International Journal of Research in Marketing 37 (2020) 336–355

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Learning in retail entry☆

Nathan Yang*

Desautels Faculty of Management, McGill University, Canada
Université de Montréal, Pavillon Lionel-Groulx, Centre interuniversitaire de recherche en économie quantitative, CP 6128, succursale Centre-ville, Montréal, QC H3C 3J7, Canada
CIRANO, 1130 Rue Sherbrooke Ouest #1400, Montréal, QC H3A 2M8, Canada
GERAD, André-Aisenstadt Building, 2920, Chemin de la Tour, 4th floor, room 4465, Montreal, Quebec H3T 1J4, Canada
IVADO, 6650 St Urbain St suite 350, Montreal, Quebec H2S 3G9, Canada

ARTICLE INFO

Article history:
First received on January 25, 2019 and was under review for 8 months
Available online 21 November 2019

Area Editor: Tülin Erdem

Keywords:
Bayesian learning
Dynamic discrete choice
Location intelligence
Market structure
Retail strategy
Social learning
Unobserved heterogeneity

ABSTRACT

Retailers may face uncertainty about the profitability of local markets, which provide opportunities for learning when making entry decisions. To quantify these informational benefits, I develop an empirical framework for studying dynamic retail entry with uncertainty and learning (from others). Using novel data about fast food chains, I estimate the model with a forward simulation estimation approach augmented with particle filtering as a way to flexibly account for unobserved firm beliefs about market profitability. The estimates confirm the presence of uncertainty and learning. Most importantly, simulations using the estimated model demonstrate that learning from others may indeed help mitigate some of the uncertainty.

☆ This research was previously circulated as “March of the Chains: Herding in Restaurant Locations” and is a substantially revised version of the second chapter from my thesis done while at the University of Toronto. I am indebted to Victor Aguirregabiria, Avi Goldfarb, Lu Han, and Junichi Suzuki for their patient advice and guidance. I also thank Jason Blevins, Ron Borkovsky, Andrew Ching, Victor Couture, Rahul Deb, April Franco, Matthew Grennan, Günter Hitsch, Masakazu Ishihara, Panle Jia, Vineet Kumar, Mara Lederman, Joshua Lewis, Hongju Liu, Michael Luca, Robert McMillan, Pedro Mira, Jeffrey Prince, Nathan Schiff, Qiaowei Shen, William Strange, K. Sudhir, Otto Toivanen, Mauricio Varela, Maria Ana Vitorino, Kitty Wang, Mo Xiao, and Juanjuan Zhang for helpful discussions. Also, much gratitude to seminar and conference participants at the Bank of Canada, Charles River Associates, University of Chicago Booth, CMU Tepper, Purdue Krannert, Econometric Society Meeting, FTC Microeconomics Conference, Johns Hopkins University, Kansas State University, MIT Sloan, QME Annual Meeting, SUNY-Stony Brook, University of Connecticut School of Business, University of Toronto, Urban Economics Association Annual Meeting, and Yale School of Management for helpful comments. Finally, I gratefully acknowledge funding from Social Sciences and Humanities Research Council of Canada and the NET Institute grant #11-16 (www.netinst.org).
* Desautels Faculty of Management, McGill University, Bronfman Building 1001 rue Sherbrooke Ouest, Montreal, Quebec Canada H3A 1G5.
E-mail address: nathan.yang3@mcgill.ca.

https://doi.org/10.1016/j.ijresmar.2019.09.005
0167-8116/

1. Introduction

For brick-and-mortar retailers, entering new geographic markets is capital intensive and often entail uncertainty.[1] Market-specific risks are especially challenging to evaluate, as these risks would have to be assessed at every geographic market, unlike firm-level, category-level, or product-level risk that can be resolved via scalable marketing intelligence research.

Location intelligence (i.e., data-driven site selection) for future retail outlet openings is often research-intensive and not as scalable, whereby real estate managers leverage whatever information they can collect on local traffic, demographics, safety information, commercial mix, and other factors for each prospective outlet location.[2] Presence of customers is critical for retailers, and thus, identifying this presence remains a priority. Ultimately, being able to assess the market is vital to the strategic growth plan as where a retailer places its store will likely determine whether it succeeds or fails.[3] As an anecdote that speaks to the efforts retailers place towards location intelligence, interviews with real estate managers reveal that a popular method for assessing the amount of foot traffic at a prospective location has been to send an employee to that site and do a time-consuming count of the number of passerby.[4] One potential location intelligence "short-cut" that has been proposed in this industry is to make inferences based on how well other retailers in the market have done in terms of entry and survival (i.e., learning from others).[5] The value of learning from others towards market-specific risk assessment is that the entry decisions of a competitor may potentially serve as a credible signal of whether or not a market is indeed profitable enough to open a new store, as the competitors likely invested some time towards researching the locations they eventually enter. Furthermore, an outlet's ability to survive upon entry may be indicative of favorable market conditions. Therefore, my research is focused on determining empirically whether or not these past entry decisions by competitors can indeed help reduce the market-specific uncertainty faced by brick-and-mortar retailers. Understanding the role that this additional information plays will be critical to provide improved location intelligence for retailers.

To assess the role of learning from others in entry decisions, I choose the Canadian fast food industry from its beginning in 1970 to 2005, as the empirical setting. This setting is chosen for a few reasons. First, the informational gains from learning may be especially pronounced in this retail sector as the firms typically do not typically have access to loyalty programs for granular customer purchase information in order to forecast market viability. Second, this industry relies on its brick-and-mortar locations to grow their sales, thereby making entry and exit a critically important decision with respect to distribution; in fact, the fast food industry was among the pioneers of location intelligence.[6] Finally, Canada was the first international market for almost all of the American fast food retail chains, so it seems entirely plausible that the retailers faced uncertainty about these new markets (e.g., Beamish, Jung, & Kim, 2003; Bell & Shelman, 2011; Schragger, 2001) especially given the novelty of the fast food restaurant concept.

However, empirically quantifying the impact of learning is challenging because to quantify learning effects, one needs to make inferences about the beliefs firms have about the geographic markets (i.e., whether the markets are expected to be profitable or not). These beliefs are difficult, if not impossible, to observe directly as data, so they are ultimately treated as unobservable (e.g., Ching, Erdem, & Keane, 2013). On top of their unobservable nature, beliefs are evolving over time as retailers absorb more information about the markets. Furthermore, these beliefs interact with each retailer's expectation of what they anticipate their rivals will do today, tomorrow, and onwards. Consequently, researchers have at best been able to quantify positive externalities between seemingly rival retailers (e.g., Shen & Xiao, 2014; Toivanen & Waterson, 2005; Yang, 2012), falling short of narrowing down these effects further to assess the role of learning.

In light of these challenges, my main contribution is to develop a new estimable model of entry that provides the following features. As in typical models of dynamic retail entry (e.g., Arcidiacono, Bayer, Blevins, & Ellickson, 2016; Beresteanu, Ellickson, & Misra, 2010; Hollenbeck, 2017; Suzuki, 2013), my framework allows firms to be both forward looking and strategic with respect to their entry decisions. In particular, the most significant contribution of my model is that I allow firms to potentially face uncertainty about whether or not a specific geographic market is viable. Under the presence of uncertainty, firms have an opportunity to learn, whereby observing entry/survival of rivals has the potential to reduce the uncertainty they face. The extent to which learning can actually reduce this uncertainty is an empirical question, and is one of the key effects that my research aims to uncover. Despite the ample theoretical treatment of decisions (involving sunk costs) with learning (e.g., Bikhchandani, Hirshleifer, & Welch, 1998; Caplin & Leahy, 1998; Chamley, 2004; Chamley & Gale, 1994; Gale, 1996), there exists no empirical counterpart.

Estimating this model is not trivial either, given the complexity in strategic and learning dynamics. In light of these complexities, the estimation approach needs to (at a minimum) accommodate for unobserved states (i.e., beliefs about market viability)

[1] See, "Tim Hortons Abruptly Shuts Multiple U.S. Locations, Taking Customers, Staff by Surprise" (*Huffington Post Canada*, November 20, 2015).

[2] See "How Fast Food Chains Pick Their Next Location" (*Fast Company*, August 25, 2014) for more details.

[3] See, "How Big Data Helps Chains Like Starbucks Pick Store Locations – An (Unsung) Key To Retail Success" (*Forbes*, April 24, 2014).

[4] See the Appendix for more details.

[5] See, "10 Things to Consider When Choosing a Location for Your Business" (*Entrepreneur*, May 20, 2015), where they state: "Ask about previous tenants. If you're opening a restaurant where five restaurants have failed, you may be starting off with an insurmountable handicap - either because there's something wrong with the location or because the public will assume your business will go the way of the previous tenants."

[6] As described in Love (1995), Ray Kroc (McDonald's co-founder) was known to investigate thoroughly every prospective market using any information (e.g., population density, income per household) he could obtain from city almanacs. It is through these efforts that McDonald's eventually created a separate real estate department whose sole purpose is to identify viable locations for their outlets.

that may evolve both stochastically and endogenously. For this reason, I augment the Bajari, Benkard, and Levin (2007) forward simulation approach with particle filtering, as first proposed by Blevins, Khwaja, and Yang (2018) under the context of dynamic retail expansion. The use of particle filtering allows me to obtain posterior distributions for the firm beliefs about market viability, which is not possible to obtain using traditional approaches for estimating dynamic games. Particle filtering is particularly powerful in the context of learning, as firms posterior beliefs are likely continuous, time-varying, and endogenous unobserved states. A distinct advantage of particle filtering (over alternative methods) is that it helps preserve all of these features about the beliefs.

The estimated model provides a number of managerially relevant insights. First, my structural estimates show that all of the retail chains likely face uncertainty (i.e., non-negligible estimates for prior variance). Second, I am able to confirm that the uncertainty retail chains face over time (i.e., posterior variance) is diminishing. Third, I demonstrate via simulations of the estimated model that the presence of rival incumbents can induce a firm to update its beliefs about the markets viability. In particular, a hypothetical one-time increase in the initial number of rival incumbents has the effect of increasing a firm's assessment of the markets viability (i.e., posterior mean), while at the same time, temporarily decreasing the firms uncertainty (i.e., posterior variance) about the market's viability. Taken together, this research shows that uncertainty matters, and that learning from others provides a practical uncertainty reduction strategy in location intelligence.

2. Model

This section provides a description of the dynamic retail entry model that will be estimated. The model accommodates for forward looking firms, as well as serial correlation that evolves both exogenously (i.e., time-varying unobserved heterogeneity) and endogenously (i.e., perceived beliefs about market profitability). I first describe the model primitives, followed by a discussion of the learning process that contributes to the endogenous serial correlation.

2.1. Background

Consider a retail industry with J firms, indexed by $i = 1, 2, \ldots, I$. At the beginning of each time period t, the firms have two possible actions to choose from $a_{imt} \in \{0, 1\}$; namely, they decide simultaneously whether to be active (i.e., $a_{imt} = 1$), or not (i.e., $a_{imt} = 0$) in local market $m = 1, 2, \ldots, M$. Therefore, the market structure can be captured by $a_{mt} = \{a_{imt}\}_{\forall i}$.

Firms are forward-looking with a discount rate of $\beta \in (0, 1)$, so their objective is to maximize the discounted payoffs $\sum_s^\infty \beta^{t+s}\Pi_{imt+s}$, where Π_{imt+s} is the one-shot payoff of firm i at period $t + s$. The one-shot payoff from being active in a market can be written as follows:

$$\Pi_i(a_{mt}, x_{imt}, z_{mt}, q_m; \theta) = \theta_1 x_{imt} + \theta_2 \sum_{-i \neq i} a_{-imt} + \theta_3 + \theta_4(1 - a_{imt-1}) + z_{mt} + q_m - \varepsilon_{imt}. \tag{1}$$

Choosing not to enter yields a pay-off a zero. As in standard retail entry models, the one-shot payoff is affected by the observable market characteristics that may include firm-specific states (x_{imt}), competitive conditions $\left(\sum_{-i \neq i} a_{it}\right)$, fixed costs ($\theta_3$), and entry costs ($\theta_4$).[7] Like Seim (2006), I also assume that each chain receives a privately known and idiosyncratic shock ε_{imt}, which one may interpret as some form of managerial ability that is known only to the firm i.

Departing from the standard model, there are two additional contributors of market profitability, which I denote as z_{mt} and q_m. First, I allow there to be serially correlated market profitability (z_{mt}) that is known to the firms but not known to the econometrician. This serially correlated process evolves according to an autoregressive process as represented by the transition equation,

$$z_{mt} = \rho z_{mt-1} + \nu_{mt}, \tag{2}$$

where ρ captures the persistence of the serially correlated process, while ν_{mt} is i.i.d. white noise.[8] Second, the remaining component of profitability (q_m) is unknown to both the firm and the econometrician.

To reiterate, there are three main sources of information that affect firm decisions in my model. There is information that is known to both the firm and econometrician (x_{imt}), information that is known only to the firm (z_{mt}), and information that is unknown to both the firm and the econometrician (q_m). In the next section, I will describe the process and channels by which firms learn about q_m.

[7] Note that I normalize exit costs to be zero, as they cannot be separately identified from entry and fixed costs (Aguirregabiria & Suzuki, 2014).

[8] A similar autoregressive process for the time-varying unobserved heterogeneity has also been used in single-agent dynamic discrete choice models like Fang and Kung (2010).

2.2. Learning

Firms can make inferences about q_m through Bayesian learning. Before receiving any information via learning, I assume that for each market m, firm i has a normal prior about unobserved market profitability:

$$q_m \sim N\left(q^i_{m1}, \sigma^i_{m1}\right). \tag{3}$$

In other words, before receiving any information, firms perceive the true value of the unobserved component of profitability to be distributed normally with mean q^i_{m1} and variance σ^i_{m1}. So in the first period, the firm's information set is $Q^i_1 = \left\{q^i_{m1}, \sigma^i_{m1}\right\}_{\forall m}$.

Learning takes place at the beginning of each time period t. There are two ways in which firms can learn. First, there is learning from experience, in which their own past experience in the market (i.e., a_{imt-1}) is informative about that market's profitability. Second, there is learning from others, in which the observed past experiences of others in the market is informative (i.e., $\check{a}_{-imt-1}$). Here, I define $\check{a}_{-imt-1} = 1$ if $a_{-imt-1} = 1$ for any $-i \neq i$. What this means is that observations of past entry behaviors $\check{a}_{-imt-1}$ generate noisy signals about the true value of unobserved profitability q_m. The observation of $a_{imt-1} = 1$ generates the following noisy signal:

$$\tilde{q}^i_{mt} = q_m + e^i_{mt}, \tag{4}$$

where $e^i_{mt} \sim N(0, \sigma^i_e)$. Here, the variability of signals generated by learning from experience is denoted by σ^i_e. Furthermore, because $\check{a}_{-imt-1}$ is perfectly observable in most retail industries, firms can engage in learning from others by receiving additional noisy signals in the form

$$\tilde{q}^{-i}_{mt} = q_m + f^i_{mt}, \tag{5}$$

where $f^i_{mt} \sim N(0, \sigma^i_f)$. In this case, variability of signals generated by learning from others is captured by σ^i_f.[9] Note that in both types of learning, these signals do not fully reveal the true value of q_m.

Before describing the Bayesian updating process for learning, I digress first and offer a brief discussion about plausible interpretations of the information content in the signals under the context of retail entry. In learning from experience, the signals $\tilde{q}^i_{mt}$ may contain information about past sales/profits, as well as market-specific idiosyncracies that managers or franchisees uncovered during their tenure. In learning from others, the signals $\tilde{q}^{-i}_{mt}$ may contain information about the observed number of customers who walk into rival stores, as well as inferences about profitability made from how long existing incumbents are able to survive in the market.

In this formulation of learning, the priors and signals are conjugates, as both the priors about unobserved market profitability as well as the signal noise are assumed to be normally distributed. This assumption allows me to derive formulas for Bayesian updating the firms' beliefs as new information arrives (DeGroot, 1970). Given these distributional assumptions, I can write the perceived market profitability (denoted q^i_{mt}) recursively at time t as follows:

$$q^i_{mt} = \sigma^i_{mt}\left[\left(1/\sigma^i_e\right)\tilde{q}^i_{mt-1}a_{imt-1} + \left(1/\sigma^i_f\right)\tilde{q}^{-i}_{mt-1}\check{a}_{-imt-1} + \left(1/\sigma^i_{mt-1}\right)q^i_{mt-1}\right] \tag{6}$$

$$\sigma^i_{mt} = \frac{1}{1/\sigma^i_{mt-1} + \left(1/\sigma^i_e\right)a_{imt-1} + \left(1/\sigma^i_f\right)\check{a}_{-imt-1}} \tag{7}$$

The level of uncertainty faced by firm i at time t is captured by $\left\{\sigma^i_{mt}\right\}_{\forall m}$, where this uncertainty is affected by the number of signals received. Given the posterior mean $\left(q^i_{mt}\right)$ and variance $\left(\sigma^i_{mt}\right)$, the information set in period t is $Q^i_t = \left\{q^i_{mt}, \sigma^i_{mt}\right\}_{\forall m}$. Note here that the posterior variance of perceived market profitability declines as more signals are received, and in the limit, perceived market profitability converges towards the true market profitability. Furthermore, the informativeness of signals associated with learning from experience and learning from others may be different (i.e., $\sigma^i_e \neq \sigma^i_f$).

2.3. Equilibrium

Based on the specification above, we can summarize the payoff relevant states as

$$S_{imt} = \left(x_{imt}, z_{mt}, a_{mt-1}, Q^i_t\right). \tag{8}$$

[9] These noisy signals may be interpreted as factors that drive competing chains to enter a market that are not perfectly correlated with the markets true quality. For example, competing chains may be opening new stores in a specific market because talented entrepreneurs in a region suddenly become available to serve as franchisees; in order to secure these franchisees, the chains may have an incentive to set up the store even if the current market conditions do not justify the stores opening.

The entire set of payoff relevant sets in market m at time t can be summarized as $S_{mt} = \{S_{imt}\}_{\forall i}$. Given these state variables, firms form strategies $\varrho_i : S_{imt} \rightarrow \{0,1\}$. Here, a_{imt} can be described as a mapping from the payoff relevant states (S_{imt}) to strategies ϱ_i. I assume that firms follow a Markov Perfect Equilibrium such that the strategy profile (e.g., Ericson & Pakes, 1995), $\varrho = \{\varrho_i\}_{\forall i}$, satisfies the following condition for all i:

$$V_i\left(S_{mt}, \varrho | \varrho_i^*, \varrho_{-i}^*\right) \geq V_i\left(S_{mt}, \varrho | \varrho_i, \varrho_{-i}^*\right) \tag{9}$$

for all ϱ_i and all states S_{mt}, where $V_i(\cdot)$ is a Bellman equation defined using the following recursive expression:

$$V_i(S_{mt}, \varrho) = E[\Pi_i(S_{imt}, \varrho(S_{imt})) + \beta E(V_i(S_{imt+1}, \varrho) | S_{imt}, \varrho_i(S_{imt}), \varrho_{-i}(S_{-imt})) | S_{mt}]. \tag{10}$$

In other words, the strategy profile in equilibrium is defined such that no firm i has an incentive to deviate from the optimal strategy ϱ_i^*, in that no alternative strategy (ϱ_i) yields higher expected discounted profits than ϱ_i^* while its rivals use strategies ϱ_{-i}^*. Note that in the Bellman equation, $\varrho_i(S_{imt}) = a_{imt}$ and $\varrho_{-i}(S_{-imt}) = a_{-imt}$ can be thought of as implied actions by firms i and $-i$, under the strategies ϱ_i and ϱ_{-i} when states are S_{imt} and S_{-imt}, respectively.

3. Estimation

Based on the model I have introduced in the previous section, the structural parameters that need to be estimated are

$$\Psi = \left(\theta, \rho, \left\{q_{m1}^i, \sigma_{m1}^i, \sigma_e^i, \sigma_f^i\right\}_{\forall i}\right). \tag{11}$$

The challenge of estimating these parameters is the presence of serially correlated and unobserved states. First, the firms make decisions based on the market profitability state (z_{mt}) that is observed to them, but unobserved to the econometrician. Furthermore, firms make assessments about q_m, and form beliefs q_{mt}^i that are a function of unobserved signals $\left(\text{i.e., } \tilde{q}_{mt}^i,\ \tilde{q}_{mt}^{-i}\right)$, which ultimately make q_{mt}^i a serially correlated and unobserved state as well; recall that q_{mt}^i is serially correlated as it can be defined recursively as a function of past beliefs q_{mt-1}^i. Also, because q_{mt}^i is affected by signals associated with learning from experience (i.e., a_{imt-1}) and learning from others (i.e., $\check{a}_{-imt-1}$), this serially correlated variable is also endogenous. To allow for these serially correlated states with continuous support that may also evolve endogenously, I employ a recently developed method for estimating retail entry/expansion models by Blevins et al. (2018) that combines Bajari et al. (2007) two-step approach with particle filtering techniques from machine learning. The key advantage of this two-step approach is that it allows me to introduce continuous observed and unobserved states in a dynamic game. Furthermore, this particle filtering approach accommodates for serially correlated unobserved states with continuous support that may evolve endogenously (i.e., posterior beliefs about perceived market profitability).

Note that an alternative approach to integrate out the unobserved states is Arcidiacono and Miller's (2011) iterative Expectation-Maximization method, which has been adopted by Igami and Yang (2016), as well as Hollenbeck (2017). The method I use has a few distinct advantages over Arcidiacono and Miller (2011). First, I do not need to make an *a priori* assumption about the number of unobserved market types. While many applications of Arcidiacono and Miller (2011) choose an arbitrary number of unobserved types, Igami and Yang (2016) make use of Kasahara and Shimotsu's (2009) identification results to find the minimum number of market types. Second, particle filtering is uniquely well-equipped to handle time-varying heterogeneity that may be both continuous and endogenous. The presence of learning in my model makes this advantage especially pronounced, since the posterior beliefs evolve endogenously based on past decisions of their own and their rivals, and that allowing for continuously distributed posterior beliefs is important as learning model estimates are likely affected by the coarseness of the grids used to discretize the posterior beliefs (Ching et al., 2013).

I now provide a brief overview of the steps in estimation, and leave the more technical details in the Appendix. This estimation approach follows two main steps. The first step is the particle filtering algorithm, which I use to obtain the posterior distribution for the unobserved states. The particle filtering proceeds by recursion, in which posterior weights help dictate how the unobserved states are re-drawn in each successive time period. This estimated posterior distribution then serves as an input for the second step of estimation, which leverages forward simulations as in Bajari et al. (2007). With the posterior distribution for the unobserved states, along with the inferred policy functions (for entry) and transitions for the exogenous observed states, I then forward simulate sequences of period profits into the future, and compare the expected values from equilibrium strategies versus strategies that deviate from the equilibrium. The criterion used in this second step ultimately penalizes cases in which the candidate parameters result in higher expected values for off-equilibrium strategies relative to equilibrium strategies.

4. Empirical setting

This section describes the empirical setting that I use to estimate the model of industry dynamics with learning. In particular, I will describe the industry and data, followed by a discussion about the key features of the data that satisfy the identification conditions I laid out in a preceding section. Finally, I proceed with exploratory analysis to highlight interesting patterns in the data.

4.1. Industry background

I study the evolution fast food outlets that primarily serve hamburgers from 1970 to 2005. In particular, I focus my attention on the five largest chains operating in Toronto, the largest and most populated Canadian city. These chains include A&W, Burger King, Harvey's, McDonald's and Wendy's. Although their relative standings have changed over time, these five chains remained as the most dominant forces in Canada's fast food industry by the end of 2005.

No other chains with national presence entered the industry but failed as a whole. Hence, the set of five 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 offered 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 first outlet in Canada was opened in 1986, while Taco Bell's first outlet in Canada was opened in 1981.

Canada, in general, has become a very important foreign market for American retail chains since the 1970s. Canada provides American chains a real growth option,[10] without the risk associated with more exotic markets overseas (Holmes, 2011). Not surprisingly, American chains tend to launch in Canada first 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. By 2005, McDonald's has 842 outlets in Canada, A&W has 252 outlets, then Burger King follows with 183, and 181 for Wendy's nationally. The largest domestic chain, Harvey's, boasts a store count of 217 outlets in Canada.

4.2. Market definition

I consider a Forward Sortation Area (FSA) as a local market. FSA designations are defined as the first three digits of a postal code and are loosely based on population. They are on average 1.8 square miles in many Canadian cities, and thus, comparable to American Census Tracts. Note that the markets I use are smaller than those used in other studies on retail competition and agglomeration. For example, Toivanen and Waterson (2005) and Yang (2012) use Local Authority Districts in the 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 (2014) focus on city markets in China. Because this research is focused on understanding retail clustering, we need a market definition that is as small as possible, but also large enough that we do not observe outlets located on land that borders two neighboring FSAs. 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 the largest Canadian city, Toronto. This selection then leaves me with a panel with over 100 markets over the course of 36 years.

I choose Toronto as the focal city of analysis for two main reasons, one substantive and one practical. First, Toronto has historically housed at least 20% of Canada's population. Not surprisingly, all of the fast food chains have their central Canadian headquarters in Toronto, even though each city typically has a local office. We see this focus towards Toronto reflected in the distribution of fast food outlets across Canada, as Toronto outlets account for nearly 20% of the country's total number of fast food outlets. Second, an important element of my analysis is to use the best possible and most relevant information about market characteristics for explaining location strategy decisions among the chains. Toronto, from a practical perspective, offers the best opportunity to control for market conditions using the richest data with fine granularity at both the geographic and time dimensions. A number of important market characteristics I utilize for my analysis are only available for the City of Toronto, or are offered by other cities but of a much poorer quality that is unsuitable for empirical analysis (e.g., inconsistent data construction over time, missing information for earlier years, not granular enough).

4.3. Data description

My data consists of four main parts. The first part contains hand-collected information about fast food location choice, while the remaining portions are market characteristics including demographics, local business activity, and transportation infrastructure from various sources. The choice of key market characteristics was largely guided by my conversations with fast food executives in charge of store development and expansion.

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 Toronto. This method allows me to identify the opening year based on the first year in which a particular outlet is listed in the phone directory, the closing year based on the last year in which a particular outlet is listed in the phone directory, and the outlet's location based on the exact address listed in the phone directory.

[10] Franchised chain growth in Canada is still markedly smaller than growth in America. Kosová and Lafontaine (2010) show that growth is about 29 percentage points lower in Canada as compared to the States.

Table 1
Summary statistics.

Variable	Mean	Std. dev.
Chain store decisions		
A&W active	0.0359	0.1862
Burger King active	0.055	0.228
Harvey's active	0.1618	0.3683
McDonald's active	0.4809	0.4997
Wendy's active	0.0294	0.169
Demographics		
Population	23,753	12,909
Property value ($ CDN)	550,520	281,721
Income ($ CDN)	78,545	38,007
Local business activity		
Number of entertainment establishments	0.012	0.1161
Number of amenity establishments	0.5248	1.2905
Number of specialty establishments	0.0575	0.2726
Number of malls	0.238	0.5343
Number of Home Depot locations	0.0185	0.1348
Transportation infrastructure		
Number of type 1 traffic lights	0.0231	0.1522
Number of type 2 traffic lights	4.8775	6.0274
Number of type 3 traffic lights	0.7209	1.1132
Number of subway stations	0.5109	0.9869
Firm-specific variables		
Nearby A&W outlets	0.485	0.912
Nearby Burger King outlets	0.689	1.233
Nearby Harvey's outlets	1.917	2.554
Nearby McDonald's outlets	6.972	6.736
Nearby Wendy's outlets	0.38	0.820
N	3672	

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). With the exception of one or two 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, whereby the first 3 digits gives me the FSA index. 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 FSA. We see from Table 1 that Harvey's and McDonald's have the largest presence in Toronto, as they are active in about 16% and 48% of the observations respectively.

Taken together, the market characteristics included in my data serve to emulate the "information sets" among the fast food executives who make restaurant location decisions. A number of the characteristics I have included are often used in *pro forma* analysis, which is often performed before committing to a new restaurant location, by the fast food chains.[11]

The first set of market characteristics capture local demographics. The first variable is population, which is available from the Canadian Census Profiles for the years 1986, 1991, 1996, 2001 and 2006 at the FSA level. I impute the missing years using the inferred population growth rates.[12] Additional information obtained from the Census includes the average income, and the average property value. Each FSA contains about 24,000 people, and on average, they have an income of $79,000 (CDN). Property value, on average, exceeds $500,000 (CDN) in Toronto. Note that the population density is quite high as the FSAs are in Toronto, which is Canada's largest and most metropolitan city.

In addition to local demographics, my data contains information about local business activity at the FSA level. Business activity may capture how viable businesses are in general across various parts of the city over time. First, I obtained detailed information from the City of Toronto's Municipal Licensing and Standards - Business Licenses and Permits about the number of newly approved and renewed business licenses over time for a variety of different types of business establishments. There are 71 different establishment categories that require licenses, so I narrow down these categories and focus on three main groups. Namely the number of licenses for entertainment, amenity and service, and specialty store businesses.[13] In addition to this licensed establishment data, I include information about the geographic diffusion of shopping centers in Toronto. The

[11] Special gratitude to Patricia Simiele (of McDonald's Canada) for giving me a detailed walk-through of their scouting and research process for new restaurant locations.
[12] Please refer to the Appendix for details.
[13] Please refer to the Appendix for details.

Table 2
Sample variance of firm decisions within and across markets.

Chain	Within	Across
A&W	0.0192	0.0346
Burger King	0.0323	0.0496
Harvey's	0.0995	0.1132
McDonald's	0.1479	0.1864
Wendy's	0.0183	0.0280

exact address and opening years for all shopping centers in Toronto are widely available through Wikipedia.[14] The remaining component of local business activity data consists of information about the location and entry timing of all Home Depot outlets, which I obtained by calling each Home Depot store and asking which year they opened. Home Depot presence is particularly relevant in my analysis as they have co-location arrangements with Harvey's, in which Harvey's agrees to operate small-scale outlets within Home Depot outlets.[15] At the very least, the presence of Home Depot outlets provides yet another proxy for the amount of business activity.

Furthermore, I augment the data with information about transportation infrastructure at the FSA level. Transportation infrastructure help capture how much traffic certain areas in the city experience. The first component includes information about the number of traffic lights in a FSA over time, which is obtained from the City of Toronto's Transportation Services - Traffic Management Centre. There are three different types of traffic lights. The first type have neither pedestrian countdown signals nor accessibility features for the blind, the second type have pedestrian countdown signals, and the third type have both pedestrian countdown signals and accessibility features.[16] The remaining piece of transportation infrastructure information I use includes the roll-out of subway stations, which I obtain using Wikipedia.

Finally, I construct a set of firm-specific variables using the number of incumbent own-brand outlets in adjacent FSA markets; these variables will serve as exclusion restrictions for identification purposes, which I will elaborate on in subsequent sections.

4.4. Identification of the model

I will now discuss the conditions needed for the parameters in my model to be identified. The way in which I interpret identification will be along the lines that "parameters of a model are identified given the assumed model structure" (Ching et al., 2013).

4.4.1. Condition 1: rich within and across-market variation

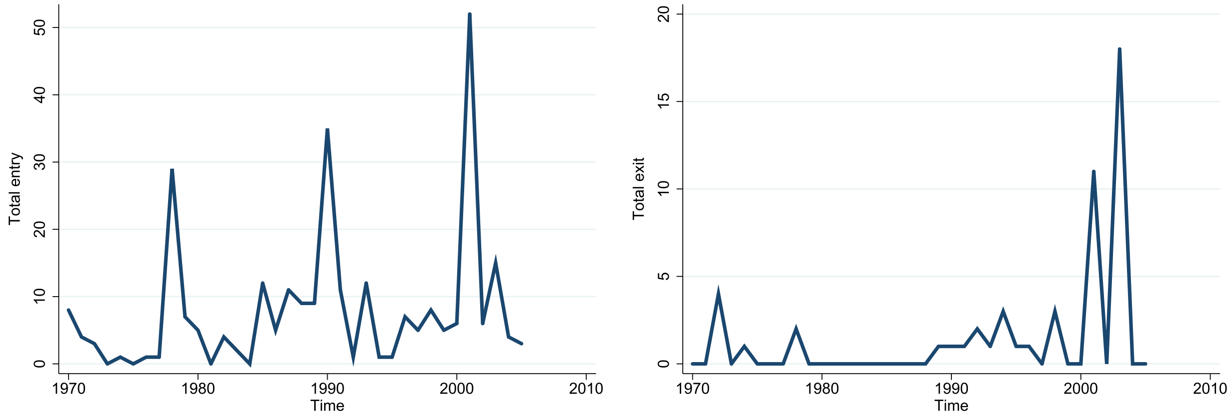
The first condition is that the panel data have variation in firm decisions both within and across markets will help with identification under the presence of unobserved heterogeneity. In particular, this rich panel variation can be used in conjunction with the parametric distributional assumptions for identifying the evolution of the unobserved serially correlated components (ω_{mt}) in the pay-off function; this general strategy in controlling for the unobserved states has also been used in Pakes (1986). For the exogenous and commonly known unobserved state, z_{mt} can be projected onto its lagged value (z_{mt-1}) to obtain the auto-covariance between z_{mt} and z_{mt-1}, which helps identify the degree of persistence (ρ). While the posterior beliefs about unobserved profitability $\left(q^i_{mt}\right)$ are also projected onto their lagged values $\left(q^i_{mt-1}\right)$ as a way to identify the parameters related to learning $\left(q^i_{m1}, \sigma^i_{m1}, \sigma^i_e, \sigma^i_f\right)$, the nature of Bayesian updating in learning makes the evolution of q^i_{mt} analytically distinct from the evolution of z_{mt} in an important dimensions.[17] Namely, q^i_{mt} is directly affected by the lagged entry decisions of both the firm itself as well as its rivals, while z_{mt} is independent of past decisions (a_{imt-1} and a_{-imt-1}).

A simple analysis of the variance for each chain's decision (a_{imt}) confirms that the data does indeed have such variation. Table 2 shows that there is both within-market and across-market variation. Here, I calculate the within-market variance as the variance across time in a_{imt} (averaged across all markets $m = 1, \ldots, M$), and analogously, I calculate the across-market variance as the variance across markets in a_{imt} (averaged across all years $t = 1, \ldots, T$). For some chains, such as Harvey's and McDonald's, this variation is particularly large. Furthermore, a general pattern we see is that the across-market variation is slightly more pronounced than the within-market variation.

4.4.2. Condition 2: initial conditions observed

The second condition is that the initial entry decision (a_{im0}) need to be observed in order to identify the prior mean and variance $\left(q^i_{m1},\ \sigma^i_{m1}\right)$ for the component of profitability (q_m) that firms face uncertainty about and learn; this similar argument has also been used in various empirical settings to identify the priors in Crawford and Shum (2005), Chintagunta, Jiang, and Jin (2008), and Dickstein (2015). Furthermore, differences in the initial conditions across firms will allow for heterogeneous priors across firms. Even if it is not feasible to observe the initial time periods, it is important to observe entry decisions as

[14] In the event that such information is not available on Wikipedia, I called the shopping centers directly for the opening year information.
[15] I thank Melissa Pannozzo (of Harvey's Canada) and Sandra Shaw (of Home Depot) for providing such insights.
[16] Please refer to Kapoor and Magesan (2014) for more details about the traffic light roll-out data.
[17] As Erdem and Keane (1996) point out, the "Bayesian learning model implies a particular form of state dependence and serial correlation."



(a) Entry (b) Exit

Fig. 1. Dynamics of entry and exit over time.

far back as possible, as beliefs about unobserved market profitability may be converging towards complete information. In the complete information scenario, Bayesian updating will be negligible, and the priors will essentially drop out of the model (Ching et al., 2013; Osborne, 2011) and hence cannot be identified if the data does not capture early periods where firms likely faced uncertainty.

This condition is trivially satisfied by the fact that my data spans as far back as 1970, which is roughly when the hamburger chains first entered Canada (Love, 1995); more specifically, A&W in 1971, Burger King in 1971, Harvey's in 1966, McDonald's in 1970, and Wendy's in 1973.

4.4.3. Condition 3: long panel

The third condition is that each market is observed over the course of many time periods, which is a common argument made in research about learning (e.g., Ackerberg, 2003; Che, Erdem, & Sabri Öncü, 2015; Ching, 2010; Ching et al., 2013; Coscelli & Shum, 2003; Crawford & Shum, 2005; Dickstein, 2015; Erdem & Keane, 1996; Narayanan & Manchanda, 2009; Narayanan, Chintagunta, & Miravete, 2007). Given the Bayesian updating formula, the signal noises associated with learning from experience $\left(\sigma_e^i\right)$ and learning from others $\left(\sigma_f^i\right)$ are identified by switches in a_{imt} and a_{-imt} respectively. Using a similar point as Dickstein (2015), each time period that is observed for a given cross-section increases the number of possible sequences of entry decisions. Under the context of my model, the number of empirical moments is 2^T, which will exceed the number of parameters provided that T is large enough. These sequences ultimately help generate variation in the information sets $\left(q_{mt}^i, \sigma_{mt}^i\right)$ across markets and firms (Ching et al., 2013).

In light of this condition, I note that there are 36 years of observations for each market (i.e., long panel). Thus, there are many opportunities to observe variation in both entry and exit (see Fig. 1). Since entry and exit reflect changes in a_{imt}, my data likely has enough inter-temporal variation to satisfy the third condition. I also show in Fig. 2 that there is variation in the average age of incumbent rivals at the time a chain enters a particular market; in particular, it is worth noting that Burger King appears to enter markets with rivals who have been active for a large number of years, while McDonald's tends to enter markets with rivals who are new to the market. This variation is relevant as well since the amount that firms are able to learn from one another may be affected by this source of variation.

4.4.4. Condition 4: rich support for observed states

The fourth condition is that the observed exogenous variables (x_{imt}) have rich support. As discussed in Kasahara and Shimotsu (2009), the maximum number of unobserved types that can be identified in a dynamic discrete choice game is of order $|x_{imt}|^{(T-1)/2}$. This condition can be achieved with a combination of a large number of observed states and/or the presence of observed states with continuous support. Furthermore, the previous requirement that T is large will also help with this condition. Under these conditions, the number of unobserved types that can be identified in a game approaches infinity in the limit, which is helpful as Bayesian updates of posterior beliefs in typical learning models are naturally continuous.

Here, I show that the observed market characteristics (x_{imt}) have rich support, as per Kasahara and Shimotsu's (2009) identification results for dynamic games with unobserved heterogeneity. A simple counting exercise confirms that this condition will hold as well. For example, if we consider only the discrete states (see Table 3) for McDonald's, then $|x_{imt}| = 3 \times 13 \times 5 \times 4 \times 2 \times 3 \times 39 \times 8 \times 6 \times 37 = 8{,}760{,}960 \simeq 2^{23}$, which means that the maximum number of unobserved types that can be identified from the data is

$$|x_{imt}|^{(T-1)/2} = (2^{23})^{(36-1)/2} \simeq 2^{391} \tag{12}$$

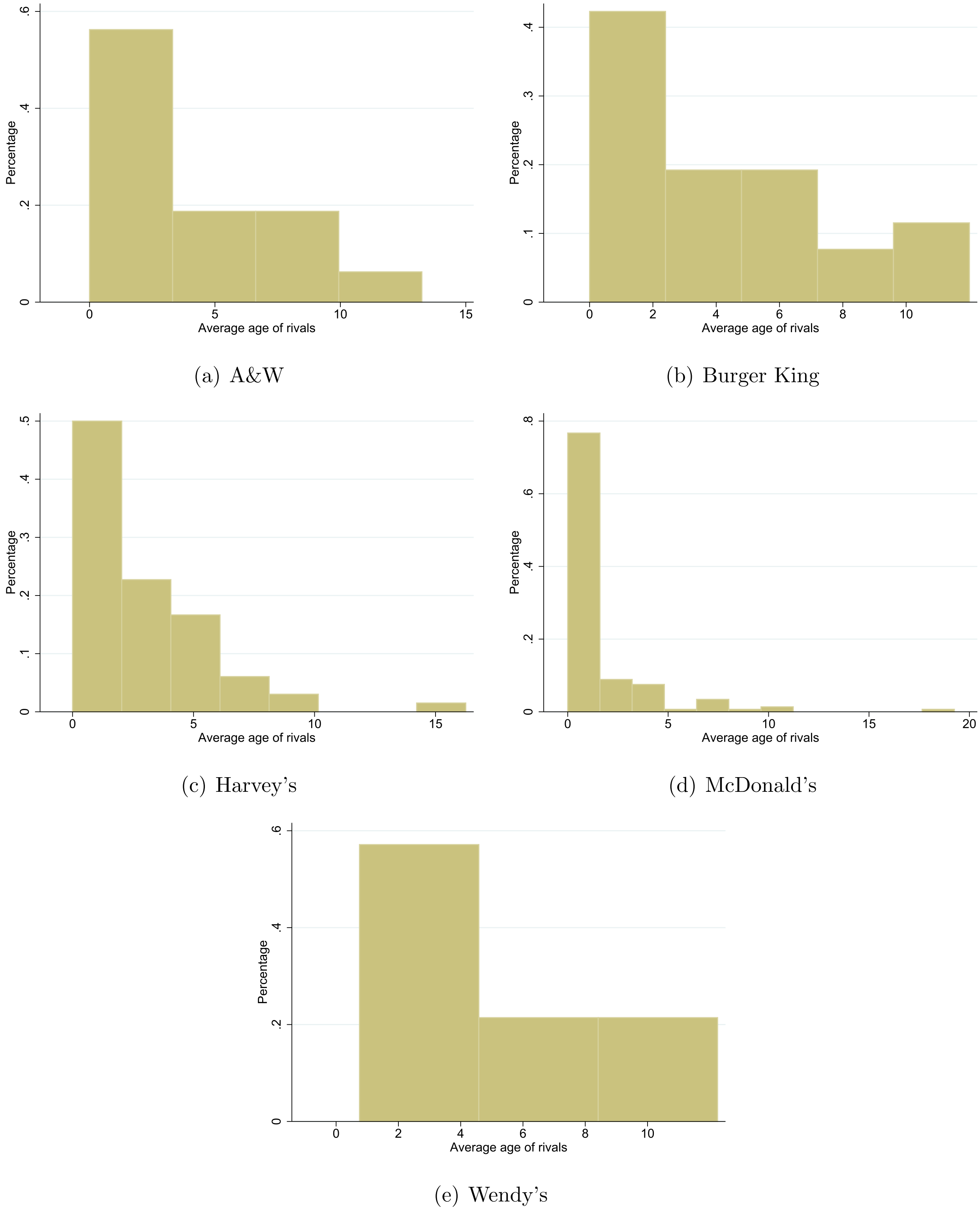


Fig. 2. Average age of rivals upon entry.

Note also that this calculation is conservative as it uses only the discrete observable market characteristics. If we include population, income, and property value, $|x_{imt}|^{(T-1)/2}$ will be trivially equal to ∞, as these market characteristics have continuous support. Recall that it is important that I can identify continuously distributed unobserved states, especially with respect to the posterior beliefs in learning.

Table 3
Size of support for observable discrete state variables.

Variable	Size of support
Number of entertainment establishments	3
Number of amenity establishments	14
Number of specialty establishments	5
Number of malls	4
Number of Home Depot locations	2
Number of type 1 traffic lights	3
Number of type 2 traffic lights	39
Number of type 3 traffic lights	8
Number of subway stations	6
Nearby A&W outlets	6
Nearby Burger King outlets	6
Nearby Harvey's outlets	12
Nearby McDonald's outlets	37
Nearby Wendy's outlets	6

4.4.5. Condition 5: valid exclusion restrictions

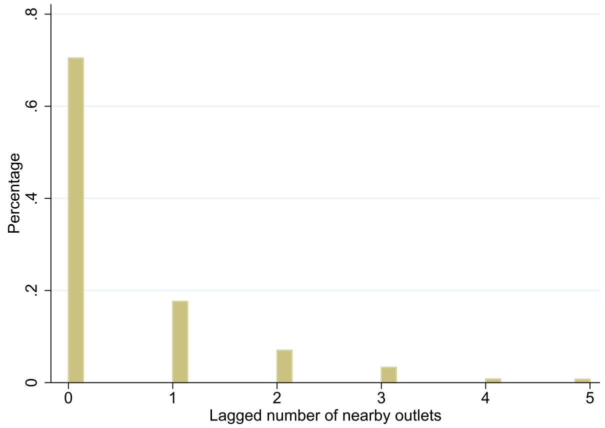
The fifth and final condition is that there needs to be firm-specific payoff shifters in the form of exclusion restrictions for nonparametric identification of a dynamic game (Bajari, Chernozhukov, Hong, & Nekipelov, 2009); these exclusion restrictions will help identify the strategic interactions (θ_2), fixed costs (θ_3), and entry costs (θ_3). Note that in standard models of entry, a_{imt-1} would qualify as an exclusion restriction, as in those models, lagged entry of rivals does not enter a firm's pay-off directly. However, the learning process in the model makes a_{imt-1} an invalid exclusion restriction as the rivals' past entry decisions enter a firm's payoff directly via its Bayesian updates of unobserved profitability. For this reason, it is imperative to have at least one firm-specific pay-off shifter in the set of exogenous market characteristics (x_{imt}). Such variation would then create unique market shocks that induce only a subset of chains to enter a market. This requirement is along the lines of the instruments needed to identify microeconomic models of social interaction (e.g., Manski, 1993), as strategic interactions can be interpreted as a form of peer effects.

Given the data available, I am able to make use of the fact that the retail chains may be affected by the presence of own-brand outlets in neighboring FSA markets. Insights from the fast food executives[18] suggest that the threat of cannibalization or encroachment serves as a pertinent risk factor when deciding on whether to enter a new market or not; if the new market in consideration is too close to nearby markets they are already active in, then entering may not be as viable. Furthermore, research about the same industry by Igami and Yang (2016) corroborates this claim by providing evidence of cannibalization effects. For this reason, I use the lagged number of active own-brand outlets in adjacent markets as an exclusion restriction. This variable would be an appropriate exclusion restriction as it has a direct impact on the chain in question, but will not have a direct impact on the payoffs of rival chains. It is also worth mentioning that this exclusion restriction inspired by Igami and Yang's (2016) insights about cannibalization has also been used to develop similar instrumental variables for chain store growth in Couture and Handbury's (2015) study about urban development. Note that in theory, it is entirely possible that the lagged number of own-brand outlets in nearby markets may have a positive effect, say via economies of density (e.g., Holmes, 2011; Nishida, 2015). However, the discussion in Igami and Yang (2016) suggest that such density economies are unlikely to outweigh cannibalization concerns. The authors point to evidence that at least for McDonald's, the outlets do not radiate outward from the headquarters over time. Furthermore, their interviews with the store-development officer at McDonald's indicates that supply chain efficiencies are not taken into account by their team when deciding where and when to enter.

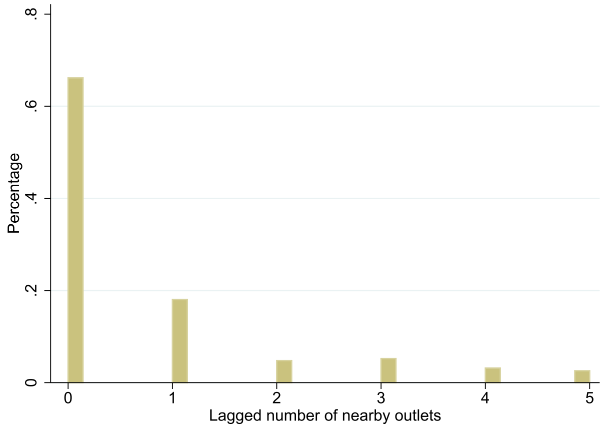
Fig. 3 illustrates the variation in this exclusion restriction. For many of the chains, the lagged number of nearby outlets ranges from 0 to 5, while for McDonald's, the number of nearby outlets can be as large as 30. Furthermore, Table 4 highlights that the number of nearby outlets provides some unique chain-specific variation, as the cross-chain correlations are not equal to 1; in fact, they all lie below 0.9. Therefore, it is likely that this variable generates unique shocks that affect only a subset of the chains across markets and time.

Table 5 confirms that firms are indeed averse to entering markets that are surrounded by neighboring markets that they already have significant presence in, which is in line with the insights from both the industry and past research. Furthermore, note that the number of own-brand outlets in adjacent markets has a statistically significant relationship with entry for all of the chains, and the F-stats are no smaller than 55.41 (i.e., all have p-values of 0.000) across each of the chains. Therefore,

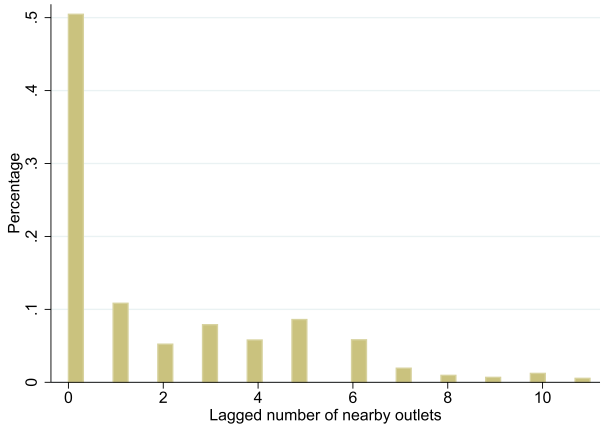
[18] Please refer to the Appendix for more details about my phone and face-to-face interviews with the executives.



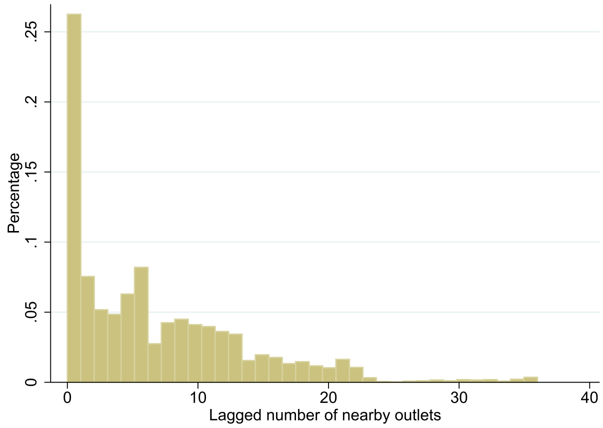
(a) A&W



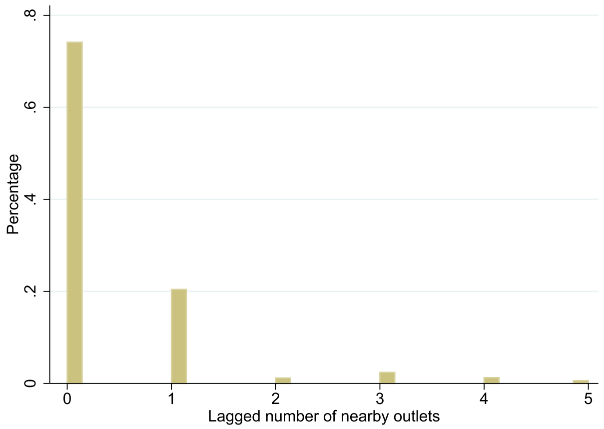
(b) Burger King



(c) Harvey's



(d) McDonald's



(e) Wendy's

Fig. 3. Distribution of the lagged number of outlets in adjacent markets.

I interpret these results as further validating the use of this proposed exclusion restriction. Harvey's appears to be most sensitive to encroachment, followed by Burger King, Wendy's, McDonald's, and A&W. This result is broadly consistent with Igami and Yang (2016), who show that own-brand business stealing effects are more pronounced for non-McDonald's chains.

Table 4
Cross-chain correlations in the lagged number of nearby outlets.

Variables	A&W	Burger King	Harvey's	McDonald's	Wendy's
A&W	–				
Burger King	0.607	–			
	(0.000)				
Harvey's	0.530	0.618	–		
	(0.000)	(0.000)			
McDonald's	0.580	0.755	0.771	–	
	(0.000)	(0.000)	(0.000)		
Wendy's	0.574	0.819	0.528	0.648	–
	(0.000)	(0.000)	(0.000)	(0.000)	

Table 5
Relationship between entry and the number of nearby outlets of same brand.

	(1)	(2)	(3)	(4)	(5)
	A&W	Burger King	Harvey's	McDonald's	Wendy's
Lagged own neighboring outlets***	−0.0249*,**,***	−0.0429***	−0.0572***	−0.0349***	−0.0375***
	(0.00335)	(0.00313)	(0.00208)	(0.000975)	(0.00320)
Constant	0.0239***	0.0254***	0.0522***	0.238***	0.0152***
	(0.00280)	(0.00361)	(0.00619)	(0.00871)	(0.00251)
Market fixed effects	Yes	Yes	Yes	Yes	Yes
Time dummy	Yes	Yes	Yes	Yes	Yes
Observations	3672	3672	3672	3672	3672
R^2	0.02	0.05	0.17	0.26	0.04

Clustered standard errors in parentheses.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

4.5. Descriptive analysis of the data

I conduct exploratory analysis of the data in this section. The purpose of this exercise is to uncover relevant factors behind a retail chain's entry and exit decision. I will also use this section to illustrate potential caveats of relying on reduced patterns alone to provide insights about learning in fast food entry.

For the reduced form analysis, I will use the following linear probability regression to describe the firm's decision a_{imt}:

$$a_{imt} = \alpha x_{imt} + \zeta_1 \check{a}_{-imt-1} + \zeta_2 a_{-imt-1}(1 - a_{imt-1}) + \eta_i + \eta_m + \eta_t + \xi_{imt}, \tag{13}$$

where x_{imt} are the observed market characteristics, η_i is a chain dummy, η_m is a market fixed effect, η_t is a time dummy, and ξ_{imt} is the i.i.d. error term. I am also interested in seeing how a rival's past presence ($\check{a}_{-imt-1}$) affects chain i,[19] as captured by the parameter ζ_1. Note also I allow this rival presence effect to have a differential impact for potential entrants (relative to incumbents), as reflected through the parameter ζ_2. One would expect $\zeta_1 < 0$, given that the Canadian fast food industry has been shown to be quite competitive (Igami & Yang, 2016). The main hypothesis of interest in this case would be whether there is attenuation in the competition effect in cases where the chain i has more to gain from learning (i.e., potential entrant who has not had as many opportunities to learn from experience); there would be evidence of attenuation in the competition effect if $\zeta_2 > 0$.[20]

Table 6 provides the linear probability regression results. The first specification has no controls, time dummies, and market fixed effects. The second specification has no controls, but with time dummies and market fixed effects. Next, I add demographic information in the third specification. For the fourth specification, I include also the local business activity data. The fifth column includes the infrastructure information. Finally the sixth specification includes the firm-specific exclusion restrictions in the form of nearby own-brand outlets in adjacent markets. These six specifications are displayed as columns 1 to 6 respectively in the table.

Before discussing the key insights about the impact that rivals have on entry, I will first describe some observable drivers of entry. It is worthwhile describing the impact of various market characteristics on entry, since in my setting, I have particularly unique data that can also shed light on the impact of other retail activity and traffic infrastructure on fast food entry, both of

[19] Recall from an earlier section that I define $\check{a}_{-imt-1}$ such that $\check{a}_{-imt-1} = 1$ if $a_{jmt-1} = 1$ for any $j \neq i$.
[20] This intuition is similar to Ackerberg (2001), who looks at the difference in the advertising effects between inexperienced and experienced customers as a way to measure the informativeness of advertising. In my scenario, instead of the informativeness of advertising, I am interested in the informativeness of rival incumbents' past entry decisions.

Table 6
Estimates from linear probability regressions.

	(1)	(2)	(3)	(4)	(5)	(6)
Rival incumbents	−0.0118**	−0.254***	−0.260***	−0.261***	−0.264***	−0.260***
	(0.00455)	(0.00565)	(0.00568)	(0.00568)	(0.00568)	(0.00568)
Rival incumbents × New entrant	0.0406***	0.0328***	0.0335***	0.0334***	0.0335***	0.0323***
	(0.000620)	(0.000607)	(0.000623)	(0.000633)	(0.000633)	(0.000649)
log(Population)			0.0263***	0.0263***	0.0232***	0.0233***
			(0.00699)	(0.00700)	(0.00701)	(0.00700)
log(Property value)			0.0592***	0.0563***	0.0511***	0.0534***
			(0.00730)	(0.00763)	(0.00768)	(0.00767)
log(Income)			−0.0208*	−0.0189*	−0.0158	−0.0189*
			(0.00902)	(0.00914)	(0.00914)	(0.00913)
Number of entertainment establishments				−0.0129	−0.00767	−0.00846
				(0.0167)	(0.0167)	(0.0167)
Number of amenity establishments				−0.00238	−0.00473*	−0.00477*
				(0.00193)	(0.00199)	(0.00199)
Number of specialty establishments				0.00392	0.00205	0.000248
				(0.00757)	(0.00756)	(0.00755)
Number of malls				0.0112	0.00583	0.00804
				(0.0116)	(0.0119)	(0.0118)
Number of Home Depot locations				−0.00377	0.00557	0.00395
				(0.0164)	(0.0165)	(0.0165)
Number of type 1 traffic lights					−0.0215	−0.0361
					(0.0295)	(0.0295)
Number of type 2 traffic lights					0.00395***	0.00400***
					(0.000756)	(0.000755)
Number of type 3 traffic lights					−0.0221***	−0.0215***
					(0.00418)	(0.00417)
Number of subway stations					0.0312***	0.0309***
					(0.00678)	(0.00677)
Lagged own neighboring outlets						−0.00656***
						(0.000752)
Constant	0.0299***	−0.0765***	−0.889***	−0.874***	−0.820***	−0.805***
	(0.00523)	(0.00548)	(0.129)	(0.130)	(0.130)	(0.130)
Chain dummy	Yes	Yes	Yes	Yes	Yes	Yes
Market fixed effects	No	Yes	Yes	Yes	Yes	Yes
Time dummy	No	Yes	Yes	Yes	Yes	Yes
Observations	18360	18360	18180	18180	18180	18180
R^2	0.38	0.45	0.45	0.45	0.46	0.46

Clustered standard errors in parentheses.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

which have not yet been explored in empirical research (Aguirregabiria & Suzuki, 2014). When discussing the impact of various market characteristics on entry, I will focus on column 6, which provides the complete set of market shifters. Focusing on the demographics, we see that population and property value both have positive effects on entry, while income has a negative effect; the negative income effect is consistent with the notion that fast food is an "inferior" good. In terms of local business activity, the number of amenity establishments (e.g., car rental, hair salon, holistic center, laundry, and parking) actually have a negative effect on entry. Furthermore, we see noticeable effects of infrastructure on market attractiveness, as the number of type 2 traffic lights as well as the number of subway stations have positive effects on entry; these findings are consistent with the fact that both of these variables provide good measures of foot traffic.[21]

The main findings from this exploratory analysis are with respect to the effect from rival incumbents (ζ_1) and its interaction with whether or not the chain is a potential entrant (ζ_2). In all the specifications, we see that the effect from rival incumbents is negative, as one would expect in a competitive industry. Furthermore, this business stealing effect appears to be dampened for new entrants. Thus, I conclude that there exists raw patterns in the data that would be consistent with learning from others. However, there are important reasons why we should believe that these patterns are not exclusively generated by learning, which ultimately motivate a more structural approach to studying learning in industry dynamics.

First, the fact that ζ_1 becomes more negative and/or ζ_2 becomes less positive as we progressively introduce more controls for market attractiveness suggest that unobserved market heterogeneity is a relevant and important driver behind entry decisions. Note that this empirical result about heterogeneity is not unique, as previous studies have also demonstrated the biasing

[21] Recall that type 2 traffic lights accommodate for pedestrians with their countdown features.

properties of heterogeneity. This identification issue was brought up early on by Berry (1992). Furthermore, in Arcidiacono and Miller (2011), Monte Carlo simulations demonstrate that competition effects will be underestimated when unobserved heterogeneity is ignored. On a similar note, a recent application by Igami and Yang (2016) confirms that the omission of unobserved heterogeneity attenuates the competition effect estimates. Consequently, biased estimates of competition will lead to misleading counterfactual analysis (Orhun, 2013). Finally, Blevins et al. (2018) have demonstrated that model fit improves when serial correlation is properly accounted for.

Second, even if one is able to perfectly control for time-varying unobserved heterogeneity, it is not entirely clear what the endogenous mechanism is behind the apparent attenuation in the competition effect from these empirical patterns alone. For instance, demand externalities may be a factor as well if there is a contemporaneous positive effect of one rival store's presence, in the form of increased traffic or exposure towards that market (e.g., Cosman, 2014; Datta & Sudhir, 2011; Eppli & Benjamin, 1994; Konishi, 2005; Jardim, 2015; Thomadsen, 2012; Vitorino, 2012; Zhu, Singh, & Dukes, 2011). Absent the well-defined structure of a Bayesian learning process (and forward-looking behavior), it is very difficult to distinguish between demand externalities and learning from others in a reduced-form manner (Ching et al., 2013).

Finally, failure to allow for forward looking incentives may lead to misleading conclusions, as some firms likely have preemptive motives (e.g., Igami & Yang, 2016) and may be engaged in strategic learning (e.g., Bikhchandani et al., 1998; Chamley, 2004; Chamley & Gale, 1994; Gale, 1996). Interestingly, the biases associated with competition and learning are somewhat nuanced when we take into account all these model features. For example, improper accounting of unobserved heterogeneity may lead us to understate the negative effect of competition and overstate the positive effect of learning from others in the short-run. However, since firms are forward looking, understating the negative effect of competition will mean that the long-run value of preempting rivals is going to be understated. Thus, any early entry behavior we observe may be improperly attributed to a low option value of learning via strategic delay, and consequently, bias downwards the positive effect of learning from others in the long-run.

Taken together, these descriptive empirical patterns along with the corresponding discussion motivate a deeper analysis of the underlying drivers behind the attenuation we see in the competition effect among new entrants (relative to incumbents). Furthermore, the shortcomings of this reduced form analysis justifies my reliance on the proposed structural model that jointly accounts for forward looking behavior, strategic interactions, Bayesian learning, and flexibly specified unobserved heterogeneity.

5. Main results

In the first-stage of estimation, I use $R = 5000$ simulations to approximate the distribution of ω_{mt} via the particle filtering procedure described earlier.[22] The policy function approximations are estimated using sieve maximum likelihood, which includes all exogenous variables, their interactions up to a second order, and interactions between exogenous variables and ω_{mt}.

For the second-stage estimation, I generate $K = 1000$ random inequalities, where each inequality compares the forward-simulated value functions under an equilibrium strategy versus under an alternative strategy. To generate the values under an alternative strategy, I randomly perturb the coefficients of the estimated reduced form policy function, and forward simulate beginning from an initial state (which is drawn randomly from the sample).

5.1. Summary of estimates

Table 7 contains the estimates of the key structural parameters. This includes the payoff parameters, the degree of persistence in the law of motion for the serially correlated z_{mt}, as well as the parameters that characterize the learning process behind unobserved profitability (q_m). Standard errors are obtained via bootstrapping.[23]

In terms of the key payoff parameters,[24] my estimates confirm that the retail chains operate in a competitive landscape. The presence of rivals has a noticeable business-stealing effect, as one would expect. Furthermore, fixed costs appear to be lower than entry costs, which suggests that the cost of operating in a market is less than the cost of entering a market. Entering a new market may be particularly costly as it likely involves investments in real estate, costs from research during pro forma analysis, and expenditures from the recruitment/retention of store managers or owner-operators.

The estimates also shed light on the time-varying unobserved heterogeneity. I find evidence that z_{mt} is persistent, which suggests that there remains some information about market profitability that is known to all the firms, but unknown to the econometrician. Furthermore, because the estimated persistence does equal unity, the market-specific heterogeneity also varies with time.

I will now discuss the key insights from the estimates are regarding the uncertainty firms face about market profitability, as well as the learning they engage in as a means to resolve this uncertainty. First, there is evidence of uncertainty as the inferred prior variance is not equal to zero for all of the chains. Second, there is heterogeneity in uncertainty and learning across the

[22] Blevins (2016) demonstrates in his Monte Carlo analysis that it is possible to obtain unbiased estimates using as little as $R = 50$ simulations when estimating a single-agent dynamic model via particle filtering.
[23] I use 100 replications with replacement from a sample of 102 markets for 36 years.
[24] For the sake of brevity, the remaining estimates in the payoff function (i.e., impact of various market characteristics) are found in the Appendix.

Table 7
Estimates for the model parameters.

	Estimate	Std. error
Payoff function		
Competition	−0.358	0.000
Fixed costs	−0.420	0.002
Entry costs	−1.822	0.025
Serial correlation		
Persistence (ρ)	0.699	0.340
Prior mean $\left(q_{m1}^i\right)$		
A&W	0.422	0.204
Burger King	0.917	0.444
Harvey's	0.793	0.384
McDonald's	0.960	0.463
Wendy's	0.656	0.317
Prior variance $\left(\sigma_{m1}^i\right)$		
A&W	0.036	0.017
Burger King	0.844	0.407
Harvey's	0.929	0.450
McDonald's	0.683	0.331
Wendy's	0.763	0.368
Variance of own signal $\left(\sigma_e^i\right)$		
A&W	0.032	0.015
Burger King	0.279	0.135
Harvey's	0.046	0.022
McDonald's	0.098	0.047
Wendy's	0.821	0.397
Variance of rival signal $\left(\sigma_f^i\right)$		
A&W	0.740	0.358
Burger King	0.395	0.191
Harvey's	0.655	0.316
McDonald's	0.169	0.081
Wendy's	0.710	0.343

Note: Standard errors are from bootstrapping across markets.

chains. For example, McDonald's has the largest prior mean and lowest variance in signals from its observations of rivals, while A&W has the lowest prior variance and lowest variance in signals from its own past experience. In other words, the results suggest that McDonald's has an *ex ante* opportunistic view about the viability of markets, and is also able to extract information effectively from the observation of rival's past presence, while A&W is most confident about its *ex ante* assessment of the markets and is able to learn effectively from its past experiences in the market. Finally, the variance of signals in learning from experience is lower than the variance of signals in learning from others for most of the firms. Therefore, learning from experience is more effective at resolving the uncertainty than learning from others.

As past rival presence has informational value, these results point to an additional source of attenuation in the competition effect, above and beyond the usual explanation of unobserved market heterogeneity. Therefore, learning from others may provide a new explanation for why markets appear "less competitive" than they really are. I will explore further the impact that learning from others has on dampening business-stealing effects.

5.2. Simulation analysis

The structural estimates allow me to conduct model simulations and counterfactuals. Through these simulations, I uncover more details about the learning process, as well as its role in potentially dampening the effects of rival competition. To carry out the simulations, I implement a forward simulation approach akin to Benkard, Bodoh-Creed, and Lazarev (2010). The forward simulations are initialized using market characteristics and market structure for the first year for each market. Using the estimates, inferred policy functions, and the SUR process for the exogenous market characteristics, I then forward simulate the entry/exit behaviors, the serially correlated unobserved state z_{mt}, and posterior beliefs about q_m. In each market, I simulate 500 sample paths (with length of 36 years).

Table 8 summarizes the simulated posterior beliefs $\left(q_{mt}^i, \sigma_{mt}^i\right)$ in equilibrium. A few patterns emerge from these results. First, the posterior mean $\left(q_{mt}^i\right)$ is lower than the prior mean $\left(q_{m1}^i\right)$ for all of the chains. So it appears that the Bayesian updates

Table 8
Summary statistics for simulated beliefs.

Variable	Mean	Std. dev.
Posterior mean $\left(q^i_{mt}\right)$		
A&W	0.019	0.076
Burger King	0.03	0.152
Harvey's	0.023	0.13
McDonald's	0.029	0.158
Wendy's	0.026	0.113
Posterior variance $\left(\sigma^i_{mt}\right)$		
A&W	0.004	0.006
Burger King	0.041	0.138
Harvey's	0.031	0.152
McDonald's	0.026	0.112
Wendy's	0.059	0.128

are leading to more pessimistic views about q_m for all of the chains. Second, there remains to be heterogeneity in beliefs - albeit in lesser degree as compared with the prior beliefs - across the chains, and Burger King and McDonald's are still *ex post* the most optimistic as compared with their rivals. Furthermore, there is also heterogeneity in posterior beliefs across markets and time, as indicated by the standard deviation of q^i_{mt} and σ^i_{mt}. Third, we see noticeable changes in the precision of their beliefs, as the posterior variance $\left(\sigma^i_{mt}\right)$ is markedly smaller than the prior variance $\left(\sigma^i_{m1}\right)$ for all of the chains. Furthermore, A&W and McDonald's continue to have the most precise beliefs after learning takes place.

I illustrate further the diminishing posterior variance with Fig. 4. Here I plot the simulated evolution of σ^i_{mt} for each of the chains (averaged across markets). In general, we see that the posterior variance approaches zero over time. Furthermore, there is heterogeneity in the rate at which the posterior variance approaches zero; for example, Harvey's and McDonald's experience the more discontinuous and sudden improvements in their *ex post* precision regarding q_m, while the improvements in precision are slightly smoother for A&W, Burger King, and Wendy's.

The next set of simulations I conduct are aimed to investigate the potential dampening that learning from others may have on business stealing effects. To proceed, I consider a hypothetical scenario in which each chain is deciding whether or not to enter markets with a *one-time* increase of four rival incumbents at $t = 1$, such that $\sum_{-i \neq i} a_{-imt-1} = 4$. I then forward simulate the counterfactual industry dynamics as before except with this change in the initial state. The fact that there are more rival competitors initially will increase the effect of business stealing $\left(\theta_2 \sum_{-i \neq i} a_{-imt}\right)$. On the other hand, their presence may provide positive signals about the markets.

Fig. 5 illustrates the impact of increasing the initial number of rival incumbents. The first panel plots the difference over time between the counterfactual and equilibrium posterior means (averaged across markets and across firms), the second panel plots the difference over time between the counterfactual and equilibrium posterior variances (averaged across markets and across firms), and the third panel plots the difference over time between the counterfactual and equilibrium absolute value of business stealing effects (averaged across markets and across firms). These plots confirm that indeed the presence of rival incumbents will induce firms to update their beliefs in a positive manner over time. Furthermore, the posterior variance is reduced, especially so in the intermediate-term. At the same time, the presence of rival incumbents assures a more competitive retail landscape, as the business stealing effect is clearly larger as well; note that the business stealing effect is at its highest immediately after the initial market's endowment of four rival incumbents, and diminishes over time as a subset of rival incumbents exit the market. Taken together, these plots suggest that learning from others may dampen the business-stealing effects of rival presence, and especially so over time. In terms of magnitudes, the elevated business stealing effect is still quite dominating, and is larger than positive updates of the posterior beliefs by multiples of 3 or more.

The counterfactual simulations of the model primitives also demonstrates a possible incentive to strategically delay entry (e.g., Bikhchandani et al., 1998; Chamley, 2004; Chamley & Gale, 1994; Gale, 1996). Entering a market immediately following the entry of rivals may not be optimal for the following reasons. First, the maximum degree that uncertainty can be reduced may occur at least a few years following the entry of rivals. Second, the posterior beliefs about the market may increase over time. Finally, eventual exit of a subset of rivals will ultimately reduce the business stealing effects in the future.

6. Conclusion

This research offers new insights about learning from others that may help brick-and-mortar retailers in their location intelligence. By quantifying the role of learning in uncertainty reduction, retailers can turn to additional (and perhaps less research-intensive) sources of information when they search for new locations to set up their stores. I develop these insights by introducing a novel model of dynamic retail entry with learning, and estimating the model using state-of-the-art techniques from dynamic games and statistical learning. Through simulation analysis of the estimated model, I give the first empirical evidence that uncertainty exists, diminishes over time, and can be partially reduced via learning from others. While these insights are framed around the Canadian fast food industry, they can be extended well beyond this specific retail sector. In general, any

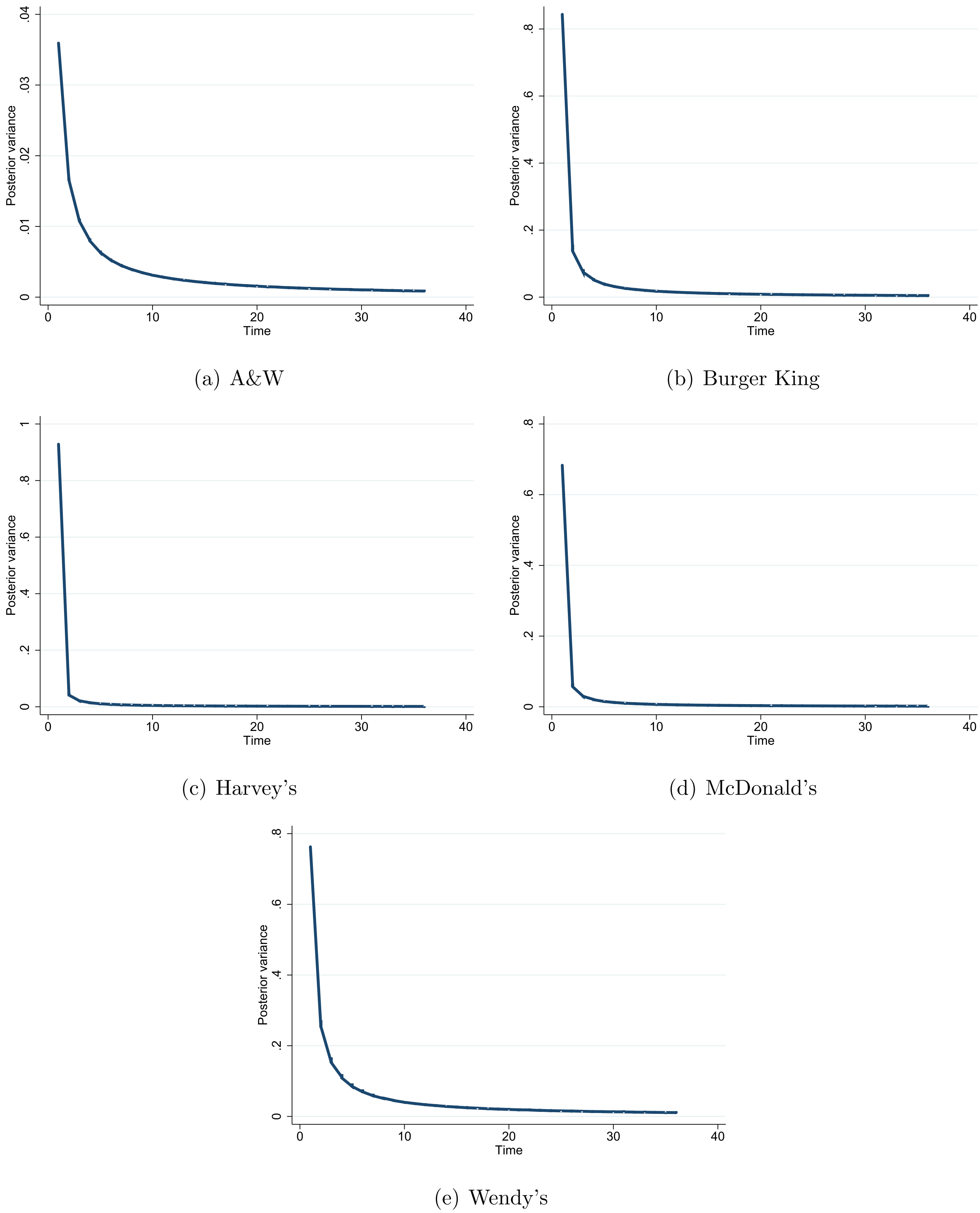
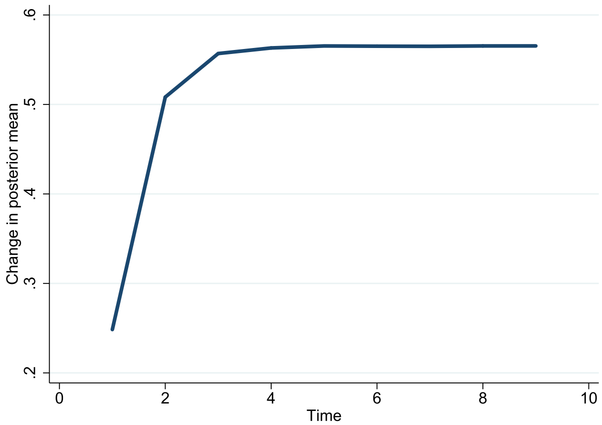
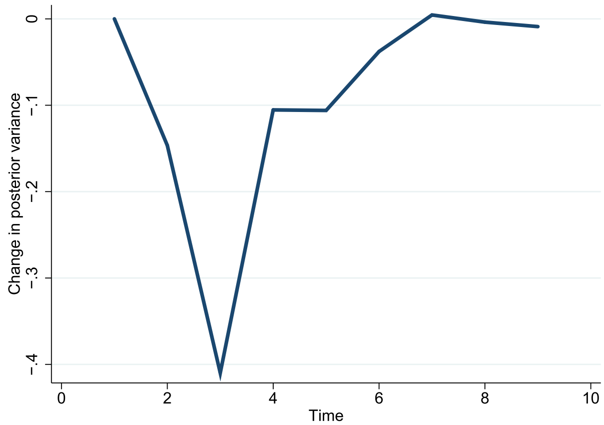


Fig. 4. Diminishing uncertainty over time.

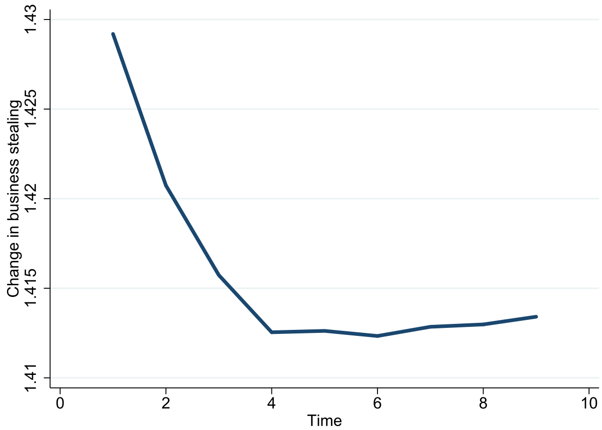
retail sector that does not benefit from granular consumer-level data (e.g., loyalty programs, mobile application tracking) may find these insights particularly valuable for their site selection strategies. By leveraging the uncertainty reduction benefits of learning from others, retailers can better avoid entering markets that turn out to be unprofitable in the long-run.



(a) Posterior Mean (q^i_{mt})



(b) Posterior Variance (σ^i_{mt})



(c) Business Stealing ($\theta_2 \sum_{-i \neq i} a_{-imt}$)

Fig. 5. Impact of hypothetical one-time increase in the initial number of rival incumbents.

Given the challenging nature of my research objective, this paper has made two important contributions: modeling and estimation. Standard models of dynamic retail entry are unable to capture the key learning dynamics, while standard estimation approaches are unable to uncover key patterns in beliefs as they evolve via learning. To that end, I believe these methodological innovations could potentially be adapted for other retail settings (e.g., coffee chains, convenience stores, fast casual restaurants). The extended discussion about identification offers a salient set of verifiable conditions to determine whether or not another empirical setting can be analyzed using the methodological framework I have presented.

There are limitations of my empirical framework. In particular, my analysis abstracts away from joint entry decisions across multiple markets. That said, a promising direction for future research would be to extend my empirical framework to allow for dynamic retail network decisions. Being able to capture these network dimensions may allow future researchers to better understand an exploitation-exploration trade-off where retailers choose between markets they have already discovered with less uncertainty (i.e., exploitation strategy) versus markets they have yet to discover that carry more uncertainty (i.e., exploration strategy). A better characterization of this trade-off would offer insights that would further improve location intelligence efforts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2019.09.005.

References

Ackerberg, D. (2001). Empirically distinguishing informative and prestige effects of advertising. *RAND Journal of Economics*, *32*, 316–333.
Ackerberg, D. (2003). Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, *44*, 1007–1040.
Aguirregabiria, V., & Suzuki, J. (2014). Identification and counterfactuals in dynamic models of market entry and exit. *Quantitative Marketing and Economics*, *12*, 267–304.
Arcidiacono, P., Bayer, P., Blevins, J.R., & Ellickson, P.B. (2016). Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *The Review of Economic Studies*, *83*(3), 889–931.
Arcidiacono, P., & Miller, R. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, *79*, 1823–1867.
Bajari, P., Benkard, L., & Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, *75*, 1331–1370.
Bajari, P., Chernozhukov, V., Hong, H., & Nekipelov, D. (2009). Nonparametric and semiparametric analysis of a dynamic discrete game. *Working paper*.
Beamish, P., Jung, J., & Kim, H.-H. (2003). *American fast food in Korea*. Richard Ivey School of Business Foundation. March 2003.
Bell, D., & Shelman, M. (2011). KFC'S radical approach to China. Harvard Business Review, November 2011.
Benkard, L., Bodoh-Creed, A., & Lazarev, J. (2010). Simulating the dynamic effect of horizontal mergers: U.S. Airlines. *Working paper*.
Beresteanu, A., Ellickson, P., & Misra, S. (2010). The dynamics of retail oligopoly. *Working paper*.
Berry, S. (1992). Estimation of a model of entry in the airline industry. *Econometrica*, *60*, 1331–1370.
Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, *12*, 151–170.
Blevins, J.R. (2016). Sequential Monte Carlo methods for estimating dynamic microeconomic models. *Journal of Applied Econometrics*, *31*(5), 773–804.
Blevins, J., Khwaja, A., & Yang, N. (2018). Firm expansion, size spillovers and market dominance in retail chain dynamics. *Management Science*, *64*, 4070–4093.
Caplin, A., & Leahy, J. (1998). Miracle on sixth avenue: Information externalities and search. *The Economic Journal*, *108*, 60–74.
Chamley, C. (2004). *Rational herds: Economic models of social learning*. Cambridge: Cambridge University Press.
Chamley, C., & Gale, D. (1994). Information revelation and strategic delay in a model of investment. *Econometrica*, *62*, 1065–1085.
Che, H., Erdem, T., & Sabri Öncü, T. (2015). Consumer learning and evolution of consumer brand preferences. *Quantitative Marketing and Economics*, *13*, 173–202.
Ching, A. (2010). Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization*, *28*, 619–638.
Ching, A.T., Erdem, T., & Keane, M.P. (2013). Learning models: An assessment of progress, challenges, and new developments. *Marketing Science*, *32*(6), 913–938.
Chintagunta, P., Jiang, R., & Jin, G. (2008). Information, learning and drug diffusion. *Quantitative Marketing and Economics*, *7*, 399–443.
Coscelli, A., & Shum, M. (2003). An empirical model of learning and patient spillovers in new drug entry. *Journal of Econometrics*, *122*, 213–246.
Cosman, J. (2014). Industry dynamics and the value of variety in nightlife: Evidence from Chicago. *Working paper*.
Couture, V., & Handbury, J. (2015). Urban revival in America, 2000 to 2010. *Working paper*.
Crawford, G., & Shum, M. (2005). Uncertainty and learning in pharmaceutical demand. *Econometrica*, *73*, 1137–1173.
Datta, S., & Sudhir, K. (2011). The agglomeration-differentiation tradeoff in spatial location choice. *Working paper*.
DeGroot, M.-H. (1970). *Optimal statistical decisions*. New York: McGraw-Hill.
Dickstein, M. (2015). Efficient provision of experience goods: Evidence from antidepressant choice. *Working paper*.
Ellison, G., Glaeser, E., & Kerr, W. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, *100*, 1195–1213.
Eppli, M., & Benjamin, J. (1994). The evolution of shopping center research: A review and analysis. *Journal of Real Estate Research*, *9*, 5–32.
Erdem, T., & Keane, M. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, *15*, 1–20.
Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, *62*, 53–82.
Kung, E.Fang, H. (2010). Why do life insurance policyholders lapse? The roles of income, health and bequest motive shocks. *NBER working paper no. 17899*.
Gale, D. (1996). What have we learned from social learning? *European Economic Review*, *40*, 617–628.
Hollenbeck, B. (2017). The economic advantages of chain organization. *RAND Journal of Economics*, *48*, 1103–1135.
Holmes, T. (2011). The diffusion of Wal-Mart and the economies of density. *Econometrica*, *79*, 253–302.
Igami, M., & Yang, N. (2016). Unobserved heterogeneity in dynamic games: Cannibalization and preemptive entry of hamburger chains in Canada. *Quantitative Economics*, *7*, 483–521.
Jardim, E. (2015). All in the mix: Spillovers and the agglomeration of neighborhood retail. *Working paper*.
Kapoor, S., & Magesan, A. (2014). Paging inspector sands: The cost of public information. *American Economic Journal: Economic Policy*, *6*, 92–113.
Kasahara, H., & Shimotsu, K. (2009). Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica*, *77*, 135–175.
Konishi, H. (2005). Concentration of competing retail stores. *Journal of Urban Economics*, *58*(3), 488–512.
Kosová, R., & Lafontaine, L. (2010). Survival and growth in retail and service industries: Evidence from franchised chains. *The Journal of Industrial Economics*, *58*, 542–578.
Love, J. (1995). *McDonald's: Behind the arches*. New York: Bantam Books.
Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*, 531–542.
Narayanan, S., Chintagunta, P., & Miravete, E. (2007). The role of self selection, usage uncertainty and learning in the demand for local telephone service. *Quantitative Marketing and Economics*, *5*, 1–34.
Narayanan, S., & Manchanda, P. (2009). Heterogeneous learning and the targeting of marketing communication for new products. *Marketing Science*, *28*, 424–441.
Nishida, M. (2015). Estimating a model of strategic network choice: The convenience-store industry in Okinawa. *Marketing Science*, *34*, 20–38.
Orhun, Y. (2013). Spatial differentiation in the supermarket industry: The role of common information. *Quantitative Marketing and Economics*, *11*, 3–37.
Osborne, M. (2011). Consumer learning, switching costs, and heterogeneity: A structural examination. *Quantitative Marketing and Economics*, *9*, 25–70.
Pakes, A. (1986). Patents as options: Some estimates of the value of holding European patent stocks. *Econometrica*, *54*, 755–784.
Schragger, R. (2001). The anti-chain store movement, localist ideology, and the remnants of the progressive constitution, 1920–1940. *Iowa Law Review*, *90*, 1011.
Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics*, *37*, 619–640.
Shen, Q., & Xiao, P. (2014). McDonald's and KFC in China: Competitors or companions. *Marketing Science*, *33*, 287–308.
Suzuki, J. (2013). Land use regulation as a barrier to entry: Entry from the Texas lodging industry. *International Economic Review*, *54*, 495–523.
Thomadsen, R. (2012). Seeking an aggressive competitor: How product line expansion can increase all firms' profits. *Journal of Marketing Research*, *49*, 349–360.
Toivanen, O., & Waterson, M. (2005). Market structure and entry: Where's the beef? *RAND Journal of Economics*, *36*, 680–699.
Vitorino, M.-A. (2012). Empirical entry games with complementarities: An application to the shopping center industry. *Journal of Marketing Research*, *49*, 175–191.
Yang, N. (2012). Burger King and McDonald's: Where's the spillover? *International Journal of the Economics of Business*, *19*, 255–281.
Zhu, T., Singh, V., & Dukes, A. (2011). Local competition, entry, and agglomeration. *Quantitative Marketing and Economics*, *9*, 129–154.