Dynamic Spillovers from the Rival: Evidence from the Entry and Expansion of KFC and McDonald's in the Chinese Cities[1]

Aamir Rafique HASHMI
Department of Economics
National University of Singapore

Ping XIAO
Department of Marketing
NUS Business School
National University of Singapore

October 23, 2015

[1] This is a thoroughly revised version of an earlier paper that circulated under the title "Entry- and Sunk-Cost Spillovers from the Rival: Evidence from Entry and Expansion of KFC and McDonald's in Chinese Cities."

Abstract

We specify and estimate a dynamic game to understand the entry and expansion of KFC and McDonald's in 246 Chinese cities from 1990 to 2007. Our main research question is how the rival's presence affects a chain's entry and expansion decisions. The structural estimates and numerical value functions enable us to quantify the spillovers from the rival and provide further insights into the entry and expansion behavior of the chains. We find that the rival's presence in the market has a positive spillover effect on the cost of entry or expansion. However, it also has a negative effect on the per outlet profit. The net effect differs across the two chains. We find that each additional McDonald's outlet in the city adds about ¥3.8 million to KFC's value whereas an additional KFC outlet subtracts about ¥4.6 million from McDonald's value.

Keywords: Dynamic games; Dynamic entry; Entry Games; Spillovers; Fastfood industry; China

JEL Codes: L13; L81; M31

1 Introduction

The decisions about market selection and expansion are at the heart of the growth strategy of any business.[1] When we think about market selection, a number of questions naturally arise: How do firms choose which markets to enter? How do firms expand after entering a new market? How does the market structure affect these decisions? Are there any demand or cost spillovers from the rival(s)? Economic theory does not provide precise answers to these questions because the answers depend on a number of factors including the nature of competition in the industry, unique characteristics of each market and other institutional factors. Hence we need industry-specific empirical models to answer these questions.

In this paper, we study the empirical relationship between market structure and the entry and expansion decisions of retail chains.[2] 'Entry' refers to the decision of opening the first outlet in a market and 'expansion' refers to the decision of adding more outlets to a market conditional on entry. Our empirical setting is the entry and expansion of two major Western fastfood chains, Kentucky Fried Chicken (KFC) and McDonald's, into various Chinese cities.[3] In November 1987 KFC opened its first outlet in Beijing and in October 1990 McDonald's opened its first outlet in Shenzhen. The two chains have dramatically increased their presence in the Chinese cities since then. By the end of 2007, KFC had close to 2000 outlets in 230 cities and McDonald's had more than 1000 outlets in about 140 cities in China. The rapid expansion of the two chains into numerous Chinese cities within the last couple of decades provides an excellent opportunity to study the driving forces behind their entry and expansion decisions and how the two chains influence each other in a strategic environment.

The market entry and expansion decisions of a chain in each period are likely to be affected by a variety of factors. First, such decisions are guided by the chain's overall expansion strategy in the country and the market characteristics including the government policy implementation on foreign investment. Second, they may depend on the number of existing outlets that the chain owns. A greater presence in a city may suggest lower set-up cost of additional outlets in the same market due to the economies of scale. However, it also increases the potential problem of within chain cannibalization. Third, a chain's decision to enter a new market or to increase the number of outlets in an existing market may be affected by the rival's presence in a number of ways. On the one hand, a stronger presence by the rival increases competition, which can discourage the entry or expansion of the focal chain. On the other hand, the establishment of the rival chain may create demand or cost spillovers that may benefit the focal chain to enter or expand in the same market.

To quantify such spillover effects, we estimate a dynamic structural model of entry. In the model the two chains simultaneously decide whether to enter a market (if they are not already in) or how many new outlets to open in the market based on the factors discussed above and their expectations about the future evolution of the market structure. Our primary interest is to understand how the rival's presence in a market affects a chain's entry and expansion decisions. The dynamic

[1] In a recent video interview, McKinsey & Company's Chris Bradly (a principal at McKinsey's Sydney office) said: "If you look at what drives the growth of companies, it turns out that selection at a micromarket level is much more important than trying to gain market share. In fact, 80 percent of growth is explained by decisions about where to compete or by market selection." [The video interview is available at `http://www.mckinsey.com/insights/strategy/the_art_of_strategy` (last accessed: November 2, 2013)]

[2] We define market structure broadly to include a chain's own outlets, its rivals' outlets and some other market-specific characteristics.

[3] We define market to be a city and use the words "market" and "city" interchangeably. In Section 2 we justify our choice of treating each city as an independent market.

considerations are important in our empirical setting. For example, a chain's entry into a new market affects the market structure which may lead to a different set of future actions by the chain and its rival and hence different future outcomes. Hence the spillovers from the rival are *dynamic* as they sum up the current and future effects of an additional rival outlet on a chain's value.

In this paper we use the method proposed by Bajari et al. (2007) (from here on, BBL) to estimate the model. A key element of BBL's method in the context of our study is to assume that the observed entry and expansion decisions are equilibrium outcomes. This allows us to recover the structural parameters of the game without solving for equilibrium. We find that a greater rival presence lowers the outlet set-up costs for both chains and significantly lowers the entry cost for McDonald's. However, it lowers per outlet profits for both chains. The net effect differs across the two chains. Our estimates show that each additional McDonald's outlet in the city adds about ¥3.8 million to KFC's value whereas an additional KFC outlet subtracts about ¥4.6 million from McDonald's value.

1.1 Related Literature

This paper connects to a number of literatures. First, it is closely related to the literature on empirical models of entry [Bresnahan and Reiss (1991), Berry (1992) and Berry and Reiss (2007)]. Mazzeo (2002), Seim (2006) and Zhu et al. (2009) extend the framework to endogenously consider product differentiation along with entry decisions in static game-theoretic environments. We contribute to this literature by using a dynamic-game framework to study the relationship between market structure and the firms' entry and expansion decisions.

Second, and most importantly, it is related to the empirical literature on retail chain entry and expansion. This literature can be classified into three groups depending on whether the underlying model is static or dynamic and whether it allows for strategic interactions among the firms.

The first group includes papers that model entry as a static game [Toivanen and Waterson (2005), Jia (2008) and Nishida (2013)]. Toivanen and Waterson (2005) allow for a positive effect from the rival in a chain's market entry decision using data from the United Kingdom's fastfood industry. They model the duoploy between Burger King and McDonald's as a static sequential game. Their structural estimates suggest that a rival's presence in the market leads a firm to expect a larger market. They argue that the spillover effect mainly operates through learning. Jia (2008) studies the positive spillover effects when locating multiple stores in nearby regions in the context of Wal-Mart's market entry decisions. Her modelling strategy involves a static three-stage game. Nishida (2013) extends Jia's approach to consider the decision of how many stores to open using data from the convenience-store industry of Japan, but constrains the spillover effect to be positive as in Jia (2008). The aforementioned papers build on game-theoretic models of static entry.

The second group consists of papers that employ single-agent dynamic models to study a chain's location and expansion decisions [Holmes (2011) and Toivanen and Waterson (2011)]. Holmes (2011) estimates the economies of density in the store location decisions of Wal-Mart in the United States. He considers cannibalization of sales by nearby stores of the same chain but does not model competition effect from other retail chains. He finds that the economies of density are substantial and extend beyond the savings in trucking costs. Toivanen and Waterson (2011) study the expansion of McDonald's in the United Kingdom up to 1990, a period when it can reasonably be considered a monopoly in the UK market. They find positive cannibalization effect from own outlets on the demand side but economies of density on the cost side.

The third group includes papers that model entry and expansion as dynamic games [Aguirregabiria and Magesan (2012) and Yang (2012)]. Aguirregabiria and Magesan (2012) use the dataset in Toivanen and Waterson (2005). Their main empirical contribution is to show that when a chain's beliefs about its rival's strategy are biased, the chain runs the risk of underestimating the potential competition effects from the rival after its own entry. Yang (2012) studies the effect of information spillovers from incumbent rivals on a chain's decision to enter a new market. Prior to entry, potential entrants are uncertain about the size of the market that they are about to enter. The past entry and exit decisions of incumbent rivals reduce this uncertainty and allow the potential entrants to make more informed entry decisions. The present paper is closest to Yang (2012) as both papers estimate dynamic games to quantify the spillovers from rivals. However, this paper differs from his in a number of ways. First, we focus on entry and expansion decisions of the chains while his focus is on binary entry and exit decisions. Second, we allow for multiple spillover effects within the chain and across the chains while he focuses solely on information spillovers from rivals. Third, our empirical setting is the fastfood market in China while he studies the Canadian fastfood market.

Finally, this paper provides yet another application of the recent methods to estimate dynamic games. The leading methodology papers in this literature are Aguirregabiria and Mira (2007), Bajari et al. (2007), Pakes et al. (2007) and Pesendorfer and Schmidt-Dengler (2008). Ackerberg et al. (2007) provide a comparison of these methods.

We have organized the rest of the paper as follows. In Section 2 we provide an overview of the Chinese fastfood industry and introduce our dataset. In Section 3 we present our model. In Section 4 we discuss our estimation strategy. We report and discuss the estimation results in Section 5 and conclude in Section 6.

2 Preliminaries

In this section we provide an overview of the history of Western chains in the Chinese fastfood market and introduce our dataset.

2.1 Industry Overview

The history of the Western fastfood in China began in November 1987 when KFC opened its first outlet in Beijing. Three years later, in October 1990, McDonald's opened its first outlet in Shenzhen and the competition between the two formally began in the world's largest emerging market.

In Figure 1 we plot the total number of outlets for the two chains from 1990 to 2007. Although globally McDonald's is far ahead of KFC in terms of the number of outlets, their positions are reversed in the Chinese market. Not only did KFC enter the country first, its expansion has also been much faster than that of McDonald's. By the end of 2007, KFC had close to 2000 outlets in China whereas McDonald's had slightly more than 1000.

In Table 1 we report, for selected years, the total number of outlets together with the number of cities entered by each chain. The first thing to note is that although KFC has been ahead of McDonald's throughout the period, it has extended its lead very sharply since 2000. The second thing to note is that most of the KFC's lead is due to its entry in more cities. By 2007, McDonald's had entered 142 cities (out of the total 246 in our sample) whereas KFC was present in 230 cities, a lead of 62%. If we compare within city density of outlets, McDonald's had 7.1 outlets per city while KFC had 8.4, a lead of 18%. The third thing to note is that the scale of entry and expansion varies over the sample period. This variation is helpful to identify the parameters of interest later.

Figure 1: Total outlets in China (by the end of the year)

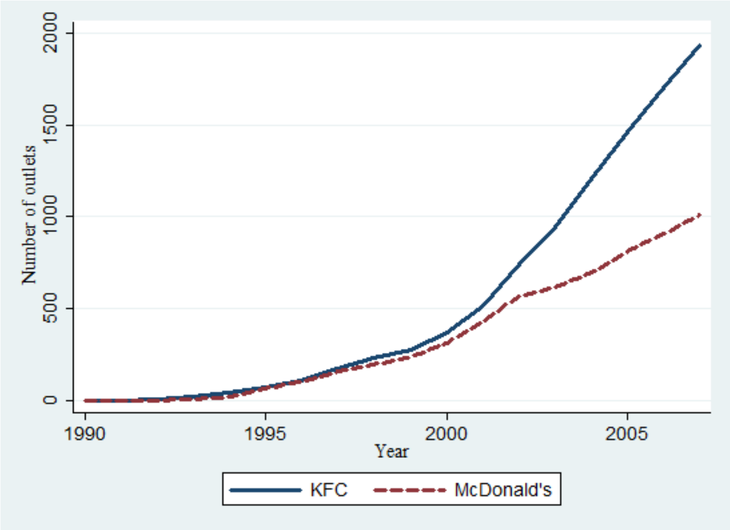

Table 1: Number of cities entered and outlets opened (by the end of the year)

	KFC			McDonald's		
Year	Cities	Outlets	Outlets/City	Cities	Outlets	Outlets/City
1990	2	4	2.0	1	1	1.0
1995	21	75	3.6	11	69	6.3
2000	72	370	5.1	56	315	5.6
2005	198	1462	7.4	117	812	6.9
2007	230	1932	8.4	142	1012	7.1

Table 2, in which we report statistics about the first mover in each city, provides further evidence that KFC has been much more aggressive than McDonald's in entering new cities. By the end of 2007, at least one of the two chains was present in 236 cities. KFC was the first mover in 177 (75%) of these cities and McDonald's in 33 (14%). In the remaining 26 (11%) cities, both chains entered in the same calendar year. On the average, McDonald's waited 3.5 years to enter a city where KFC was already present.

We model the entry and expansion of the Western fastfood chains in the Chinese market over the 1990-2007 period as a duopoly game between KFC and McDonald's.[4] Although lately some other Western fastfood chains have entered the Chinese market, their presence is still modest by the end of our sample period.[5] Similarly, none of the local Chinese fastfood chains is large enough

[4]It is a popular perception that KFC and McDonald's are close rivals in the Chinese market. For example, The Wall Street Journal Asia reported on February 29, 2012: "McDonald's Corp. is launching a new ad campaign in China ... to win market share from its dominant rival, Yum Brands Inc.'s KFC." [p. 19]

[5]For example, Subway entered the Chinese market in 1995 but had only around 150 outlets by the end of 2007.

to be considered a strategic player in the national fastfood market. Also, the Western fastfood is quite distinct from the traditional Chinese fastfood and hence the two are not likely to be close substitutes.

Table 2: First mover

First Mover	No. of Cities	(%)
KFC	177	(75.0)
McDonald's	33	(14.0)
Both	26	(11.0)
Total	236	(100.0)

We do not distinguish between the franchised and non-franchised outlets mainly due to the following two reasons. First, during our sample period the franchised outlets accounted for a small percentage of the total for both chains.[6] Second, our focus in this paper is on entry and expansion decisions. These decisions are made at the national level by the top management teams of the chains and the franchisees have little say in them.

2.2 Data

Our basic unit of analysis is a city. Our sample consists of 246 Chinese cities spanning an eighteen-year period from 1990 to 2007.[7] We track entry and expansion of KFC and McDonald's in these cities for the entire sample period.[8] We have also gathered information on important city characteristics. In the rest of this sub-section we highlight some important features of the dataset.

In Table 3 we show the frequency distribution of the number of outlets at the end of the sample period. There are 16 cities (6.5% of total) in the sample in which KFC has yet to enter and 104 cities (42.3%) in which McDonald's has yet to enter. The number of cities with 1 or 2 KFC outlets is 131 (53.3%) and with 1 or 2 McDonald's outlets it is 89 (36.2%). There are also a number of cities with a sizable presence by both chains. For example, there are 39 cities (15.8%) with 11 or more KFC outlets and 23 cities (9.4%) with 11 or more McDonald's outlets.

In the upper panel of Table 4 we show the frequency distribution of the number of outlets for the entire sample. This information is important because the number of outlets in a city is a key state variable in our model. There are many zeros in the data. Out of a total number of 4,428 observations (246 cities *times* 18 years), there are 3,168 (71.5%) observations with zero outlets for KFC and 3,638 (82.2%) observations with zero outlets for McDonald's. The number of observations

Burger King entered in 2005 and had fewer than 50 outlets by the end of 2007.

[6]KFC opened its first franchised outlet in 2000 and by the end of 2005 it had only 37 franchised outlets, which were 2.5% of its total stock of 1462 outlets. (`http://franchise.business-opportunities.biz/2006/04/21/kfc-makes-it-easier-to-buy-a-franchise` [last accessed: October 7, 2013]) McDonald's launched a pilot franchise program in 2004 and had only 6 franchised outlets by early 2010. (`http://www.reuters.com/article/2010/05/06/us-mcdonalds-china-idUSTRE6451W420100506` [last accessed: March 1, 2013])

[7]The 246 cities are identified according to the *Chinese Cities Yearbook*. According to the World Bank, total population of China was 1,318 million in 2007. Out of these, 45% or 593.1 million people lived in urban areas. Total population of the 246 cities in our sample was 336.5 million in 2007. This amounts to 56.7% of the total urban population in China in that year.

[8]For details on collection and construction of the dataset see Shen and Xiao (2014).

with 1 or 2 outlets is 755 (17.1%) for KFC and 466 (10.5%) for McDonald's. The number of observations with 11 or more outlets is just 164 (3.7%) for KFC and 116 (2.6%) for McDonald's.

Table 3: Frequency distribution of the number of outlets in a city (end of 2007)

No. of Outlets	KFC No. of Cities	KFC (%)	McDonald's No. of Cities	McDonald's (%)
0	16	(6.5)	104	(42.3)
1	85	(34.6)	64	(26.0)
2	46	(18.7)	25	(10.2)
3 to 5	38	(15.4)	21	(8.5)
6 to 10	22	(8.9)	9	(3.7)
11 to 20	19	(7.7)	13	(5.3)
21 to 100	18	(7.3)	9	(3.7)
> 100	2	(0.8)	1	(0.4)
Total	246	(99.9)§	246	(100.1)§

§The sum does not add up to 100 due to rounding.

Table 4: Frequency distributions of total and new outlets (Full sample: 246 cities, 18 years)

Number of Total Outlets (at the start of the year)	0	1	2	3-5	6-10	11-20	21-100	>100	Total
KFC: Frequency	$3,168$	535	220	207	134	94	61	9	$4,428$
(Percent)	(71.5)	(12.1)	(5.0)	(4.7)	(3.0)	(2.1)	(1.4)	(0.2)	(100.0)
McDonald's: Frequency	$3,638$	364	102	128	80	62	50	4	$4,428$
(Percent)	(82.2)	(8.2)	(2.3)	(2.9)	(1.8)	(1.4)	(1.1)	(0.1)	(100.0)
Number of New Outlets (opened during the year)	**0**	**1**	**2**	**3-5**	**6-10**	**11-20**	**21-30**	**>30**	**Total**
KFC: Frequency	$3,643$	454	140	128	37	18	7	1	$4,428$
(Percent)	(82.3)	(10.3)	(3.2)	(2.9)	(0.8)	(0.4)	(0.2)	(0.0)	(100.1)§
McDonald's: Frequency	$4,018$	225	70	75	30	9	1	0	$4,428$
(Percent)	(90.7)	(5.1)	(1.6)	(1.7)	(0.7)	(0.2)	(0.0)	(0)	(100.0)

§The sum does not add up to 100 due to rounding.

In the lower panel of Table 4 we show the frequency distribution of the new outlets opened during a year. This is also important because the number of new outlets is a key decision variable in our model. Here again we see lots of zeros. There are $3,643$ (82.3%) observations with zero new outlets for KFC and $4,018$ (90.7%) such observations for McDonald's. When a chain does open new outlets, most of the time it opens either 1 or 2 outlets. There are 594 (13.5%) observations with 1 or 2 new outlets for KFC and 295 (6.7%) such observations for McDonald's. The rest of the observations (191 or 4.3% for KFC and 115 or 2.6% for McDonald's) are with 3 or more new outlets.

In Table 5 we report summary statistics for a number of variables. They include both chain-specific and city-specific variables. The first chain-specific variable is the total number of outlets of a chain in a city at the beginning of a calendar year. The average number of outlets is 1.8 for KFC and 1.2 for McDonald's. As we saw in Table 4, there is a lot of variation in the number of outlets of both chains. This is reflected in high standard deviations (SDs) of the number of outlets: 8.8 for KFC and 6.3 for McDonald's.

Table 5: Summary statistics

Variable (units)	Obs.	Mean	S.D	p_{10}	p_{90}
Chain Characteristics (KFC)					
Total outlets (number at the start of the period)	4428	1.8	8.8	0	3
New outlets (number per period)	4428	0.4	1.7	0	1
Distance from own headquarters (thousand km)	4428	1.0	0.5	0.4	1.7
Chain Characteristics (McDonald's)					
Total outlets (number at the start of the period)	4428	1.2	6.3	0	1
New outlets (number per period)	4428	0.2	1.1	0	0
Distance from own headquarters (thousand km)	4428	1.2	0.7	0.4	2.1
City Characteristics					
Population (millions)	4263	1.1	1.3	0.3	1.9
GDP p.c. (2001 prices, thousand yuans)	4042	14.8	15.8	3.7	29.3
Provincial capital city (dummy)	4428	0.12	0.32	0	1
z_1 (Open city 1993, dummy)	4428	0.07	0.26	0	0
z_2 (Open city 2001, dummy)	4428	0.08	0.27	0	0

The second chain-specific variable is the number of new outlets opened in a city during a calendar year. The average number of new outlets is 0.4 for KFC (with a SD of 1.7) and 0.2 for McDonald's (with a SD of 1.1). A notable feature of this variable in the data is the high frequency of zeros. Another notable feature is that the SD is many times higher than the mean. Both these features will influence our choice of the econometric model in the first stage of estimation below.

The third chain-specific variable is the distance from own headquarters. KFC's headquarters were in Shanghai throughout the sample period. McDonald's headquarters were in Hong Kong until 2003 and have been in Shanghai since then. The average distance of a city from the chain headquarters is one thousand kilometers for KFC and 1.2 thousand kilometers for McDonald's.

The last five variables in Table 5 are city-specific. The first one is population. The average population of a city in our sample is 1.1 million with a SD of 1.3 million. The second is the per capita gross domestic product (GDP p.c.). The average GDP p.c. (in 2001 prices) is ¥14.8 thousand with a SD of ¥15.8 thousand. The third is a dummy variable for provincial capital cities. There are 29 such cities in our sample and they are, on the average, four times larger (in terms of population) than other cities. The last two city-specific variables, z_1 and z_2, are policy dummies. In late 1992 the State Council, the chief administrative authority in China, issued a notice called "open policy" to support 22 cities in attracting foreign investment. Our dummy variable z_1 takes a value of 1 for these cities from 1993 onwards. Then in 2000, the central goverment announced its "promote-west" policy for a large-scale development of the underdeveloped western inland cities. This policy covered 51 cities in our sample. The dummy variable z_2 takes a value of 1 for these cities from 2001 onwards.

In addition to the above, we have also collected data on city-level operating income of the two chains. These data are available from 2003 to 2007 for an unbalanced panel of cities. We provide more details on these data in Appendix B. The average annual operating income *per city* for KFC is ¥328.57 million (with a SD of ¥489.83 million) and for McDonald's it is ¥83.04 million (with a SD of ¥141.36 million). The average annual income *per outlet* is ¥17.91 million (SD=¥20.29 million) for KFC and ¥9.26 million (SD=¥5.44 million) for McDonald's. We use this information for normalizations in stage two of the estimation (see Section 5.2).

3 Model

We model the entry and expansion of KFC and McDonald's into various Chinese cities as a dynamic duopoly game.[9] We restrict ourselves to a duopoly because, as we discussed in Section 2.1, KFC and McDonald's did not face any significant competition from other Western fastfood chains during the sample period. The dynamic considerations are important as today's entry and expansion decisions may have non-trivial implications for future actions of either chain. By ignoring the dynamics one may draw incorrect inferences about the impact of the underlying market structure on the observed actions.

We model the chains' decisions at the city level. This means that the chains decide about entry into and expansion within a city by looking at the state (to be defined below) of the city. This assumption is reasonable because the cities in our sample are generally hundreds of kilometers away from each other and hence can reasonably be treated as independent markets.[10] A problem with this assumption is that the fastfood chains generally decide their expansion policies at the national level and then look for suitable markets to implement the policies. For example, KFC is likely to first decide the total number of outlets that it wants to open in China in a given year and then look for suitable cities to achieve its target. We take care of this problem by introducing the age (or experience) of a chain in the country as a state variable. The idea is that the national level expansion of a chain can reasonably be approximated–at least within the sample–by fitting a time trend to its total number of outlets in the country. For example, simple quadratic trend lines can approximate the expansion of KFC and McDonald's (see Figure 1) in China very well.

In our model, time is discrete and planning horizon is infinite. The length of a period is one year. There are two fastfood chains, i and j. A chain can either be 'in' the market or 'out' of the market. If it is in the market, it has to decide how many outlets to open in the market in the current period. We call this the *expansion* decision. If it is out of the market, it has to decide whether to enter the market or stay out of it in the current period. We call this the *entry* decision. These decisions depend on the expected profit of opening outlets as well as the cost of entering a new market and the set-up costs of additional outlets.[11] In the following, we model the problem of chain i. Chain j solves a similar problem.

[9] Our model does not feature exit from a city or a reduction in the number of outlets within a city. This is because we do not observe exit or contraction in our sample. In our original dataset, we do not have complete information on 39 KFC and 24 McDonald's outlets. The outlets with missing information may potentially include exits. However since these account for a small fraction of the total number of outlets in the data and also there is no particular pattern for the missing data, we exclude them from our sample.

[10] In our sample, the average distance between a focal city and the nearest city with at least one own outlet is 327 kilometers for KFC (with a SD of 442 kilometers) and 647 kilometers for McDonald's (with a SD of 1,095 kilometers).

[11] In our model entry and set-up costs are different. Entry cost is paid when a chain opens its first outlet in the city. It includes the set-up cost of the first outlet but also includes all other direct and indirect economic costs that may be associated with the setting up of a business in a new city. The set-up costs are incurred when a chain expands within the city and opens more outlets.

3.1 State and Choice Variables

The state vector facing chain i consists of the following 10 variables:

1. $n_{i,c} \in \{0, 1, \ldots, \bar{n}\}$, the number of own outlets in city c at the beginning of the period. If $n_{i,c} = 0$, the chain does not have any outlet in the city and hence it is 'out' of the city. When $n_{i,c} > 0$, the chain is 'in' the city. To bound the state space we impose an upper limit, $\bar{n}$, on the number of outlets in a city. This can be thought of as the saturation point for the chain in the city.[12]

2. $n_{j,c} \in \{0, 1, \ldots, \bar{n}\}$, the number of rival's outlets in the city at the beginning of the period.

3. $a_i \in \{0, 1, \ldots, \bar{a}\}$, age of the chain in the country. We set $\bar{a} = 50$. It means whatever effect a chain's experience has on its entry and expansion decisions, it disappears after 50 years. We need this assumption to bound the state space.

4. $d_{i,c}$, distance of the city from the chain's own headquarters. This distance is city specific and does not change over time.[13]

5. p_c, population of the city. We assume it to follow a first-order Markov process.

6. y_c, per capita gross domestic product (GDP p.c.) of the city. This too is modelled as a first-order Markov process.

7. k_c, provincial-capital dummy.

8. $z_{1,c}$, open city 1993 (dummy). See Section 2.2, especially Table 5, for details on this and the next two state variables.

9. $z_{2,c}$, open city 2001 (dummy).

10. ε_i, a privately observed random shock to the cost of entry (if the chain is out of the city, i.e. $n_{i,c} = 0$) or to the set-up cost of a new outlet (if the chain is in the city, i.e. $n_{i,c} > 0$). In the former case we denote it by ε_i^e and in the latter case by ε_i^s. ε_i^e and ε_i^s come from different distributions. Each of these shocks is independently and identically distributed (i.i.d.). We assume these shocks to follow Gaussian distributions with zero mean and $\{\sigma^e, \sigma^s\}$ standard deviations. We estimate $\{\sigma^e, \sigma^s\}$ together with other structural parameters of the model.

We denote the state vector facing chain i of city c by $\mathbf{s}_{i,c} \in S$, where S is the state space:

$$\mathbf{s}_{i,c} = \{n_{i,c}, n_{j,c}, a_i, d_{i,c}, p_c, y_c, z_{1,c}, z_{2,c}, k_c, \varepsilon_i\}. \tag{1}$$

[12] When we forward simulate the model for estimation purposes we set $\bar{n} = 500$. This number is reasonable because in our dataset the maximum number of KFC outlets is 191 (in Shanghai in 2007) and McDonald's outlets is 117 (in Beijing in 2007). Even in the large US cities, where McDonald's and KFC outlets are ubiquitous, the total number of outlets of these chains seldom exceeds 500. For example, the total number of McDonald's outlets is 546 in New York city, 478 in Los Angeles and 398 in Chicago. The corresponding numbers for KFC are: 233, 221 and 154. [Source: yellowpages.com, last accessed on March 7, 2013]

[13] Although McDonald's moved its headquarters from Hong Kong to Shanghai in 2003, we treat this as an unexpected move in the dynamic model. In other words, if the headquarters of a chain are in Hong Kong in the current period, both chains would expect them to remain in Hong Kong in the future. Likewise, if the headquarters are in Shanghai in the current period, they would be expected to remain in Shanghai.

When chain i is out of the city we denote its state by $\mathbf{s}_{i,c}^{\text{OUT}}$ and when it is in the city, we denote it by $\mathbf{s}_{i,c}^{\text{IN}}$. When $\mathbf{s}_{i,c} = \mathbf{s}_{i,c}^{\text{OUT}}$, it means $n_{i,c} = 0$ and $\varepsilon_i = \varepsilon_i^e$. On the other hand, when $\mathbf{s}_{i,c} = \mathbf{s}_{i,c}^{\text{IN}}$, it means $n_{i,c} > 0$ and $\varepsilon_i = \varepsilon_i^s$. $\mathbf{s}_{i,c}^{\text{OUT}}$ and $\mathbf{s}_{i,c}^{\text{IN}}$ are mutually exclusive: a chain can either be in $\mathbf{s}_{i,c}^{\text{OUT}}$ or in $\mathbf{s}_{i,c}^{\text{IN}}$. Once a chain is in $\mathbf{s}_{i,c}^{\text{IN}}$ it cannot go back to $\mathbf{s}_{i,c}^{\text{OUT}}$ as there is no exit in our model.

When a chain is out of the city, i.e. it is in state $\mathbf{s}_{i,c}^{\text{OUT}}$, its choice variable is $e_i \in \{0, 1\}$. If the chain decides to enter the city in the current period, $e_i = 1$, otherwise, $e_i = 0$.

When the chain is in the city, i.e. it is in state $\mathbf{s}_{i,c}^{\text{IN}}$, its choice variable is $x_i \in \{0, 1, \ldots, \bar{x}\}$, where $\bar{x}$ is the maximum number of outlets that a chain can open in a city in one period.

Our motivation for including so many market-specific variables is to control as much for the market heterogeneity as possible.

3.2 Timing of Events

The timing of events for chain i within a period depends on whether the chain is out of or in the city. If the chain is out of the city, the timing of events is as follows. At the start of the period the chain observes its state vector $\mathbf{s}_{i,c}^{\text{OUT}}$. Given $\mathbf{s}_{i,c}^{\text{OUT}}$, and its beliefs about the entry and expansion strategies of its rival $\eta_j(\cdot)$ and $\chi_j(\cdot)$, the chain decides on $e_i \in \{0, 1\}$, i.e. whether to enter the city or stay out. If it decides to enter the city, it pays the cost of entry, $c_i^e\left(\mathbf{s}_{i,c}^{\text{OUT}}\right)$ and opens one outlet that will start operations from the start of the next period. If it decides to stay out, it does nothing and waits for the next period, when it will face the same choice with an updated $\mathbf{s}_{i,c}^{\text{OUT}}$.

If the chain is already in the city, at the start of the period it observes its state vector $\mathbf{s}_{i,c}^{\text{IN}}$. Given $\mathbf{s}_{i,c}^{\text{IN}}$ and its beliefs about the strategies of its rival, the chain decides on $x_i \in \{0, 1, \ldots, \bar{x}\}$, i.e. the number of new outlets to open in the city during the period. It then pays, if applicable, the sunk cost, $c_i^s\left(\mathbf{s}_{i,c}^{\text{IN}}\right)$, for each outlet that it decides to open. Chain j makes similar decisions simultaneously depending on its own state. All new outlets are opened for business at the end of the period and the state vector is updated. We discuss the state transitions in more detail in Section 3.5.

3.3 The Entry Problem

We are now ready to formally write the entry problem of a chain as a dynamic program.

$$\begin{aligned} V_i^{\text{OUT}}\left(\mathbf{s}_{i,c}^{\text{OUT}} | \boldsymbol{\theta}_i, \eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right), \chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)\right) \quad = \quad & \max_{e_i \in \{0,1\}} \Big\{ e_i \left[\beta E\left(V_i^{\text{IN}}\left(\mathbf{s}_{i,c}^{\prime\text{IN}}\right)\right) - c_i^e\left(\boldsymbol{\theta}_i^e, \mathbf{s}_{i,c}^{\text{OUT}}\right)\right], \\ & (1 - e_i)\, \beta E\left(V_i^{\text{OUT}}\left(\mathbf{s}_{i,c}^{\prime\text{OUT}} | \boldsymbol{\theta}_i, \eta_j\left(\mathbf{s}_{j,c}^{\prime\text{OUT}}\right), \chi_j\left(\mathbf{s}_{j,c}^{\prime\text{IN}}\right)\right)\right) \Big\}. \end{aligned} \tag{2}$$

The value function of chain i, when it is out of the city, is $V_i^{\text{OUT}}\left(\mathbf{s}_{i,c}^{\text{OUT}}\right)$. Note that the value function has i in the subscript. It implies that the two chains can potentially have different value functions. This is motivated by the reduced-form regression results that we report later in Section 5.1. We take the reduced-form results as the preliminary evidence to show that the profits and the costs of entry and expansion could potentially be asymmetric across chains and give rise to different value functions. For the same reason, functions $\pi_i(\cdot)$, $c_i^e(\cdot)$ and $c_i^s(\cdot)$ are also subscripted by i.

The value depends on the state vector $\mathbf{s}_{i,c}^{\text{OUT}}$ conditional on $\boldsymbol{\theta}_i, \eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right)$ and $\chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)$ being given. The vector $\boldsymbol{\theta}_i = \{\beta, \boldsymbol{\theta}_i^\pi, \boldsymbol{\theta}_i^e, \boldsymbol{\theta}_i^s\}$ contains all the parameters of the model that we aim to estimate for chain i. The strategy functions $\eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right)$ and $\chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)$ represent the rival strategies that chain i takes as given. Which of the rival's two strategy functions is relevant for the chain in the current period depends on whether the rival is out of or in the city. In the former case, $\eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right)$ is relevant and in the latter case $\chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)$ is relevant.

The chain decides about entry by comparing the expected net benefit of entry with the expected value of staying out for another period. The expected net benefit of entry consists of two terms. The first term, $\beta E\left(V_i^{\text{IN}}\left(\mathbf{s}_{i,c}^{\prime\text{IN}}\right)\right)$, is the discounted expected benefit of being in the city with one outlet at the start of the next period.[14] The expectation is over the future values of $n_{j,c}, p_c, y_c$ and own sunk-cost shock. We define $V_i^{\text{IN}}(\cdot)$ explicity below when we describe the expansion problem of the chain. The second term, $c_i^e\left(\boldsymbol{\theta}_i^e, \mathbf{s}_{i,c}^{\text{OUT}}\right)$, is the cost of entry that is to be paid in the current period.

The discounted expected value of staying out of the city for another period is $\beta E\left(V_i^{\text{OUT}}\left(\mathbf{s}_{i,c}^{\prime\text{OUT}}\right)\right)$. Here the expectation is over the future values of $n_{j,c}, p_c, y_c$ and own entry-cost shock.

The solution to the above problem is a strategy profile $\eta_i\left(\mathbf{s}_{i,c}^{\text{OUT}}\right)$ and optimal $e_i = \eta_i\left(\mathbf{s}_{i,c}^{\text{OUT}}\right)$.

3.4 The Expansion Problem

The expansion problem of chain i is to choose x_{ic}, the number of new outlets to open in the city during the period, to maximize the expected value of future profit streams. We can write this problem in the form of the following Bellman's equation:

$$\begin{aligned} V_i^{\text{IN}}\left(\mathbf{s}_{ic}^{\text{IN}}|\boldsymbol{\theta}_i, \eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right), \chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)\right) &= \max_{x_i \in \{0,1,\ldots,\bar{x}\}} \Big\{ \left[\pi_i\left(\mathbf{s}_{ic}^{\text{IN}}|\boldsymbol{\theta}_i^{\pi}\right) n_{ic} - c_i^s\left(\mathbf{s}_{ic}^{\text{IN}}|\boldsymbol{\theta}_i^s\right) x_{ic}\right] \\ &\quad + \beta EV_i\left(\mathbf{s}_{ic}^{\prime\text{IN}}|\boldsymbol{\theta}_i, \eta_j\left(\mathbf{s}_{j,c}^{\prime\text{OUT}}\right), \chi_j\left(\mathbf{s}_{j,c}^{\prime\text{IN}}\right)\right)\Big\}. \end{aligned} \tag{3}$$

This is the value of a chain that is in the city and facing state $\mathbf{s}_{ic}^{\text{IN}}$. Once again the value is conditional on $\boldsymbol{\theta}_i, \eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right)$ and $\chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)$. The period return consists of two terms. The first term, $\pi_i\left(\mathbf{s}_{ic}^{\text{IN}}|\boldsymbol{\theta}_i^{\pi}\right) n_{ic}$, is profit per outlet multiplied by the number of outlets. The second term, $c_i^s\left(\mathbf{s}_{ic}^{\text{IN}}|\boldsymbol{\theta}_i^s\right) x_{ic}$, is the set-up cost of a new outlet multiplied by the number of new outlets opened during the period. Parameter $\beta \in (0,1)$ is the discount factor. EV_i is the expected continuation value, where the expectation is over the evolution of $\mathbf{s}_{ic}$ and over the probability distributions of the rival's private shocks. The optimal solution to this problem is a strategy profile $\chi_i\left(\mathbf{s}_{ic}^{\text{IN}}|\boldsymbol{\theta}_i, \chi_j\right)$.

3.5 State Transitions

We now turn to the evolution of state variables in the model. The evolution of n_{ic} and n_{jc} is endogenous. Using a prime ($'$) to denote next period values

$$\begin{aligned} n'_{ic} &= n_{ic} + x_{ic}, \text{ and} \\ E\left(n'_{jc}\right) &= n_{jc} + E\left(x_{jc}\right). \end{aligned} \tag{4}$$

There is an expectation sign on x_{jc} because chain i does not observe the private cost shocks of its rival.

[14] We assume that when a chain enters a city, it opens only one outlet in the period of its entry. In our dataset, in about 14% of the cases (32 cities out of 230 for KFC and 20 cities out of 142 for McDonald's) a chain opens more than one outlet during the calendar year of its entry into a new city. Since the percentage of cases of entry with more than one outlet is small, for simplicity, we do not allow for this possibility in our model. This also helps in the separate identification of entry-cost and set-up-cost parameters.

Age, a_i, increases by one unit every period deterministically until it reaches the absorbing state $\bar{a}$. Once $a_i = \bar{a}$, it remains constant. We set $\bar{a} = 50$. Here the implicit assumption is that the additional benefits from longer experience dissipate after a chain has been in the country long enough.

Distance from headquarters, d_{ic}, remains constant over time and so do the policy dummies ($z_{1,c}$ and $z_{2,c}$) and the provincial-capital dummy k_c.

We assume population and GDP p.c. to follow first-order Markov processes that we estimate from the data.

The private shocks to the cost of entry (ε_i^e) and sunk cost of opening a new outlet (ε_i^s) are assumed i.i.d. over time, across cities and across the two chains. They are drawn from zero mean normal distributions with σ_i^e and σ_i^s standard deviations, respectively.

3.6 Equilibrium

We use the concept of Markov Perfect Equilibrium (MPE). The MPE of our model consists of the strategy profiles $\eta_i^* \left(\mathbf{s}_{i,c}^{\text{OUT}} | \boldsymbol{\theta}_i, \eta_j^*, \chi_j^*\right)$, $\eta_j^* \left(\mathbf{s}_{j,c}^{\text{OUT}} | \boldsymbol{\theta}_i, \eta_i^*, \chi_i^*\right)$, $\chi_i^* \left(\mathbf{s}_{i,c}^{\text{IN}} | \boldsymbol{\theta}_i, \eta_j^*, \chi_j^*\right)$ and $\chi_j^* \left(\mathbf{s}_{j,c}^{\text{IN}} | \boldsymbol{\theta}_j, \eta_i^*, \chi_i^*\right)$ such that $\eta_i^* \left(\mathbf{s}_{i,c}^{\text{OUT}} | \boldsymbol{\theta}_i, \eta_j^*, \chi_j^*\right)$ is an optimal solution to the problem in (2) and $\chi_i^* \left(\mathbf{s}_{i,c}^{\text{IN}} | \boldsymbol{\theta}_i, \eta_j^*, \chi_j^*\right)$ is an optimal solution to the problem in (3). Similarly $\eta_j^* \left(\mathbf{s}_{j,c}^{\text{OUT}} | \boldsymbol{\theta}_i, \eta_i^*, \chi_i^*\right)$ and $\chi_j^* \left(\mathbf{s}_{j,c}^{\text{IN}} | \boldsymbol{\theta}_j, \eta_i^*, \chi_i^*\right)$ are optimal solutions to similar problems for chain j.

3.7 Spillovers from the Rival

Suppose we know $\boldsymbol{\theta}_i$ and have numerically solved for the equilibrium of the model as defined in Section 3.6 above. This solution will give us functions V^{IN} and V^{OUT} for both chains defined over the entire state space. That means we can change a state exogenously and see how V^{IN} and V^{OUT} respond to the change. Now suppose that we compare V^{IN} for two states that differ only in the number of rival's outlets. This comparison will tell us how an exogenous change in the rival's presence will affect V^{IN} for the chain. To be specific, if an increase in the number of rival's outlets increases V^{IN} for the chain, we call it a positive spillover from the rival. In this case, more rival outlets add to the chain's value, holding everything else constant. Because V^{IN} and V^{OUT} are in numerical form, we can also quantify the size of the spillover.

The spillovers from the rival to a chain's V^{IN} and V^{OUT} represent the overall spillovers. With $\boldsymbol{\theta}_i$ in hand, we can also quantify the spillovers from the rival to various components of V^{IN} and V^{OUT}. Of special interest to us are the spillovers from the rival to the profit function, $\pi_i(\cdot)$, and the set-up-cost function, $c_i^s(\cdot)$, in (3) and the entry-cost function, $c_i^e(\cdot)$, in (2). The parameters associated with the number of the rival's outlets in these functions tell us how the rival's presence affects per-outlet profit, set-up-cost of a new outlet and the cost of entering a new city for the chain.

However, to be able to perform this type of analysis we need to estimate $\boldsymbol{\theta}_i$ and then find the equilibrium of the model as defined in Section 3.6. Both of these are non-trivial tasks under any circumstances and given the size of the state space in our model, they are virtually impossible. Our solution to this problem, as we explain in the next section, is to estimate the model parameters using an estimation procedure proposed by Bajari et al. (2007). After estimating the parameters, we can use forward simulations to find numerical values of V^{IN} and V^{OUT} for both chains. Once this is done, we can exogenously change the state vector and study the response of V^{IN} and V^{OUT} to the change.

4 Estimation

Our estimation strategy is an extension of Bresnahan and Reiss (1991). Their idea is that if we observe N homogenous firms in a market and a firm's profits are a decreasing function of the number of firms in the market then it must be the case that $\Pi(N+1) < 0$ and $\Pi(N) \geq 0$ (Π is the profit function). In words: if there are N firms in the market then $N+1$ firms would be too many for that market and $N-1$ firms too few.

We extend their idea to a dynamic setting with heterogenous multi-outlet chains. Our idea is that if a chain opens x outlets in a market in a given period, $x+1$ outlets must be too many to open for the chain in that particular market in one period and $x-1$ outlets must be too little. Similarly, when a chain enters a city (entry is irreversible in our model) at time τ, it must be the case that time $\tau-1$ is too early for entry and $\tau+1$ too late.

The objective of estimation is to estimate parameters in vector $\boldsymbol{\theta}_i = \{\boldsymbol{\theta}_i^{\pi}, \boldsymbol{\theta}_i^{e}, \boldsymbol{\theta}_i^{s}\}$.[15] A full-solution method, like Rust's Nested-Fixed-Point Algorithm [Rust (1987)], will involve solving the dynamic programming problems in (2) and (3) simultaneously for both chains for a given set of parameters at every iteration of the estimation process. Given the size of the state space for our model, computation of such equilibrium even once is practically impossible.

To solve this problem we make use of the estimation method proposed by Bajari et al. (2007). The idea is to take the observed entry and expansion decisions as equilibrium outcomes generated from a unique Markov-Perfect Equilibrium (MPE). This allows one to estimate the entry and expansion strategy profiles, $\eta^*(\cdot)$ and $\chi^*(\cdot)$, for both chains directly from the data and obviates the need to compute equilibrium for the dynamic problems in (2) and (3). With $\eta^*(\cdot)$ and $\chi^*(\cdot)$ in hand, one can use forward simulations to get numerical estimates of V^{IN} and V^{OUT}. Because these values are based on equilibrium strategy profiles, by definition they must be higher than the values generated by any alternative (non-equilibrium) strategy profiles $\eta^l(\cdot)$ and $\chi^l(\cdot)$. The idea of estimation is to pick the parameters of the value functions in such a way that the values based on $\eta^*(\cdot)$ and $\chi^*(\cdot)$ are at least as high as the values based on various alternative strategy profiles.

Our estimation proceeds in three stages. In the first stage we estimate $\eta^*(\cdot)$ and $\chi^*(\cdot)$ from the data together with transition matrices for population and GDP p.c. variables. In the second stage we forward simulate to get the numerical estimates of V^{IN} under equilibrium and non-equilibrium strategy profiles. We then use equilibrium conditions of the model to recover parameters in V^{IN} function. In the third stage we forward simulate V^{OUT} under equilibrium and non-equilibrium strategy profiles and use equilibrium conditions to recover parameters in V^{OUT}. The sequential estimation of the parameters in V^{IN} and V^{OUT} eases the burden of computation. We now provide details on each stage of estimation.

4.1 Stage One

A key element of BBL's methodology is the assumption that the data represent a unique Markov Perfect Equilibrium. With this assumption it is possible to estimate equilibrium strategies directly from the data. Ideally this should be done non-parametrically but that would require a very large dataset. A more practical option is to use a flexible parametric specification.

We use data on entry decisions of the two chains to estimate a separate Probit model for each

[15] Our estimation strategy does not identify the discount factor β hence it is fixed for identification.

chain. This gives us the probability of a chain entering a city as a function of the observed state. This is our empirical estimate of $\eta^*(\cdot)$.[16] Similarly, we use data on expansion decisions to estimate a separate Negative-Binomial (NB) model for each chain. This gives us the expected number of new outlets that a chain would open in equilibrium as a function of the observed state.[17] This is our empirical estimate of $\chi^*(\cdot)$.[18]

We also estimate the state transition matrices for GDP p.c. and population in stage one. To estimate the state-transition matrix, we first discretize the variable.[19] Let $\Pr(s, s')$ denote an element of the state-transition matrix, where s is the current period state and s' is the next period state. To find the value of $\Pr(s, s')$, we count the observations in the data with the current state s and the future state s'. We then divide this number by the number of observations in the data with the current state s.

$$\Pr(s, s') = \frac{\sum \mathbf{1}\,(\text{Current State } = s \text{ and Next-period State } = s')}{\sum \mathbf{1}\,(\text{Current State } = s)}, \tag{5}$$

where $\mathbf{1}(\cdot)$ is an indicator function and the sums are over all observations in the sample.

4.2 Stage Two

In stage two we forward simulate the model from an initial state, say $\mathbf{s}^{\text{IN}}_{i,c,0}$, to get numerical estimates of V^{IN}_i for a given parameter vector $\boldsymbol{\theta}^{\text{IN}}_i = \{\boldsymbol{\theta}^{\pi}_i, \boldsymbol{\theta}^{s}_i\}$. We can write the Bellman's equation in (3) in its sequence form as:

$$V^{\text{IN}}_i\left(\mathbf{s}^{\text{IN}}_{i,c,0}|\boldsymbol{\theta}^{\text{IN}}_i, \eta_j\left(\mathbf{s}^{\text{OUT}}_{j,c}\right), \chi_j\left(\mathbf{s}^{\text{IN}}_{j,c}\right)\right) = \max_{\{x_{i,c,t}\in\{0,1,\ldots,\bar{x}\}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\pi_i\left(\mathbf{s}^{\text{IN}}_{i,c,t}|\boldsymbol{\theta}^{\pi}_i\right) n_{i,c,t} - c^s_i\left(\mathbf{s}^{\text{IN}}_{i,c,t}|\boldsymbol{\theta}^{s}_i\right) x_{i,c,t}\right]. \tag{6}$$

We already have $\eta^*_j\left(\mathbf{s}^{\text{OUT}}_{j,c}\right)$, $\chi^*_j\left(\mathbf{s}^{\text{IN}}_{j,c}\right)$ and $\chi^*_i\left(\mathbf{s}^{\text{IN}}_{i,c}\right)$ from the first stage. Substituting $\chi^*_i\left(\mathbf{s}^{\text{IN}}_{i,c,t}\right)$ for $x_{i,c,t}$, $\eta^*_j(\cdot)$ for $\eta_j(\cdot)$ and $\chi^*_j(\cdot)$ for $\chi_j(\cdot)$, we can write (6) as

$$V^{\text{IN}}_i\left(\mathbf{s}^{\text{IN}}_{i,c,0}|\boldsymbol{\theta}^{\text{IN}}_i, \eta^*_j\left(\mathbf{s}^{\text{OUT}}_{j,c}\right), \chi^*_j\left(\mathbf{s}^{\text{IN}}_{j,c}\right)\right) = \sum_{t=0}^{\infty} \beta^t \left[\pi_i\left(\mathbf{s}^{\text{IN}}_{i,c,t}|\boldsymbol{\theta}^{\pi}_i\right) n_{i,c,t} - c^s_i\left(\mathbf{s}^{\text{IN}}_{i,c,t}|\boldsymbol{\theta}^{s}_i\right) \chi^*_i\left(\mathbf{s}^{\text{IN}}_{i,c,t}\right)\right]. \tag{7}$$

Because we do not have data on prices, quantities and costs, we cannot estimate the demand and cost functions that underlie the profit functions. Instead, following the recent empirical literature on entry and market structure [Bresnahan and Reiss (1991), Berry (1992), Toivanen and Waterson

[16] Despite observing the state and entry decisions, a researchers can only recover the probability of a decision because she does not observe the private entry-cost shocks of the chains.

[17] In the case of expansion too, the researcher can only recover the probability of each possible decision. Our Negative Binomial estimates give us the expected number of new outlets.

[18] The NB model is an appropriate model for expansion decisions because the dependent variable, the number of new outlets opened, is a count variable with many zero observations in the data (see Table 4). Also, the variance of the number of new outlets far exceeds its mean (see Table 5), hence a NB model is more suitable than a simple Poisson model. We later test the suitability of the NB model over the Poisson model for our dataset and find strong support for it (see Table A.2).

[19] In Tables A.3 and A.4 we show how we discretize GDP p.c. and population variables.

(2005), Jia (2008), etc.], we assume a linear reduced-form profit function:[20]

$$\begin{aligned}\pi_i\left(\mathbf{s}^{\mathrm{IN}}_{i,c,t}|\boldsymbol{\theta}^{\pi}_i\right) \quad = \quad & \theta^{\pi}_{i,0}+\theta^{\pi}_{i,n_i}n_{i,c,t}+\theta^{\pi}_{i,n_i^2}n^2_{i,c,t}+\theta^{\pi}_{i,n_j}n_{j,c,t}+\theta^{\pi}_{i,n_j^2}n^2_{j,c,t}+\theta^{\pi}_{i,a_i}a_{i,t}+\theta^{\pi}_{i,a_i^2}a^2_{i,t}+\\ & \theta^{\pi}_{i,d_i}d_{i,c}+\theta^{\pi}_{i,p}p_{c,t}+\theta^{\pi}_{i,y}y_{c,t}+\theta^{\pi}_{i,z_1}z_{1,c}+\theta^{\pi}_{i,z_2}z_{2,c}+\theta^{\pi}_{i,k}k_c,\end{aligned} \tag{8}$$

The per outlet profit $\pi_i\left(\mathbf{s}^{\mathrm{IN}}_{i,c,t}|\boldsymbol{\theta}^{\pi}_i\right)$ depends on the total number of one's own outlets in the city in period t, $n_{i,c,t}$, the total number of the rival's outlets, $n_{j,c,t}$, the chain's experience $a_{i,t}$, and the market characteristics including the distance of the city to the chain's headquarters, population size, GDP p.c., whether the city implemented any special policy favoring foreign investment and whether it is a provincial-capital city.

These parameterizations have the advantage of being simple but, as Berry and Reiss (2007) point out, it is difficult to attach any meaningful economic interpretation to various parameters. Also, with this parameterization we are unable to separately identify demand and cost elements within the profit functions.

Similarly, we assume a linear parameterization for the set-up cost of a new outlet:

$$\begin{aligned}c^s_i\left(\mathbf{s}^{\mathrm{IN}}_{i,c,t}|\boldsymbol{\theta}^{s}_i\right) \quad = \quad & \theta^{s}_{i,0}+\theta^{s}_{i,n_i}n_{i,c,t}+\theta^{s}_{i,n_i^2}n^2_{i,c,t}+\theta^{s}_{i,n_j}n_{j,c,t}+\theta^{s}_{i,n_j^2}n^2_{j,c,t}+\theta^{s}_{i,a_i}a_{i,t}+\theta^{s}_{i,a_i^2}a^2_{i,t}+\\ & \theta^{s}_{i,d_i}d_{i,c}+\theta^{s}_{i,p}p_{c,t}+\theta^{s}_{i,y}y_{c,t}+\theta^{s}_{i,z_1}z_{1,c}+\theta^{s}_{i,z_2}z_{2,c}+\theta^{s}_{i,k}k_c+\sigma^s_i\epsilon^s_{i,c,t}.\end{aligned} \tag{9}$$

For forward simulations, we start from an arbitrary state $\mathbf{s}^{\mathrm{IN}}_{i,c,0}$. Given the first-stage estimates we can find the next period state $\mathbf{s}^{\mathrm{IN}}_{i,c,1}$. We continue this process until β^t is almost equal to zero. We do so a large number of times to average over the private sunk-cost shock of the rival chain and to take care of the stochastic nature of population and GDP p.c. variables. We discount the expected numerical values of future states to present by using the discount factor β. This process generates present discounted values of various state variables which are based on both chains following their equilibrium strategies. Because the functions $\pi_i(\cdot)$ and $c^s_i(\cdot)$ are linear in parameters, we can use the same forward simulated values to get numerical estimates of V^{IN}_i for any candidate parameter vector $\boldsymbol{\theta}^{\mathrm{IN}}_i$. Let $V^{\mathrm{IN}}_i\left(\mathbf{s}^{\mathrm{IN}}_{i,c,0}|\boldsymbol{\theta}^{\mathrm{IN}}_i,\chi^*_i,\eta^*_j,\chi^*_j\right)$ denote this numerical estimate. This is a numerical estimate of V^{IN}_i for state $\mathbf{s}^{\mathrm{IN}}_{i,c,0}$ and a given parameter vector $\boldsymbol{\theta}^{\mathrm{IN}}_i$ when chain i follows its equilibrium strategy χ^*_i and its rival follows its equilibrium strategies η^*_j or χ^*_j.

Next we repeat these forward simulations using alternative (non-equilibrium) strategies for chain i.[21] To construct alternative strategy profiles we randomly draw 100 alternative strategy parameter vectors according to

$$\theta^{p'}=\theta^p\left(1+\varepsilon\right), \tag{10}$$

where $\varepsilon\sim N(0,1)$, θ^p is the equilibrium strategy parameter vector estimated in the first stage and $\theta^{p'}$ is the alternative strategy parameter vector. Under both equilibrium and alternative strategies, the number of outlets to open is bounded between 0 and $\bar{x}$, both inclusive. We set $\bar{x}=50$.[22] Let

[20] We denote most of the parameters in this paper by Greek letter θ, and use superscripts and subscripts to distinguish among them. We use superscripts to identify whether a parameter belongs to the profit function (superscript π) or the setup cost of a new outlet (superscript s) or the cost of entry function (superscript e). We use i in the subscript to highlight that there are separate parameters for chains i and j. The second symbol in the subscript, after i, identifies the variable to which that coefficient belongs. For example, for parameter θ^{π}_{i,n_i} in (8), n_i in the subscript indicates that this parameter measures the effect of the number of own outlets on per outlet profit.

[21] Chain j still follows its equilibrium strategies when chain i follows the alternative ones.

[22] In our dataset the maximum number of outlets that KFC has ever opened in a city within a year is 33 (in Shanghai in 2004). The maximum for McDonald's is 27 (in Shanghai in 2006).

χ_i^l denote the lth alternative strategy profile, then $V_i^{\text{IN}}\left(\mathbf{s}_{i,c,0}^{\text{IN}}|\boldsymbol{\theta}_i^{\text{IN}},\chi_i^l,\eta_j^*,\chi_j^*\right)$ denotes the numerical estimate of V^{IN} when chain i follows alternative strategy l.

If χ_i^* is indeed the Markov Perfect Equilibrium strategy profile then the following must hold at the true parameter vector $\boldsymbol{\theta}_i^{*\text{IN}}$.

$$V_i^{\text{IN}}\left(\mathbf{s}_{i,c,0}^{\text{IN}}|\boldsymbol{\theta}_i^{*\text{IN}},\chi_i^*,\eta_j^*,\chi_j^*\right)\geq V_i^{\text{IN}}\left(\mathbf{s}_{i,c,0}^{\text{IN}}|\boldsymbol{\theta}_i^{*\text{IN}},\chi_i^l,\eta_j^*,\chi_j^*\right). \tag{11}$$

But the inequality may not hold for some states if $\boldsymbol{\theta}_i^{*\text{IN}}$ is replaced by an alternative parameter vector $\boldsymbol{\theta}_i^{\text{IN}}$. BBL propose to estimate $\boldsymbol{\theta}_i^{*\text{IN}}$ by minimizing the violations of the inequality in (11). More formally, let us define a difference variable d as

$$d_i^{\text{IN}}\left(\mathbf{s}_{i0}|\boldsymbol{\theta}_i^{\text{IN}},\chi_i^l,\eta_j^*,\chi_j^*\right)=V_i^{\text{IN}}\left(\mathbf{s}_{i,c,0}^{\text{IN}}|\boldsymbol{\theta}_i^{\text{IN}},\chi_i^*,\eta_j^*,\chi_j^*\right)-V_i^{\text{IN}}\left(\mathbf{s}_{i,c,0}^{\text{IN}}|\boldsymbol{\theta}_i^{\text{IN}},\chi_i^l,\eta_j^*,\chi_j^*\right), \tag{12}$$

and a criterion function Q as

$$Q\left(\boldsymbol{\theta}_i^{\text{IN}}\right)=\sum_{\mathbf{s}_{i,c,0}^{\text{IN}}}\sum_{l=1}^{L}\left[\min\left(d_i^{\text{IN}}\left(\mathbf{s}_{i,c,0}^{\text{IN}}|\boldsymbol{\theta}_i^{\text{IN}},\chi_i^l,\eta_j^*,\chi_j^*\right),0\right)\right]^2, \tag{13}$$

where the first sum is over all states in the state space and the second sum is over L alternative strategy profiles.[23] The BBL point estimate of $\boldsymbol{\theta}_i^{*\text{IN}}$ is then given by

$$\hat{\boldsymbol{\theta}}_i^{\text{IN}}=\arg\min_{\boldsymbol{\theta}_i^{\text{IN}}}Q\left(\boldsymbol{\theta}_i^{\text{IN}}\right).$$

Once we have estimated $\hat{\boldsymbol{\theta}}_i^{\text{IN}}$ we can use the estimated parameters and find V_i^{IN} for any arbitrary state $\mathbf{s}_{i,c}^{\text{IN}}$. Let $V_i^{\text{IN}}\left(\mathbf{s}_{i,c}^{\text{IN}}|\hat{\boldsymbol{\theta}}_i^{\text{IN}},\chi_i^*,\eta_j^*,\chi_j^*\right)$ denote this value.

4.3 Stage Three

In stage three we take $V_i^{\text{IN}}\left(\mathbf{s}_{i,c}^{\text{IN}}|\hat{\boldsymbol{\theta}}_i^{\text{IN}},\chi_i^*,\eta_j^*,\chi_j^*\right)$ as given and estimate the entry-cost function $c_i^e\left(\boldsymbol{\theta}_i^e,\mathbf{s}_{i,c}^{\text{OUT}}\right)$. If chain i is in state $\mathbf{s}_{i,c}^{\text{OUT}}$, i.e. it has yet to enter the city, it solves the problem in (2). The sequence version of that problem can be written as

$$V_i^{\text{OUT}}\left(\mathbf{s}_{i,c,0}^{\text{OUT}}|\boldsymbol{\theta}_i,\eta_j\left(\mathbf{s}_{j,c}^{\text{OUT}}\right),\chi_j\left(\mathbf{s}_{j,c}^{\text{IN}}\right)\right)=\max_{\{e_{i,c,t}\in\{0,1\}\}_{t=0}^{\tau}}\left\{\sum_{t=0}^{\tau}\beta^t e_{i,c,t}\left[\beta E\left(V_i^{\text{IN}}\left(\mathbf{s}_{i,c,t}^{\prime\text{IN}}\right)\right)-c_i^e\left(\boldsymbol{\theta}_i^e,\mathbf{s}_{i,c,t}^{\text{OUT}}\right)\right]\right\}, \tag{14}$$

where $\tau\geq 0$ is the period of entry, which is the first period when chain i chooses $e_{i,c,t}=1$. Substituting $\eta_i^*\left(\mathbf{s}_{i,c,t}^{\text{OUT}}\right)$ for $e_{i,c,t}$, $\eta_j^*\left(\cdot\right)$ for $\eta_j\left(\cdot\right)$ and $\chi_j^*\left(\cdot\right)$ for $\chi_j\left(\cdot\right)$, we can write (14) as

$$V_i^{\text{OUT}}\left(\mathbf{s}_{i,c,0}^{\text{OUT}}|\boldsymbol{\theta}_i,\eta_j^*\left(\mathbf{s}_{j,c}^{\text{OUT}}\right),\chi_j^*\left(\mathbf{s}_{j,c}^{\text{IN}}\right)\right)=\left\{\sum_{t=0}^{\tau}\beta^t\eta_i^*\left(\mathbf{s}_{i,c,t}^{\text{OUT}}\right)\left[\beta E\left(V_i^{\text{IN}}\left(\mathbf{s}_{i,c,t+1}^{\text{IN}}\right)\right)-c_i^e\left(\boldsymbol{\theta}_i^e,\mathbf{s}_{i,c,t}^{\text{OUT}}\right)\right]\right\}.$$

If chain i has not entered the city yet, it considers entry at every period. If its optimal choice is to stay out, i.e. $\eta_i^*\left(\mathbf{s}_{i,c,t}^{\text{OUT}}\right)=0$, it does nothing except waiting for the next period. When the conditions are ripe for entry, it will enter. We denote the period of entry by τ. In period $t=\tau$, $\eta_i^*\left(\mathbf{s}_{i,c,t}^{\text{OUT}}\right)=1$ and the chain will pay the cost of entry $c_i^e\left(\boldsymbol{\theta}_i^e,\mathbf{s}_{i,c,t}^{\text{OUT}}\right)$. Although the entry takes place when $t=\tau$, the newly opened outlet starts business in the next period. Hence the expected gross benefit of entry at $t=\tau$ is $\beta E\left(V_i^{\text{IN}}\left(\mathbf{s}_{i,c,t+1}^{\text{IN}}\right)\right)$. The net benefit of entry is then $\beta E\left(V_i^{\text{IN}}\left(\mathbf{s}_{i,c,t+1}^{\text{IN}}\right)\right)-c_i^e\left(\boldsymbol{\theta}_i^e,\mathbf{s}_{i,c,t}^{\text{OUT}}\right)$.

[23] In our application $L=100$.

This net benefit is in current-value terms when $t = \tau$. To find its present value, we need to discount it back to period 0 by multiplying it with β^t.

For $c_i^e\left(\boldsymbol{\theta}_i^e, \mathbf{s}_{i,c,t}^{\mathrm{OUT}}\right)$ too we assume a linear parameterization:

$$\begin{aligned} c_i^e\left(\boldsymbol{\theta}_i^e, \mathbf{s}_{i,c,t}^{\mathrm{OUT}}\right) \quad = \quad & \theta_{i,0}^e + \theta_{i,n_j}^e n_{j,c,t} + \theta_{i,a_i}^e a_{i,t} + \theta_{i,a_i^2}^e a_{i,t}^2 + \theta_{i,d_i}^e d_{i,c} + \\ & \theta_{i,p}^e p_{c,t} + \theta_{i,y}^e y_{c,t} + \theta_{i,z_1}^e z_{1,c} + \theta_{i,z_2}^e z_{2,c} + \theta_{i,k}^e k_c + \sigma_i^e \epsilon_{i,c,t}^e. \end{aligned} \quad (15)$$

To estimate the parameters in the entry cost function we apply BBL's method once more. This time the forward simulations start from an arbitrary state $\mathbf{s}_{i,c,0}^{\mathrm{OUT}}$ and continue until the period of entry. When we need a numerical estimate of $E\left(V_i^{\mathrm{IN}}\left(\mathbf{s}_{i,c,t+1}^{\mathrm{IN}}\right)\right)$ in the period of entry, we use forward-simulated estimates of $V_i^{\mathrm{IN}}\left(\mathbf{s}_{i,c}^{\mathrm{IN}}|\hat{\boldsymbol{\theta}}_i^{\mathrm{IN}}, \chi_i^*, \eta_j^*, \chi_j^*\right)$ based on stage-two parameter estimates.

4.4 Identification

In BBL's method, the parameters of a model are point identified if there is a unique parameter vector that satisfies all possible equilibrium inequalities implied by (11). In practice, to make estimation computationally feasible, one has to work with a sample analog of $Q\left(\boldsymbol{\theta}_i^{\mathrm{IN}}\right)$ in (13) and compute it using a small subset of all possible initial states and a limited set of alternative policies. Srisuma (2010) discusses problems associated with this choice. He shows that it is possible to lose a lot of identification information if the alternative policies are not carefully chosen. He suggests that the set of alternative policies should be determined by how the state variables are transformed into an action. In other words, the alternatives should be informative about the structural parameters.

In practice, the alternatives are informative about the structural parameters if for the same state they suggest choices different than the equilibrium choice. In our estimation, we allow for 100 randomly chosen alternatives. We also allow the alternatives to differ remarkably from the equilibrium choices (see (10)). The large number of alternatives and the great variation in them help us identify the structural parameters in the second and third stages of estimation.

The first-stage entry strategies are identified from variation in the entry decisions in response to changes in the observed state. Similarly, the first-stage expansion decisions are identified from variation in the number of new outlets opened in a period in response to changes in state.

5 Estimation Results

We now discuss the estimation results for each stage separately.

5.1 Stage One Estimates

In stage one, we estimate the equilibrium entry and expansion strategies for the two chains. We also estimate the state-transition matrices for GDP p.c. and population variables.

Our estimates of entry strategies are in Table A.1. Both chains are more likely to enter a city if: their rival has more outlets in the city; they are more experienced in the country (as measured by the age variable); the city is closer to their headquarters, more populated, and has a higher GDP p.c. Besides KFC is much more likely to enter a provincial capital. The rival's presence has a significant and positive effect on the focal firm's probability of entry. Yet for McDonald's the effect can become negative in cities that implemented "promote-west" policy, as seen from the significantly negative coefficient of the interaction term of rival's outlets and the policy dummy.[24] This could happen possibly due to the fact the chains got some information from their rival's presence in the city before the promote-west policy. However, after the promote-west policy, they took advantage of the government's policy initiative and rival's presence became less important.

Our estimates of expansion strategies are in Table A.2. Both chains open more outlets in the cities where they already have greater presence, though this effect weakens as the number of own outlets in the city increases. The effect of the number of rival's outlets on expansion is positive and significant for KFC. KFC also expands faster as it gains more experience in the country. For McDonald's the rival's effect and the age effect are not statistically significant. Proximity to own headquarters favors expansion for KFC but has a statistically insignificant effect for McDonald's. A larger population of the city affects the expansion positively for both chains. GDP p.c. has a small negative effect on KFC's expansion and a small positive effect on McDonald's expansion. Both chains expand faster in provincial capital cities.

We estimate separate strategy functions for the two chains. Our results justify this choice and we find that the two chains respond quite differently to some state variables.

In the case of expansion strategies, we estimate highly significant over-dispersion parameter for both chains. This justifies our choice of a Negative Binomial model over a Poisson model.

We estimate the state transition matrices using the procedure described in Section 4.1. We report the estimated matrices in Appendix A (Tables A.5 and A.6). The transition matrix for GDP p.c., Table A.5, shows that although it is a persistent variable, it generally has a high probability (around a quarter) of going up. This reflects the rapid growth of the Chinese economy during the sample period (1990-2007). The transition matrix for population, Table A.6, shows a much higher level of persistence and some probability (mostly around 4%) of increase. This increase is mainly due to urbanization.

5.2 Stage Two Estimates

In stage two we estimate the parameters in the profit and set-up-cost functions $(\pi(\mathbf{s}), c^s(\mathbf{s}))$. We then use the estimated parameters to find numerical estimates of $V^{\text{IN}}(\mathbf{s})$. We report the parameter estimates in Table A.7.[25] Our primary interest is in spillover effects from the rival. The key result is that rival's presence has a positive and significant spillover effect on the cost of setting up new outlets. The reduction in the set-up cost of a new outlet could be due to more familiarity of the city with the Western food chains that makes it easier for the chain to acquire or rent buildings and make arrangements for its initial manpower needs. It could also be due to greater support from the local government to encourage foreign investment in the city. The reduced-form estimates in Table A.2 suggest that after the chains have entered a city, KFC opens more outlets in the cities with a greater rival presence whereas McDonald's opens fewer outlets in such cities. The structural estimates provide some insights into these reduced-form results. While a greater rival presence in the city reduces the set-up cost of new outlets, it also reduces per outlet profits for both chains due to the competition effect: as the rival opens more outlets, it takes away some of the demand for the focal chain's food. The net long-term effect of these opposite forces is positive for KFC and negative for McDonald's. We comment more on this net effect below.

[24] For KFC the coefficients are negative but statistically insignificant.

[25] We fix the value of the discount factor β at 0.90.

We also notice that the greater number of one's own outlets in the city is associated with lower set-up cost possibly due to the managerial or supply-chain-related economies. However, when the number of one's own outlets increases, it also creates the cannibalization problem which lowers per outlet profit. As a chain gains more experience in the country (captured by 'age'), it faces a lower set-up cost possibly due to learning effects. The per outlet profit on the other hand decreases over time possibly due to the increased competition in the food/restaurant industry. Both chains face higher set-up cost in cities that are farther away from their headquarters. The cost is lower in more populated and richer cities. Finally, the open policies are effective in reducing the cost of opening new outlets.

The last row of Table A.7 reports the percentage of the equilibrium conditions that are satisfied at the reported parameters. Recall that the idea of estimation is to choose the parameter vector θ_i^{IN} that minimizes the violation of the equilibrium conditions (see (11)). In our estimation, there are $125,800$ equilibrium conditions for KFC and $78,800$ equilibrium conditions for McDonald's. Our algorithms do a decent job of satisfying 84.9% equilibrium conditions for KFC and 77.3% for McDonald's. We note that including more state variables in profit and set-up cost functions helps us increase the percentage of equilibrium conditions that are satisfied at the estimated parameters. We also experimented with more restrictive profit and/or set-up-cost functions. In those experiments, the estimated parameters satisfied a much smaller percentage of equilibrium conditions.[26]

With the profit and set-up-cost functions in hand, we can use forward simulations to get numerical estimates of $V^{\text{IN}}(\mathbf{s})$.[27] We report some summary statistics on the numerical estimates of $V^{\text{IN}}(\mathbf{s})$ for both chains in Table 6.[28]

Table 6: Summary statistics on the numerical estimates of V^{IN} (¥m)

Chain	Obs.	Mean	S.D	p_{10}	p_{50}	p_{90}
For all observed states						
KFC	1258	351	761	51	113	792
McDonald's	788	312	556	57	113	710
For observed states with $n_i = 1$						
KFC	528	60	12	49	57	77
McDonald's	361	60	5	56	59	65

Note: The constant in the profit function, θ_{i0}^{π}, is normalized to ¥17.9m for KFC and ¥9.3m for McDonald's.

The interpretation of $V^{\text{IN}}(\mathbf{s})$ is that it is the expected discounted value of future profit streams (net of the set-up-cost of new outlets) when the chain and its rival both follow equilibrium strategies. Average numerical value of $V^{\text{IN}}(\mathbf{s})$ is ¥351 million for KFC and ¥312 million for McDonald's. These values vary a lot from city to city and over time as the state of a city changes. This is evident from high standard deviations (¥761m and ¥556m).

The numerical estimates of the value function, $V^{\text{IN}}(\mathbf{s})$, are derived under the assumption that

[26] The results are available from authors upon request.

[27] Once the parameters are known, in principle, one can solve for the numerical value function as a fixed point of the Bellman's equation in (3). However, because of the very large statespace and the interdependence between the two chains, computation of the fixed point is computationally infeasible.

[28] The summary statistics are based on simulated values for the observed states only. Using the same procedure one can compute simulated values for any state in the state space.

both chains follow their equilibrium strategies, which we estimate from the data in stage one, to enter a city or expand within a city. Once we have the numerical estimates, if we change the value of a state variable, a change in V^{IN} will tell us the effect of this exogenous change in state on the sum of the expected discounted value of future profit streams. This is a major advantage of a structural model that it alows us to *quantify*, subject to a normalization, the long-term exogenous effects of various state variables on the expected net present value of a chain's profit stream. A simple way to see these quantitative effects is to regress numerical estimates of V^{IN} on various state variables. We report the results of one such exercise in Table 7.

The first thing to note is that these simple linear regressions can explain about 97% of variation in the estimated values. This is partly due to the fact that our first-stage regressions as well as the assumed profit, set-up-cost and entry-cost functions are all linear.

Table 7: How V^{IN} varies with state

Dependent variable: V^{IN}		
Independent variable	KFC	McDonald's
Number of own outlets in the city	43.14 (0.45)	44.02 (0.48)
Number of rival's outlets in the city	3.79 (0.61)	−4.60 (0.36)
Own age (years)	−10.09 (1.28)	−6.53 (1.26)
City population (thousands)	0.02 (0.01)	0.03 (0.005)
City GDP p.c. (thousands)	0.90 (0.30)	2.69 (0.28)
Provincial capital city (dummy)	157.53 (15.49)	−73.70 (14.17)
z_1(Open city 1993, dummy)	−162.16 (15.04)	−20.21 (14.19)
z_2(Open city 2001, dummy)	−0.82 (11.51)	6.09 (12.80)
R^2	0.97	0.97
Observations	1,258	788

Notes: 1) Parameter θ_{i0}^{π} is normalized to ¥17.9m for KFC and ¥9.3m for McDonald's; 2) The numbers in parentheses are standard errors.

For both chains, the values are strongly positively related to the number of own outlets in the city: one more outlet in the city, increases the value by ¥43.14m for KFC and by ¥44.02m for McDonald's. This is approximately 2.4 times (43.14/17.9) the average annual profit per outlet for KFC and 4.7 times (44.02/9.3) for McDonald's.

The effect on V^{IN} of the number of rival outlets is different, as we mentioned above, for the two chains. For KFC, more McDonald's outlets increase value. In other words, there is an overall positive spillover from McDonald's to KFC and it is worth ¥3.79m. For McDonald's, on the other hand, more KFC outlets decrease the value by about ¥4.60m. The reason for these opposite effects is the balance of positive and negative effects from the rival. Our structural estimates suggest that for both chains a greater presence of the rival reduces per outlet profits (a negative effect) and also reduces the set-up cost of opening a new outlet (a positive effect). For KFC, the positive effect donimates the negative effect and for McDonald's the opposite is true.

Among other state variables, age affects value negatively for both chains. Population and GDP p.c. have small but positive effects on values. Being in a provincial capital generates a higher value for KFC but a lower value for McDonald's. The structural estimates also decompose these effects. They suggest that being in a provincial capital city increases per outlet profits for KFC but also increases the costs of setting up new outlets. The net effect is positive. For McDonald's, being in a provincial capital city decreases per outlet profit but also decreases the set-up cost and the net effect is negative.

The first openness dummy (z_1) lowers V^{IN} for both chains. As we shall see later, this implies a lower cost of entry into the open cities. The effect of the second openness dummy (z_2) is statistically insignificant.

5.3 Stage Three Estimates

Having obtained the numerical estimates of V^{IN}, we proceed to estimate the entry-cost function. We report the entry-cost parameter estimates in Table 8.

Table 8: Entry-cost parameters

Variable	KFC	McDonald's
$\theta^e_{i,0}$	116.97 (0.48)	73.08 (0.14)
θ^e_{i,n_j}	0.02 (0.06)	-0.06 (0.01)
θ^e_{i,a_i}	-5.25 (0.05)	-1.65 (0.02)
θ^e_{i,a_i^2}	0.07 (4.60×10^{-3})	0.02 (5.41×10^{-4})
$\theta^e_{i,p}$	-1.18×10^{-4} (4.2×10^{-5})	2.48×10^{-4} (1.06×10^{-5})
$\theta^e_{i,y}$	3.91×10^{-4} (3.11×10^{-3})	-3.48×10^{-3} (6.69×10^{-4})
θ^e_{i,d_i}	1.31×10^{-3} (5.24×10^{-5})	6.07×10^{-4} (1.37×10^{-5})
θ^e_{i,z_1}	-4.96 (0.16)	-1.23 (0.06)
θ^e_{i,z_2}	0.06 (0.11)	-0.81 (0.02)
$\theta^e_{i,k}$	3.80 (0.18)	-13.63 (0.06)
σ^e_i	4.72×10^{-3} (0.09)	3.68×10^{-3} (0.02)
EC[1]	278,400	325,400
ECS[2]	80.9%	73.0%

EC[1]: Equilibrium conditions

ECS[2]: Equilibrium conditions satisfied

Note: Numbers in parentheses are bootstrapped standard errors.

These estimates are based on V^{IN} estimates from the second stage and the observed entry (and non-entry) decisions of the two chains. The large autonomous entry costs, θ^e_0, are consistent with the observation that despite the large expected gains from entry, i.e. V^{IN}, the chains decided not to enter many cities or waited for quite a few years before entry.

Coming to the central question of spillovers from the rival, we find the effect to be different across chains. The number of rival outlets has no significant effect on the entry cost for KFC. It is consistent with KFC moving first to enter most of the cities. For McDonald's, the presence of KFC significantly lowers the entry cost. How do we reconcile these structural estimates with the reduced-form estimates in Table A.1? The reduced-form estimates suggest that both chains are more likely to enter the cities where the rival has a greater presence. The structural estimates provide the following clues for this behavior. For KFC a greater presence of the rival in the city increases V^{IN} (Table 7) but does not affect the entry cost by much and hence KFC is more likely to enter the city. For McDonald's, although a greater presence of the rival in the city reduces V^{IN} (Table 7), it also reduces the entry cost and the net effect is positive. We discuss the net effects below when we comment on the results in Table 10.

Experience in the country, as measured by age, lowers the entry cost for both chains. Population and GDP p.c. have negligible effects. Both chains find it easier to enter the cities that are closer to their respective headquarters. The cost of entry is lower for cities covered under open and promote-west policies. McDonald's finds entering a provincial capital much easier, though KFC finds it harder.

Our algorithms to estimate the parameters reported in Table 8 contained $278,400$ equilibrium conditions for KFC and $325,400$ for McDonald's. At the reported parameter estimates, 80.9% of the conditions were satisfied for KFC and 73.0% for McDonald's.

The actual decision to enter a city does not depend solely on the cost of entry. The expected value of entry plays an equally important role. The numerical estimates of V^{OUT} capture the net effect of these two factors. We report some summary statistics on the numerical estimates of V^{OUT} in Table 9.

Table 9: Summary statistics on the numerical estimates of V^{OUT} (¥ m)

Chain	Obs.	Mean	S.D	p_{10}	p_{50}	p_{90}
For all observed states						
KFC	2784	0.83	0.43	0.38	0.75	1.41
McDonald's	3254	−0.15	0.14	−0.23	−0.19	−0.04
For observed states with $entry_i = 1$						
KFC	223	1.32	0.49	0.79	1.28	1.88
McDonald's	140	−0.04	0.36	−0.24	−0.16	0.22

The estimates of V^{OUT} are small in absolute terms with an average of less than a million for KFC and a negative average for McDonald's. If we look only at the entry periods, the values are slightly bigger as one would expect.

However, what is important to understand the entry decisions is that how V^{OUT} depends on various state variables. To see this we regress V^{OUT} on relevant state variables. We report the results in Table 10.

To better interpret the results, note that the SD of V^{OUT} is 0.43 for KFC and 0.14 for McDonald's (Table 9). Also note that the ratio of $p_{90} - p_{10}$ to SD for a standard Normal distribution is 2.56 (i.e.

a spread of 2.56 SD's around the mean covers 80% of the distribution) whereas the same ratio for the distribution of the estimated V^{OUT} is 2.40 for KFC and just 1.36 for McDonald's. So a change of one SD in V^{OUT} will have a bigger effect on entry decision than does a one SD change in the case of a Normal distribution and more so for McDonald's. Keeping this in mind, we now interpret some of the results in Table 10.

Table 10: How V^{OUT} varies with state

Dependent variable: V^{OUT}		
Variable	KFC	McDonald's
Number of rival's outlets in the city	-0.07	0.06
	(0.01)	(1.56×10^{-3})
Own age (years)	0.08	-0.01
	(1.10×10^{-3})	(3.93×10^{-4})
Distance from own headquarters (000 km)	-0.04	-0.02
	(9.01×10^{-3})	(3.09×10^{-3})
City population (thousands)	1.26×10^{-4}	3.51×10^{-5}
	(8.57×10^{-6})	(3.09×10^{-6})
City GDP p.c. (thousands)	0.01	3.74×10^{-3}
	(5.69×10^{-4})	(2.03×10^{-4})
z_1(Open city 1993, dummy)	-0.13	-0.07
	(0.03)	(0.01)
z_2(Open city 2001, dummy)	-0.27	-0.03
	(0.02)	(0.01)
Provincial capital city (dummy)	0.45	0.16
	(0.02)	(0.01)
R^2	0.72	0.59
Observations	2,784	3,254

Note: Numbers in parentheses are standard errors.

Each additional McDonald's outlet reduces V^{OUT} for KFC by one-sixth (0.07/0.43) of a SD. It means that staying out of the city is less valuable for KFC if the rival has more outlets in the city. On the other hand, an additional KFC outlet in the city increases V^{OUT} for McDonald's by close to one-half (0.06/0.14) of a SD. This result suggests that the value of staying out, i.e. not entering, is higher for McDonald's if the rival has a greater presence in the city. The reason is that although more rival outlets reduce the entry-cost for McDonald's (Table 8), they also reduce its V^{IN} (Table 7) and the net result is a higher V^{OUT} for McDonald's. But then why do we see a higher probability of entry for McDonald's when there are more rival outlets in the city in the reduced-form regressions (Table A.1)? The structural estimates suggests that although V^{OUT} is higher for McDonald's when there are more rival outlets in the city, McDonald's is still more likely to enter because a greater number of own outlets in the city greatly increases V^{IN} (Table 7). Once again, the structural estimates rationalize the observed entry outcomes as optimal.

6 Concluding Remarks

We model the entry and expansion of KFC and McDonald's in the Chinese fastfood market as a dynamic game. We assume that the observed entry and expansion decisions are equilibrium outcomes. This allows us to recover the structural parameters of the game without solving for equilibrium.[29] We use the estimated model to study the dynamic spillovers from the rival.

We find that the rival's presence has significant and positive spillover effects on set-up cost for both chains. In addition, KFC's presence lowers the entry cost for McDonald's. On the other hand, a greater rival presence in the city affects the per outlet profit negatively for both chains. Taking the two effects together, the net effect is positive for KFC and negative for McDonald's. Each additional McDonald's outlet in the city adds about ¥3.8 million to KFC's value whereas an additional KFC outlet subtracts about ¥4.6 million from McDonald's value. As mentioned earlier, these inter-chain effects are 'dynamic spillovers' since they have summed up the current and future effects of an additional rival outlet on a chain's value.

The very large state space of our dynamic model is a mixed blessing. On the one hand, a rich set of state variables allows us to satisfy most of the equilibrium conditions during structural estimation and also provides better understanding of the determinants of profits, set-up costs and entry costs. On the other hand, it makes the computation of equilibrium practically impossible and hence we cannot do any counterfactual analysis. An interesting counterfactual could have been looking at the entry and expansion of one of the two chains in the absence of the other.

We pick the structural parameters that rationalize the data in the light of our model. If the fast-food chains act according to a model that is different from ours, our structural parameter estimates may not reflect the true entry or set-up costs faced by the chains. This is not a limitation of our paper alone, this is a limitation of structural estimation in general: the structural estimates are as credible as the theory on which they are based.

Another limitation of our analysis is that we cannot separately identify demand-side and supply-side spillovers to per outlet profit. This is because we do not have data on prices, quantities and costs. Future work could estimate separate demand and supply models and embed them into the period profit functions. Such extensions will be helpful to disentangle the two types of spillovers.

References

Ackerberg, D., Benkard, C. L., Berry, S. & Pakes, A. (2007), Econometric Tools for Analyzing Market Outcomes, *in* J. J. Heckman & E. E. Leamer, eds, 'Handbook of Econometrics', Vol. 6A, Elsevier Science, Amesterdam: North-Holland.

Aguirregabiria, V. & Magesan, A. (2012), Identification and Estimation of Dynamic Games when Players' Beliefs are not in Equilibrium. Working Paper 449, University of Toronto.

Aguirregabiria, V. & Mira, P. (2007), 'Sequential Estimation of Dynamic Discrete Games', *Econometrica* **75**(1), 1–53.

Bajari, P., Benkard, C. L. & Levin, J. (2007), 'Estimating Dynamic Models of Imperfect Competition', *Econometrica* **75**(5), 1331–1370.

Berry, S. & Reiss, P. (2007), Empirical Models of Entry and Market Structure, *in* M. Armstrong & R. H. Porter, eds, 'Handbook of Industrial Organization', Vol. 3, Elsevier Science, Amesterdam: North-Holland.

[29] Although the computational burden of BBL's estimation method is relatively low, it is nontrivial nonetheless. For example, the forward simulations to estimate the setup-cost parameters took about 14 days of CPU time and those to estimate the entry-cost parameters took a total of 307 days of CPU time. These simulations were run on HP Xeon Quad-core processors with CPU speeds of 1.7GHz each.

Berry, S. T. (1992), ‘Estimation of a Model of Entry in the Airline Industry’, *Econometrica* **60**(4), 889–917.

Bresnahan, T. F. & Reiss, P. C. (1991), ‘Entry and Competition in Concentrated Markets’, *Journal of Political Economy* **99**(5), 977–1009.

Holmes, T. J. (2011), ‘The Diffusion of Walmart and Economies of Density’, *Econometrica* **79**(1), 253–302.

Jia, P. L. (2008), ‘What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry’, *Econometrica* **76**(6), 1263–1316.

Mazzeo, M. (2002), ‘Product Choice and Oligopoly Market Structure’, *RAND Journal of Economics* **33**(2), 221–242.

Nishida, M. (2013), Estimating a Model of Strategic Network Choice: The Convenience-Store Industry in Okinawa. Working Paper.

Pakes, A., Ostrovsky, M. & Berry, S. (2007), ‘Simple Estimators for the Parameters of Discrete Dynamic Games (with entry/exit examples)’, *Rand Journal of Economics* **38**(2), 373–399.

Pesendorfer, M. & Schmidt-Dengler, P. (2008), ‘Asymptotic Least Squares Estimators for Dynamic Games’, *Review of Economic Studies* **75**(3), 901–928.

Rust, J. (1987), ‘Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher’, *Econometrica* **55**(5), 999–1033.

Seim, K. (2006), ‘An Empirical Model of Firm Entry with Endogenous Product-Type Choices’, *RAND Journal of Economics* **37**(3), 619–642.

Shen, Q. & Xiao, P. (2014), ‘McDonald’s and KFC in China: Competitors or Companions?’, *Marketing Science* **33**(2), 287–307.

Srisuma, S. (2010), Estimation of Structural Optimization Models: A Note on Identification. Working Paper, available at https://sites.google.com/site.tangsrisuma (last accessed on July 16, 2013).

Toivanen, O. & Waterson, M. (2005), ‘Market Structure and Entry: Where’s the Beef?’, *The RAND Journal of Economics* **36**(3), 680–699.

Toivanen, O. & Waterson, M. (2011), Retail Chain Expansion: The Early Years of McDonalds in Great Britain. Centre for Economic Policy Research, Discussion Paper No. 8534.

Yang, N. (2012), An Empirical Model of Industry Dynamics with Common Uncertainty and Learning from the Actions of Competitors. *Mimeo.*

Zhu, T., Singh, V. & Manuszak, M. D. (2009), ‘Market Structure and Competition in the Retail Discount Industry’, *Journal of Marketing Research* **46**(4), 453–466.

A Appendix Tables

Table A.1: First-stage policies (Entry)

Dependent variable: Enter = 1; Do not enter = 0		
Independent variable	KFC	McDonald's
1. Number of rival's outlets	0.15 (0.07)	0.11 (0.03)
2. Age	0.16 (0.06)	0.29 (0.06)
3. Age squared	0.001 (0.002)	-0.01 (0.00)
4. Distance from headquarters	-3.49×10^{-4} (8.69×10^{-5})	-4.50×10^{-4} (8.26×10^{-5})
5. Population	2.27×10^{-4} (6.92×10^{-5})	3.04×10^{-4} (5.88×10^{-5})
6. GDP per capita	0.02 (0.00)	0.02 (0.00)
7. Provincial Capital (dummy)	0.70 (0.18)	0.04 (0.19)
8. Number of rival's outlets × Open 1 (dummy)	-0.08 (0.26)	0.11 (0.11)
9. Number of rival's outlets × Open 2 (dummy)	-0.88 (0.59)	-0.22 (0.11)
R^2	0.26	0.21
Observations	2784	3254

Note: The regression model is Probit. The numbers in parentheses are standard errors.

Table A.2: First-stage policies (Expansion)

Dependent variable: Number of own new outlets opened in the city		
Independent variable	KFC	McDonald's
1. Number of own outlets	0.08 (0.01)	0.09 (0.02)
2. Number of own outlets squared	-9.82×10^{-4} (1.43×10^{-4})	-0.02 (0.00)
3. Number of own outlets cubed	3.26×10^{-6} (5.86×10^{-7})	9.81×10^{-6} (3.23×10^{-6})
4. Number of rival's outlets	0.04 (0.01)	-0.03 (0.02)
5. Number of rival's outlets sqaured	-1.14×10^{-3} (3.15×10^{-4})	-5.48×10^{-4} (2.46×10^{-4})
6. Number of rival's outlets cubed	7.97×10^{-6} (2.14×10^{-6})	2.18×10^{-6} (1.03×10^{-6})
7. Age	0.19 (0.07)	-0.03 (0.10)
8. Age squared	-0.01 (0.00)	-0.01 (0.01)
9. Distance from headquarters	-3.52×10^{-4} (8.57×10^{-5})	8.37×10^{-5} (1.11×10^{-4})
10. Population	9.53×10^{-5} (2.77×10^{-5})	1.10×10^{-4} (4.45×10^{-5})
11. GDP per capita	-0.01 (0.00)	0.01 (0.00)
12. Provincial Capital (dummy)	0.54 (0.11)	0.62 (0.17)
Over-dispersion parameter (χ^2)	151.62	180.86
Over-dispersion parameter (p-value)	0.000	0.000
Pseudo R^2	0.21	0.18
Observations	1258	788

Note: The regression model is Negative Binomial. The numbers in parentheses are standard errors except for the over-dispersion parameter, for which the number in parentheses are p-values of the χ^2 distribution.

Table A.3: Discretization of GDP per capita (in thousand yuan)

State	Range in the data	Discretized value
1	$0 < y < 4$	2
2	$4 \leq y < 8$	6
3	$8 \leq y < 12$	10
4	$12 \leq y < 16$	14
5	$16 \leq y < 20$	18
6	$20 \leq y < 30$	25
7	$30 \leq y < 50$	40
12	$50 \leq y$	60

Table A.4: Discretization of Population (in thousands)

State	Range in the data	Discretized value
1	$0 < P < 250$	125
2	$250 \leq P < 500$	375
3	$500 \leq P < 750$	625
4	$750 \leq P < 1,000$	875
5	$1,000 \leq P < 1,500$	$1,250$
6	$1,500 \leq P < 2,000$	$1,750$
12	$2,000 \leq P$	$3,000$

Table A.5: Estimated transition matrix for GDP p.c. (thousand yuans)

Current\Next	0-4	4-8	8-12	12-16	16-20	20-30	30-50	>50
0-4	**0.69**	0.29	0.01	0.00				
4-8	0.02	**0.75**	0.21	0.02	0.00	0.00	0.00	
8-12	0.00	0.03	**0.72**	0.24	0.01	0.00	0.00	
12-16	0.00	0.00	0.03	**0.64**	0.31	0.02		
16-20		0.00	0.01	0.05	**0.53**	0.40	0.00	0.00
20-30			0.01	0.02	0.03	**0.74**	0.21	
30-50					0.00	0.02	**0.83**	0.14
>50							0.05	**0.95**

Note: Row sums may not be equal to 1 due to rounding.

Table A.6: Estimated transition matrix for population (thousands)

Current\Next	0-250	250-500	500-750	750-1000	1000-1500	1500-2000	>2000
0-250	**0.81**	0.17	0.01			0.01	
250-500	0.00	**0.94**	0.04	0.00	0.01	0.01	0.00
500-750	0.00	0.01	**0.93**	0.04	0.01	0.00	0.00
750-1000		0.00	0.01	**0.94**	0.04		0.00
1000-1500	0.00	0.00	0.01	0.01	**0.94**	0.04	0.01
1500-2000		0.01			0.01	**0.93**	0.04
>2000						0.00	**1.00**

Note: Row sums may not be equal to 1 due to rounding.

Table A.7: Stage-two structural parameter estimates

	KFC		McDonald's	
Parameter	Estimate	Standard Error	Estimate	Standard Error
$\theta^{\pi}_{i,0}$	17.9	Normalization	9.3	Normalization
θ^{π}_{i,n_i}	-6.85×10^{-3}	(1.19×10^{-3})	-9.15×10^{-3}	(1.37×10^{-3})
θ^{π}_{i,n_i^2}	7.47×10^{-6}	(1.50×10^{-6})	9.39×10^{-6}	(2.05×10^{-6})
θ^{π}_{i,n_j}	-2.09×10^{-3}	(2.53×10^{-4})	-4.95×10^{-5}	(2.67×10^{-4})
θ^{π}_{i,n_j^2}	1.78×10^{-6}	(4.92×10^{-7})	-5.19×10^{-7}	(6.21×10^{-7})
θ^{π}_{i,a_i}	-0.70	(0.01)	-0.18	(0.03)
θ^{π}_{i,a_i^2}	7.48×10^{-3}	(1.87×10^{-4})	1.13×10^{-3}	(3.33×10^{-4})
θ^{π}_{i,d_i}	1.82×10^{-4}	(4.67×10^{-5})	5.39×10^{-5}	(3.14×10^{-5})
$\theta^{\pi}_{i,p}$	-1.15×10^{-5}	(6.93×10^{-6})	1.20×10^{-6}	(3.51×10^{-6})
$\theta^{\pi}_{i,y}$	7.95×10^{-6}	(2.45×10^{-4})	-2.74×10^{-4}	(1.19×10^{-4})
θ^{π}_{i,z_1}	-0.47	(0.07)	-0.11	(0.04)
θ^{π}_{i,z_2}	-0.06	(0.08)	-0.09	(0.03)
$\theta^{\pi}_{i,k}$	0.31	(0.07)	-1.57	(0.51)
$\theta^{s}_{i,0}$	109.86	(0.99)	69.81	(2.07)
θ^{s}_{i,n_i}	-0.12	(0.02)	-0.17	(0.03)
θ^{s}_{i,n_i^2}	1.76×10^{-4}	(3.91×10^{-5})	2.54×10^{-4}	(4.92×10^{-5})
θ^{s}_{i,n_j}	-0.02	(3.48×10^{-3})	-2.32×10^{-3}	(2.26×10^{-3})
θ^{s}_{i,n_j^2}	2.73×10^{-5}	(7.99×10^{-6})	1.18×10^{-6}	(5.68×10^{-6})
θ^{s}_{i,a_i}	-4.92	(0.08)	-1.42	(0.20)
θ^{s}_{i,a_i^2}	0.06	(1.24×10^{-3})	0.01	(2.59×10^{-3})
θ^{s}_{i,d_i}	1.80×10^{-3}	(4.30×10^{-4})	4.65×10^{-4}	(2.79×10^{-4})
$\theta^{s}_{i,p}$	-1.35×10^{-4}	(5.67×10^{-5})	-2.41×10^{-5}	(3.08×10^{-5})
$\theta^{s}_{i,y}$	-5.77×10^{-3}	(1.33×10^{-3})	-6.23×10^{-3}	(9.86×10^{-4})
θ^{s}_{i,z_1}	-4.06	(0.59)	-0.92	(0.33)
θ^{s}_{i,z_2}	-0.53	(0.69)	-0.82	(0.27)
$\theta^{s}_{i,k}$	2.47	(0.61)	-14.17	(4.56)
σ^{s}_{i}	0.53	(0.17)	0.52	(0.36)
EC[1]	125,800		78,800	
ECS[2]	84.9%		77.3%	

[1]EC = Equilibrium conditions

[1]ECS = Equilibrium conditions satisfied

Note: Standard errors are bootstapped.

B Data on Operating Income

We have compiled the data on operating income from a number of sources. The data for the years 2003-2005 are from *Statistical Yearbook of China Restaurants* (Issues 2004, 2005 and 2006). The data for years 2006 and 2007 are from *Statistical Yearbook of China Chain Stores of Retail Trades and Catering Services* (Issues 2007 and 2008). Instead of reporting a precise number for the operating income, the yearbooks report a range. We take the average of the lower and upper limits of the range as our measure of operating income. We expand our dataset by using information from the following online sources: 1)`http://www.docin.com/p-69630763.html`; 2)`http://www.baiyan.cc/newsInfo.asp?typenumber=0002&id=106`; 3)`http://news.sohu.com/20060902/n245129062.shtml`; 4)`http://finance.qq.com/a/20090327/000223.htm`; and 5)`http://www.canyin168.com/glyy/glzx/hyfx/201103/28410.html`. We convert the data into constant 2001 yuans by using the Consumer's Price Index. We report some summary statistics in Table B.1 below.

Table B.1: Annual operating income (2001 prices, million yuans)

	KFC	McDonald's
Per City		
Mean	328.57	83.04
S. D.	489.83	141.36
Per Outlet		
Mean	17.91	9.26
S. D.	20.29	5.44
Observations	84	170

Our dataset on operating income primarily covers big cities. For example, in our full sample the average number of outlets per city for the 2003-2007 period is 4.9 for KFC and 2.9 for McDonald's. If we restrict the sample to cities for which we have data on operating income, the average number of outlets per city becomes 21.3 for KFC and 9.3 for McDonald's. Consequently, the static profits that we estimate may not be representative of the average profits in the entire country. One way to see this is to use country-level sales figures for the two chains, assume a reasonable mark-up and compute country-level profits. We can then compare them with the country-level profits implied by our estimates. According to The Wall Street Journal (2013) China's fastfood sales were $47 billion in 2007. The market shares of KFC and McDonald's were 6.5% and 2.3%, respectively. This leads to the sales of about $3.055 billion for KFC and $1.081 billion for McDonald's. If one assumes a markup of 30% on cost, the implied profits are $0.705 billion for KFC and $0.249 billion for McDonald's. We estimate average profit per outlet to be Y17.234 million for KFC and Y10.228 million for McDonald's. The exchange rate between yuan and dollar was 7.5910 yuan per dollar on July 2, 2007 (Source: www.finance.yahoo.com). Using this exchange rate our estimated average profit per outlet works out to be $2.270 million for KFC and $1.347 million for McDonald's. These translate into country-level estimated profits of $4.386 billion (= $2.270m × 1932 outlets) for KFC and $1.363 billion (= $1.347m × 1012 outlets). These simple back-of-the-envelope calculations show that we might be over-estimating the static profits by a factor of 6.2 (4.386/0.705) for KFC and by a factor of 5.5 (1.363/0.249) for McDonald's.