connection between clustering and learning. Empirical analysis in the earlier sections has already shown us that the entry/exit patterns we see in the data are consistent with the story of learning; but such analysis does not actually show how uncertainty/learning will impact retail concentration, as uncertainty

Table 11: 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

into markets so as to avoid being the first entrants into a market, whereby being first yields no informational spillover that they can get a free-ride off of (Chamley, 2004). Strategic delay would ultimately generate the pattern in Figure 4 where the herding behavior is more pronounced in the latter years, as the option value of delay falls.

Figure 5: The number of instances in which a retailer avoids a market others failed in.

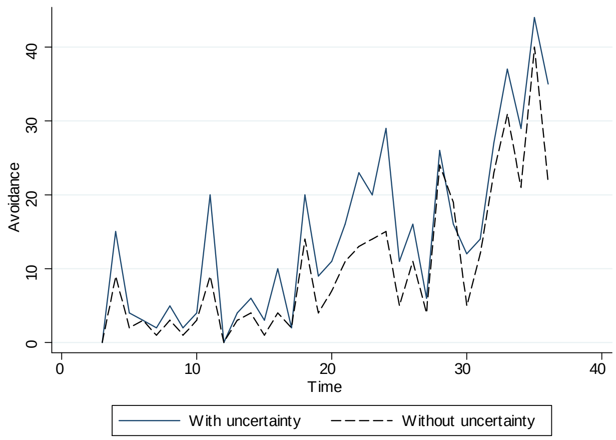
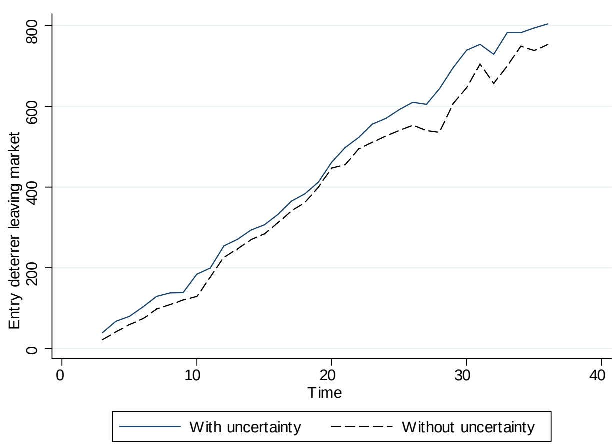


Figure 6: Number of instances in which entry deterrer exits a market.



7 Concluding remarks

The primary objective of my paper has been to understand an overlooked force behind clustering of retail chains, a challenging problem that current research in industrial organization, marketing, and urban economics has not yet explored. Using unique data with rich time and geographic variation from Canada's fast food industry, I develop a new oligopoly model of entry and exit that accounts for learning and unobserved heterogeneity. Using the model, I derive a simple DID test for learning, that when applied to my data, shows that learning is present in the fast food industry. Through counterfactual analysis of an estimated model, I show that an industry facing uncertainty and learning is more clustered than an industry facing no uncertainty and learning, thereby showing a connection between learning and agglomeration.

In future work, researchers may wish to consider that firms can potentially learn about profitability through their own experience in similar or neighboring markets. For example, a retail chain may learn through its past experience that low income markets are better than high income markets for generating demand if low income households have a greater propensity to consume unhealthy and salty food. Such experiences should then induce the chain to focus primarily on these markets in the future. My analysis has abstracted away from such learning behavior. However, it may be worthwhile considering this extension for future work as doing so can introduce rich heterogeneity in the ex ante beliefs that can ultimately be identified by data, when information about realized revenue is not available. With such a model, one can determine which types of markets are riskier than others. Such insight would especially be useful if retail managers have limited resources for conducting real estate research across markets, and wish to allocate their local headquarters optimally.

Finally, the DID (regression) test for learning I present need not be restricted to the fast food industry. It could in principle be applied to a more general class of social interaction models. For example, this test could be used to determine whether learning from peers is present in the adoption of new technologies, or in the consumption of new experience goods (i.e., word-of-mouth). The general strategy for this strand of empirical research is to first identify credibly a peer effect, and then run a series of ad hoc falsification tests that suggest that these peer effects are most likely driven by learning.

References

[1] Aguirregabiria, V. and Mira, P. (2007). Sequential Estimation of Dynamic Discrete Games. Econometrica 75.

[2] Aguirregabiria, V. and Mira, P. (2010). Dynamic Discrete Choice Models: A Survey. Journal of Econometrics 156, 38-67.

[14] Chi, F. and Yang, N. (2010). Twitter Adoption in Congress. Review of Network Economics 10.

[15] Conley, T. and Udry, C. (2010). Learning about a New Technology: Pineapple in Ghana. American Economic Review 100, 35-69.

[16] Damar, H-E. (2009). Why Do Payday Lenders Enter Local Markets? Evidence from Oregon. Review of Industrial Organization 34, 173-191.

[17] Datta, S. and Sudhir, K. (2011). The Agglomeration-Di¤erentiation Tradeo¤ in Spatial Location Choice. Working paper.

[18] Deutsch, C. (1993). Commercial Property: Ladies' Mile; From the Carriage Trade to the Category Killers. New York Times, June 23, 1993.

[19] Dubé, J-P., Hitsch, G., and Rossi, P. (2010). State Dependence and Alternative Explanations for Consumer Inertia. RAND Journal of Economics forthcoming.

[20] Dunne, T., Klimek, S., Roberts, M. and Xu, Y. (2009). Entry, Exit, and the Determinants of Market Structure. NBER Working paper.

[21] Eckert, A. and West, D. (2008). Firm Survival and Chain Growth in a Privatized Retail Liquor Store Industry. Review of Industrial Organization 32.

[22] Ellickson, P. Houghton, S. and Timmins, C. (2010). Estimating Network Economies in Retail Chains: A Revealed Preference Approach. NBER Working paper.

[23] Ellison, G., Glaeser, E. and Kerr, W. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. American Economic Review 100, 1195-1213.

[24] Eppli, M. and Benjamin, J. (1994). The Evolution of Shopping Center Research: A Review and Analysis. Journal of Real Estate Research 9.

[25] Feinberg, R. (2008). Explaining the Credit Union Entry Decision, and Implications for Performance. Review of Industrial Organization 33, 81-91.

[26] Goolsbee, A. and Klenow, P. (2002). Evidence on Learning and Network Externalities in the Di¤usion of Home Computers. Journal of Law and Economics 45, 317-43.

[27] Holmes, E. (2010). U.S. Apparel Retailers Map an Expansion to the North. Wall Street Journal March 29, 2010.

[28] Holmes, T. (2011). The Di¤usion of Wal-Mart and the Economies of Density. Econometrica 79, 253-302.

[29] Jia, P. (2008). What happens when Wal-Mart comes to town: an empirical analysis of the discount retailing industry. Econometrica 76.

[30] Jovanovic, B. (1982). Selection and the Evolution of Industry. Econometrica 50.

[31] Jovanovic, B. (1989). Observable Implications of Models with Multiple Equilibria. Econometrica 57, 1431-1437.

[43] Rob, R. (1991). Learning and Capacity Expansion under Demand Uncertainty. The Review of Economic Studies 58, 655-75.

[44] Schlosser, E. (2001). Fast Food Nation. HarperCollins: New York.

[45] Seim, K. (2006). An empirical model of …rm entry with endogenous product-type choices. RAND Journal of Economics 37.

[46] Shen, Q. (2010). A Dynamic Model of Entry and Exit in a Growing Industry. Working paper.

[47] Shen, Q. and Liu, H. (2012). Dynamic Learning and Exit in Competitive Markets. Working paper.

[48] Shen, Q. and Xiao, P. (2011). McDonald's and KFC in China: Competitors or Companions. Working paper.

[49] Smith, A. (2006). Encyclopedia of Junk Food and Fast Food. Greenwood Press: Westport.

[50] Strauss, M. (2011). U.S. retail invasion gathers pace. The Globe and Mail, January 21, 2011.

[51] Suzuki, J. (2010). Land Use Regulation as a Barrier to Entry: Evidence from the Texas Lodging Industry. Working paper.

[52] Tamer, E. (2003). Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. Review of Economic Studies 70, 147-167.

[53] Thomadsen, R. (2007). Product Positioning and Competition: The Role of Location in the Fast Food Industry. Marketing Science 26, 792-804.

[54] Thomadsen, R. (2010). Seeking an Aggressive Competitor: How Product Line Expansion Can Increase All Firms' Pro…ts. Working paper.

[55] Toivanen, O. and Waterson, M. (2005). Market structure and entry: where's the beef? RAND Journal of Economics 36.

[56] Toivanen, O. and Waterson, M. (2011). Retail Chain Expansion: The Early Years of McDonald's in Great Britain. Working paper.

[57] Tucker, C. and Zhang, J. (2010). Growing Two-Sided Networks by Advertising the User Base: A Field Experiment. Marketing Science 29, 805-814.

[58] Varela, M. (2010). The Costs of Growth: Estimating Entry Costs with Endogenous Growth Rates. Working paper.

[59] Vitorino, M-A. (2008). Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry. Journal of Marketing Research forthcoming.

[60] Yang, N. (2010). Burger King & McDonald's: Where's the Spillover? International Journal of the Economics of Business forthcoming.

[61] Zhang, J. (2010). The Sound of Silence: Observational Learning in the U.S. Kidney Market. Marketing Science 29, 315-355.

[62] Zhu, T. and Singh, V. (2009). Market Structure and Competition in the Retail Discount Industry. Journal of Marketing Research 46,453-466.

[63] Zhu, T. and Singh, V. (2009). Spatial competition with endogenous location choices: An application to discount retailing. Quantitative Marketing and Economics 7, 660-667.

[64] Zhu, T., Singh, V. and Dukes, A. (2011). Local competition, entry, and agglomeration. Quantitative Marketing and Economics forthcoming.

8 Appendix

8.1 Details about how variables are imputed

I impute the population in 1999 using the inferred exponential population growth rate between 1996 and 2001, and the population in 1990 using the exponential growth rate between 1991 and 1996. Observations before 1986 are imputed using a convex combination of the national growth rate and the growth rate pertaining to 1986 to 1991. I place a greatest weight on the annual national growth rate for years closest to 1970, and greatest weight on the 1986-1991 growth rate for years approaching 1986. I am also able to obtain the geographic area (in sq km) for each FSA from the Census of Canada. These values are later used to calculate the population density for each FSA market.

I impute income and property value in a similar manner as population. The di¤erence is that for the years before 1986, I use a convex combination of the national in‡ation rate and the rate of return pertaining to 1986 to 1991. Because the proportion of residents who work in/out of an FSA market was not available for each Census, I use the information available for 2006.

8.2 Applying Aguirregabiria and Mira's (2007) representation lemma

I will now demonstrate how the MPE can be expressed using only the conditional choice probabilities, states, and model primitives. As before, $_{mt}$ denotes the state. We can express the speci…c values associated with being active and not as the following:

$$_i(1 \quad _{mt}\ \varsigma) \qquad {}^P_i(\quad _{mt}\ \varsigma) + \sigma \ {}_i^{X;P}(1 \quad _{mt}\ \varsigma)' \, {}^{-P}_i$$

$$_i(0 \quad _{mt}\ \varsigma) \qquad \sigma \ {}_i^{X;P}(0 \quad _{mt}\ \varsigma)' \, {}^{-P}_i$$

where ${}_i^{X;P}(1 \quad _{mt}\ \varsigma)$ and ${}_i^{X;P}(0 \quad _{mt}\ \varsigma)$ are transition probability vectors, and ${}^{-P}_i$ is a vector of integrated values across all possible states. Because the decision variable is discrete, we can write the integrated value as

$$
\begin{aligned}
{}_i^P(\quad _{mt}\ \varsigma) \qquad & \quad _i(\quad _{mt}\ \varsigma) \ _i(1 \quad _{mt}\ \varsigma) + (1 \quad _i(\quad _{mt}\ \varsigma)) \ _i(0 \quad _{mt}\ \varsigma) + {}^P_{imt} \\
= & \quad _i(\quad _{mt}\ \varsigma)[\ {}^P_i(\quad _{mt}\ \varsigma) + \sigma \ {}_i^{X;P}(1 \quad _{mt}\ \varsigma)' \, {}^{-P}_i] \\
& + (1 \quad _i(\quad _{mt}\ \varsigma))[\sigma \ {}_i^{X;P}(0 \quad _{mt}\ \varsigma)' \, {}^{-P}_i]
\end{aligned}
$$

where ${}^P_{imt} = P(\ ^- \ (\ _i(\quad _t\ \varsigma \ ,))$, and ${}^P_{imt}$ is derived using the assumption that a_{imt} has an iid

8.3 Details about the estimation procedure

The estimation algorithm can be described as follows:

1. Generate a grid of possible values for $<^g$ 2 [0 1].

2. Estimate non-parametrically the initial CCP vector $\hat{\ }^{;g}$. Alternatively, draw them randomly from a uniform distribution.

3. As in Arcidiacono and Miller (2011), initialize j at the predicted probability from a ...tted probit model of entry using the ...rst year's worth of data.

4. Given ${}_{mt}$, $\hat{\ }^{;g}$, and $<^g$, generate a sequence of posterior beliefs for each ...rm and market $f\hat{<}^{;g}_{mt} g_{\forall ;m;t}$.

5. Given ${}_{mt}$, $\hat{\ }^{;g}$, $<^g$, and $f\hat{<}^{;g}_{mt} g_{\forall ;m;t}$, compute:

$$
{}_m^{;g} = \frac{j \prod_t [\ {}_i(\ -{}_i(\ {}_{mt}\ \varsigma_G)\ \ {}_{mt} j\ )]}{\sum_{j'} j' \prod_t [\ {}_i(\ -{}_i(\ {}_{mt}\ \varsigma')\ \ {}_{mt} j\ )]} \tag{28}
$$

6. Use ${}_m^{;g}$ to calculate j according to:

$$
j^{;g} = \frac{\sum_m {}_m^{;g}}{\ } \tag{29}
$$