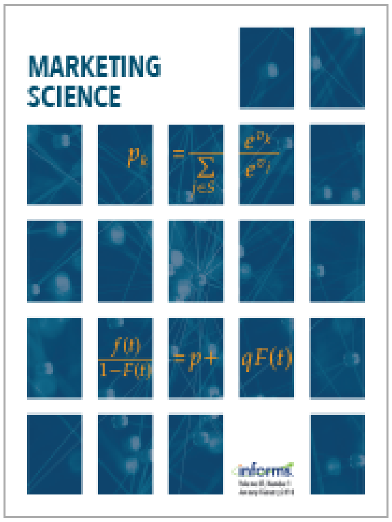
This article was downloaded by: [165.123.228.21] On: 30 March 2021, At: 17:30
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA



Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

McDonald's and KFC in China: Competitors or Companions?

Qiaowei Shen, Ping Xiao

To cite this article:
Qiaowei Shen, Ping Xiao (2014) McDonald's and KFC in China: Competitors or Companions?. Marketing Science 33(2):287-307. https://doi.org/10.1287/mksc.2013.0824

Full terms and conditions of use: https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2014, INFORMS

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

MARKETING SCIENCE
Vol. 33, No. 2, March–April 2014, pp. 287–307
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2013.0824
© 2014 INFORMS

McDonald's and KFC in China: Competitors or Companions?

Qiaowei Shen
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, qshen@wharton.upenn.edu

Ping Xiao
NUS Business School, National University of Singapore, Singapore 119245, bizxp@nus.edu.sg

In this paper, we study the entry and expansion decisions of McDonald's and KFC in China using an originally assembled data set on the two chains' expansion in the China market from their initial entry up to year 2007. We analyze how the presence of a rival affects each firm's strategies. The results indicate that a rival's presence has a net positive effect on a chain's expansion decision. We focus on testing two possible explanations for a positive rival impact: market learning and demand expansion. First, we derive a set of theoretical predictions on how a chain's optimal expansion decision would react to its rival's expansion patterns when market learning versus demand expansion is the driving force of the rival's positive influence. The empirical analysis based on these predictions consistently suggests that market learning is more likely to explain the positive effect of KFC on McDonald's and that demand expansion is more plausible with McDonald's positive spillover on KFC. In other words, the results are consistent with the presence of KFC signaling market demand potential and growth to McDonald's and the presence of McDonald's helping to cultivate consumer taste and generate demand for Western fast food, which benefits KFC.

Keywords: entry; market learning; demand expansion; emerging market; fast-food chains
History: Received: December 27, 2011; accepted: October 17, 2013; Preyas Desai served as the editor-in-chief and Michel Wedel served as associate editor for this article. Published online in *Articles in Advance* January 17, 2014.

1. Introduction

The decision to enter an emerging market characterized by high demand potential and high risk is always intriguing: When to enter? Where to enter? How many stores to open? These are important decisions faced by international firms that are interested in expanding their businesses in emerging markets. In this paper, we study the entry and expansion decisions of McDonald's and KFC in China, especially how the two players have influenced each other.

Over the past two decades, these two major fast-food chain restaurants have flourished in the world's largest emerging market. Although most Chinese consumers in big cities are now familiar with the two American brands, the concept of Western fast food was novel when KFC opened its first outlet in Beijing in November 1987.[1] By 2007, however, KFC had entered 240 cities with 2,000 outlets and McDonald's had entered approximately 140 cities with more than 1,000 outlets. In this fast expansion process, it is intriguing to ask how the presence of a rival affected each firm's strategies—specifically, decisions on whether to enter a new market or how many new outlets to open in an existing market.

On the one hand, KFC and McDonald's are considered direct competitors in the Western fast-food category in the China market. On the other hand, there could be positive spillover effects from the rival in this specific context. For example, although China, with its large population and remarkable economic growth, is an attractive market, it also presents substantial uncertainty to foreign companies. Therefore market learning through both one's own information and the information revealed from a rival's action could be important. In addition, traditional Chinese food is distinct from Western food, and it takes time for local consumers to accept typical Western food such as hamburgers. The presence of the rival chain could facilitate consumer learning of the new product and expand the potential demand that benefits all the firms in the same category.

We assembled original data on the entire entry and expansion history of KFC and McDonald's in China from 1987 to 2007 through multiple sources. The data include the number of new outlets KFC and McDonald's opened each year in each city and the local characteristics of these cities. Our empirical analysis of chains' expansion decisions reveals that both

[1] Both McDonald's and KFC opened their restaurants in Hong Kong before entering mainland China. The China market discussed in this paper thus excludes Hong Kong.

McDonald's and KFC are more likely to increase the number of outlets in a city where their rival has a larger scale of presence, controlling for city characteristics and own-chain effect. In other words, the rival's presence has a positive and significant effect that outweighs the competition effect in this setting.

The main objective of the paper is to empirically test which potential source is more likely to be the driving force of the positive rival effect in this context, and we focus on two potential sources: (1) market learning (i.e., the entry behavior of a rival may reveal valuable information about the market size and growth rate that a firm can utilize in its subsequent decisions) and (2) demand expansion (i.e., the existence of a rival chain may help to cultivate consumers' taste for Western-style fast food and expand the potential demand). We develop a theoretical framework about a chain's optimal expansion decision that predicts that a chain would react differently to the rival's expansion pattern depending on whether it infers information from the rival's actions or benefits from the rival's role in cultivating consumers. The empirical analysis based on these predictions suggests that the positive rival effect for KFC and McDonald's is likely to be driven by different sources. More specifically, we find that market learning is more likely to explain the positive effect that KFC exerts on McDonald's, and demand expansion is more likely to explain McDonald's positive spillover on KFC.

Although, in essence, our research is on firms' entry and expansion decisions, the uniqueness of the study lies in the context of decision making. First, it happens as the two U.S.-based firms are exploring a large emerging market, which requires substantial learning from the firm side. Second, firms' expansion history in this market is also a process that builds up a new industry category, which is considerably different from entry in a mature market or industry. The context of two global fast-food chain restaurants expanding in the world's largest emerging market simultaneously provides us with a unique opportunity to potentially identify the positive spillover effects in the course of firms' interactions.

Meanwhile, this study has important managerial implications for multinational firms that are interested in expanding their business in an emerging market. Previous research on emerging market entry has largely focused on the industry level. A large literature in the international economics field has studied the entry mode, performance, and impact of foreign direct investment on the emerging economy (e.g., Borensztein et al. 1998, Isobe et al. 2000, Javorcik 2004). We take a more microlevel perspective by modeling firms' entry and expansion behavior based on individual firm decision making. We focus on the dynamic interactions between the entering firms in their exploration of the emerging market, an area that still awaits more rigorous studies. Our research shows that the competing firms may well be "companions" when expanding in an emerging market with significantly different market environments and consumer preference.

Our work is closely related to Toivanen and Waterson (2005), who, using McDonald's and Burger King's store-opening data in the United Kingdom between 1991 and 1995, find that a rival's presence increases the probability of entry. This finding reverses the prediction of the standard entry model, which holds that all else being equal, a firm will enter the market without a competitor if it can choose between two comparable markets. Toivanen and Waterson attribute the positive effect to learning; i.e., a firm learns the market size through its rival's presence. Yang (2013) further explores the role of learning in fast-food chain agglomerations by estimating a structural model of entry and exit decisions under common market uncertainty, using data from the fast-food industry in Canada. He finds that fast-food chains face significant uncertainty before entry and that learning from the incumbent chains is a key driver of fast-food clustering.

This paper also relates to the increasing empirical work on the (positive) spillover effects of entry. Using data on U.S. regional shopping centers, Vitorino (2012) finds complementarity among certain types of department stores. Mariotti et al. (2010) study the role of information externalities and knowledge spillovers for the agglomerative behavior in location choices of multinational enterprises in Italy. In a study of Japanese multinational corporations, Henisz and Delios (2001) show that the prior entry decisions of other organizations in the same business group have a greater influence on firms that lack experience in the market. Our work contributes to this literature by identifying different types of positive spillover effects with entry in an emerging market.

Finally, this paper also adds to the literature on the fast-food industry in general. Sault et al. (2006) investigate within-market location decisions of McDonald's and Burger King in the United Kingdom and find that learning from competitors along with product differentiation can explain the observed location patterns of these two players. Thomadsen (2007) looks at the two companies' location strategies in the U.S. market and finds that the equilibrium locations of the two fast-food players depend on the market size. Our research provides additional insights into how major fast-food chains make entry and expansion decisions in an emerging market. To the best of our knowledge, this is the first paper that investigates fast-food chains' entry decisions and the spillover effects in an emerging market.

The rest of this paper is organized as follows. We discuss the industry background and the data in §2. In §3, we present a theoretical framework on entry and generate a set of testable hypotheses under each explanation of positive rival effect. In §4, we estimate a two-stage model that links a chain's initial entry and expansion decision and test which explanation is more likely to be the driving force of positive rival effect based on the hypotheses derived from the theoretical framework. We further discuss the variation of a rival's impact in §5 and conclude in §6.

2. Industry Background and Data

2.1. The Development of KFC and McDonald's in China

The presence of Western fast food in China began in November 1987, when KFC opened its first outlet in Beijing. Following in KFC's footsteps, McDonald's (McD's) opened its first outlet in Shenzhen in October 1990. Thus the competition between KFC and McD's formally started in the world's largest emerging market. On a global scale, KFC and McD's are not competitors at the same level; McD's is far ahead of KFC in terms of its number of outlets. However, the picture is completely reversed in the China market.

By tracking KFC's entry path, we find that KFC typically enters the capital city of a province first and then expands to nearby cities. KFC chose China's capital city, Beijing, as the home of its first outlet and quickly spread to the east and south coast, which is economically more developed than the rest of China. Then KFC steadily expanded to the north, the central area, and the west. It now has a footprint in all the provinces in China except Tibet.

In contrast, McD's expansion has been relatively more conservative. McD's opened its first outlet in Shenzhen, a city in the Guangdong province that is next to Hong Kong, where McD's was headquartered at the time. In its first few years there, McD's limited its entry to only a few big cities and the cities in the southeastern part of China that are close to Shenzhen. It accelerated its expansion rate from 1999 but was dwarfed by its rival's even faster pace. By the end of 2007, McD's still had no presence in seven provinces, including Tibet.

During their 20 years of growth in China, the two chains not only expanded to dozens of new cities per year but also continuously increased the number of outlets in the cities they had already entered. Such expansion breadth and depth has contributed to the fast increase in the total number of outlets for both chains. Figure 1 depicts the diffusion of the number of cities entered by KFC and McD's. The pattern is close to the typical S-shaped diffusion, with the "takeoff" occurring around 1999. Figure 2 shows the cumulative number of outlets over time for each chain.

Figure 1 First-Outlet Diffusion

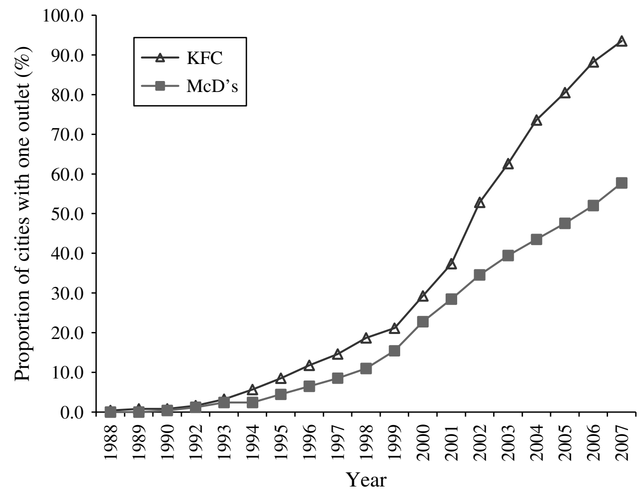
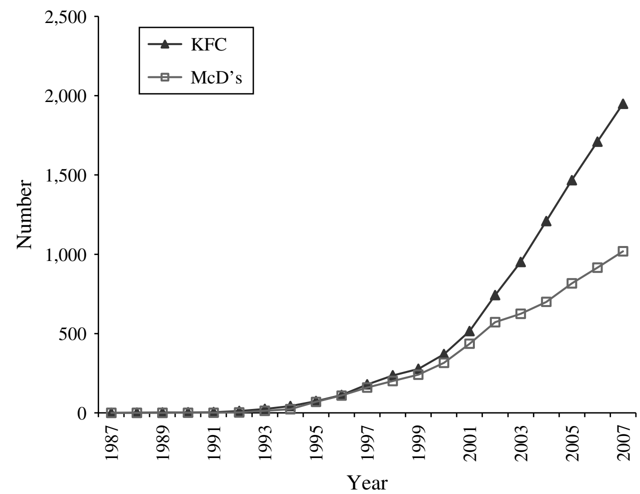


Figure 2 The Cumulative Number of Outlets Over Time



Although more players have been appearing in China's fast-food market over time, we believe it is reasonable to exclude other fast-food firms in our analysis because none of them had a significant national impact by the end of our sample period. Other foreign fast-food brands entered China much later and have a much smaller presence.[2] In terms of the local Chinese fast-food industry, despite fast growth over the years, the scale of national Chinese fast-food chains is still small within the overall fast-food industry. Moreover, the competition between the Western fast-food chains and the local Chinese fast-food chains is much weaker than the competition between the two Western chains.

[2] Subway entered China in 1995 but had only approximately 150 outlets by the end of 2007. Burger King was not present in China until June 2005 and had fewer than 50 outlets by the end of 2007.

2.2. Data and Statistics

Our data track the opening year and location of each outlet of KFC and McD's at the city level from 1987 to the end of 2007. For each city, we know when KFC and/or McD's opened their first outlet and how many additional outlets were opened each year after initial entry. The data set was constructed from both publicly sourced and hand-collected information. We obtained the outlet lists including outlet names, telephone numbers, and addresses for both KFC and McD's through the companies' official websites and online restaurants' guide. Using a local marketing research company, we obtained each individual outlet's opening year information.[3] To cross-check the accuracy of the information, we obtained outlet registration information from the Industry and Commercial Administration Bureau in a few cities that we have access to. The information from the two sources is highly consistent, which ensures the quality of the data. We also searched the milestones of the two fast-food chains in China through multiple sources and compared them with our data, e.g., the total number of outlets in China by year 2000. The results are again largely consistent. Our final data cover about 98% of the official total outlets for both KFC and McD's by 2007.

We collected city-level characteristics from volumes of the *Chinese Cities Yearbook* spanning 1989 to 2008.[4] The cities reported in the yearbooks vary in some years, but we are able to consistently identify 246 major cities across years. When merging the two data sets, we found that by the end of the sample period, 236 of the 246 cities witnessed the entry of at least one of the fast-food chains.[5] For each of the 246 cities in the data, we observe the population size (in 10,000's), geographical area (in square kilometers), and gross domestic product (GDP, in 100 million yuan (RMB)) each year.[6] We also collected the longitude and latitude information for each city, which allows us to track the distance between each city and the firms' headquarters or any other city in the network. KFC has been headquartered in Shanghai since it first entered the China market. The headquarters of McD's was in Hong Kong initially but moved to Shanghai in 2004.

The descriptive statistics of city characteristics are summarized in Table 1. We further categorize the sample into big and small cities. Big cities include municipalities directly under the administration of the central government, special autonomous regions, the capital city of each province, the 14 coastal cities with an open economy, cities in special economic zones that enjoy more favorable policies for foreign investors, and other cities with a GDP over 20 billion RMB and population over 2 million. These cities are economically more developed or have special economic and political status, which account for 25.3% of the sample, with the rest categorized as small cities. The descriptive statistics of sample breakdown for big and small cities are summarized in the middle part of Table 1. The data reveal that, on average, big cities are more densely populated and economically better developed, although big cities are not necessarily larger geographically. The bottom part of Table 1 breaks down the sample to the pre-1999 and post-1999 eras. The year 1999 seems to be the takeoff year for both KFC and McD's, when the two chains started to speed up their expansion in China. A notable difference for the two sample periods is the substantial increase in GDP after 1999.

Table 2 summarizes the descriptive statistics of the number of new outlets opened each year in each city by KFC and McD's by different sample breakdowns. On average, KFC adds 0.39 new outlets per year per city, with a much higher number in big cities and during the second period (1999–2007). A similar pattern is observed for McD's but on a smaller scale. Moreover, KFC has more experience (more years of operation) than McD's, on average, across cities. It is also evident that chains enter big cities earlier than the small cities, on average. Table 3 indicates the total number of cities entered for each chain and the cumulative number of outlets by the end of the sample period. Before 1999, both KFC and McD's entered about three new cities and added about two dozen new outlets nationally each year on average. Both chains expand much faster after 1999. During this period, McD's entered roughly 13 new markets each year, adding 90 new outlets per year. KFC on average entered 20 new cities and opened 187 new outlets each year. For both chains, the practice of franchising is very limited in the China market in our sample period.[7]

[3] There is missing information for 39 KFC outlets and 24 McD's outlets. Since the missing data only account for a small percentage of the total number of outlets, and there is no special pattern for missing data, we exclude them from our formal analysis. The stores with missing information may include exits.

[4] The information reported in each yearbook includes the statistics from the previous year. The early years of the city yearbook are only available in hardcopy, and we manually input the city characteristics information for those years. Also, we were not able to find the yearbook for 1988. Therefore, the formal analysis excludes the first entry of KFC in Beijing in 1987.

[5] There are 11 cities that KFC or McD's has entered, but we fail to match these cities with city characteristics. These are all small towns (not identified by the yearbook as a "city") that the chains entered recently, which are excluded from our analysis. Altogether, there are 18 KFC outlets and three McD's outlets in those 11 cities.

[6] The yearly GDP figures are adjusted for inflation, all in 2005 RMB.

[7] KFC opened its first franchised store in 2000, and by the end of 2005, there were 37 franchised stores in total, which account for less than 5% of KFC's total number of restaurants (see *Shanghai Daily* 2006). McD's launched a pilot franchise program in 2004 and has only six franchised restaurants by early 2010 (see Kwok 2010). Overall, franchised stores account for a small percentage of the total stores of both chains in our sample period.

Table 1 **Descriptive Statistics**

Sample	Variable	Mean	SD	Minimum	Maximum
All	*Population*	107.81	125.82	5.94	1,290.14
	Area	1,651.22	2,134.15	25.00	22,341
	GDP	196.14	484.26	0.66	9,651.97
	Distance to KFC's headquarters	1,009.06	533.83	0.00	3,818.46
	Distance to McD's headquarters	1,204.28	681.73	0.00	3,818.46
Big city	*Population*	204.45	207.30	17.53	1,290.14
	Area	1,588.39	1,643.30	77.00	12,484
	GDP	506.73	832.37	8.17	9,651.97
	Distance to KFC's headquarters	970.99	630.77	0.00	3,499.91
	Distance to McD's headquarters	1,086.55	678.49	0.00	3,499.91
Small city	*Population*	73.53	41.53	5.94	276.48
	Area	1,673.51	2,282.88	25	22,341
	GDP	80.88	96.28	0.66	1,441.86
	Distance to KFC's headquarters	1,021.88	496.36	171.64	3,818.46
	Distance to McD's headquarters	1,243.95	678.33	140.33	3,818.46
Pre-1999	*Population*	94.08	101.31	5.94	1,018.59
	Area	1,521.36	2,087.51	25	22,341
	GDP	88.97	165.84	2.71	2,853.56
	Distance to KFC's headquarters	763.17	469.21	0.00	1,778.47
	Distance to McD's headquarters	962.41	708.87	0.00	2,667.75
Post-1999	*Population*	123.64	147.60	14.08	1,290.14
	Area	1,801.03	2,177.59	50.00	19,576
	GDP	303.22	655.11	6.98	9,651.97
	Distance to KFC's headquarters	1,004.04	534.26	0.00	3,813.46
	Distance to McD's headquarters	1,032.51	569.32	0.00	2,721.84

Note. Population in 10,000; area in square kilometers; GDP in 100 million RMB; and distance to headquarters in kilometers.

Figures 3(a) and 3(b) show the histograms of the total number of outlets by city at the end of the sample period for the two chains. The long tail of the histogram reflects the significant heterogeneity across markets. Among the 246 cities in our data set, KFC is the first mover in 187 cities and McD's is the first mover in 33 cities. Among the 33 cities where McD's took the initiative and entered before KFC, 20 are small cities. There are 16 cities that KFC and McD's entered in the same year; they are all small cities.

Table 2 **Descriptive Statistics on Number of New Outlets per Year and Experience**

Variable	Sample	Mean	SD	Minimum	Maximum
KFC's new outlets per year	All	0.39	1.63	0	33
	Big	1.26	3.02	0	33
	Small	0.10	0.36	0	4
	Pre-1999	0.09	0.60	0	15
	Post-1999	0.77	2.28	0	33
McD's new outlets per year	All	0.21	1.06	0	27
	Big	0.72	2.00	0	27
	Small	0.03	0.20	0	3
	Pre-1999	0.07	0.72	0	19
	Post-1999	0.37	1.35	0	27
KFC's experience	All	1.46	2.91	0	20
	Big	3.40	4.29	0	20
	Small	0.81	1.85	0	13
McD's experience	All	0.92	2.39	0	18
	Big	2.50	3.74	0	18
	Small	0.39	1.34	0	12

Table 3 **Descriptive Statistics on Expansion**

Variable	Sample	KFC	McD's
Total number of outlets at end of sample period	All	1,932	1,012
	Big city	1,565	890
	Small city	367	122
	Pre-1999	237	200
	Post-1999	1,695	812
Number of cities entered	All	230	142
	Big city	62	57
	Small city	168	85
	Pre-1999	46	27
	Post-1999	184	115

3. Theoretical Framework

We start with a simple model that captures the key considerations in firms' expansion decision. We assume that the decision of how many new outlets to open depends on the revenue and cost associated with the additions. First, firm i's (expected) revenue (R_{it}) from its stores is formulated as [8]

$$R_{it} = (n_{it} + S_{it})[M_i^e(X_t, H_{it}, H_{jt}) - (a_1 n_{it} + a_2 S_{it}) - (b_1 n_{jt} + b_2 S_{jt})], \quad (1)$$

where n_{it} is the number of new outlets to be opened in period t and S_{it} is the stock of outlets that chain i owns at the beginning of period t in the market. If

[8] For brevity of notation, we suppress the m subscript for market in the equations in this section.

Figure 3(a) **Distribution of McD's Outlets in 2007**

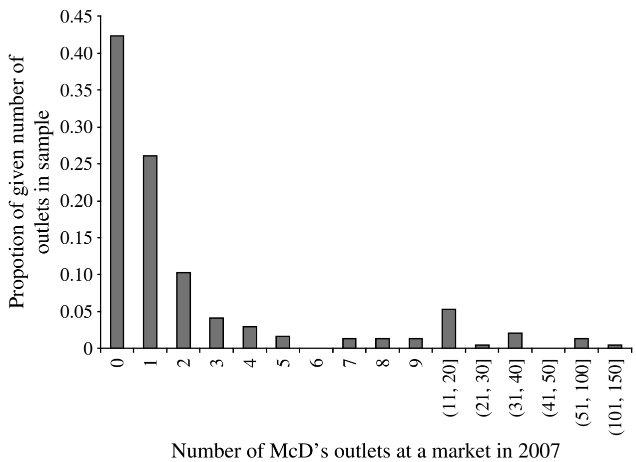
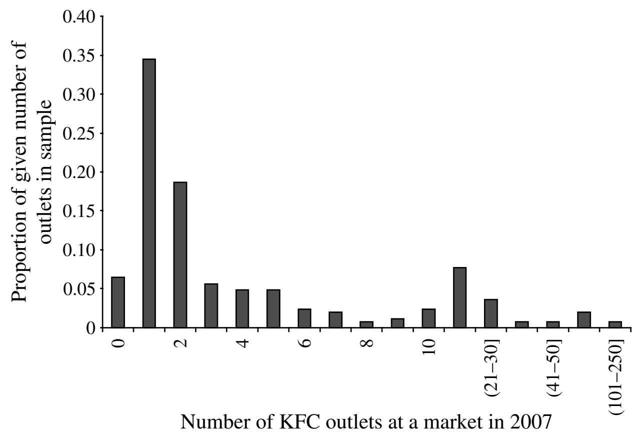


Figure 3(b) **Distribution of KFC Outlets in 2007**



the number is 0, it means the firm has not entered the market yet. The terms in square brackets can be interpreted as the average revenue per outlet, which depends on the following: (1) the expected market size for chain i, M_i^e; (2) the cannibalization of outlets from the same chain, including the existing stores before period t and the newly opened stores, with $a_1 \geq 0$ and $a_2 \geq 0$; and (3) the competition effect from the rival chain, $b_1 \geq 0$ and $b_2 \geq 0$. We use M_i^e to denote firm i's (subjective) evaluation of the market size or demand at period t, which is a function of the market characteristics, X_t, and firm i's and the rival's past behaviors, summarized by H_{it} and H_{jt}, respectively. We assume that H_{it} and H_{jt} can positively affect demand evaluation; we discuss the mechanisms below.

Second, the total cost of maintaining the existing outlets and opening new ones is modeled as a linear function:

$$C_{it} = K_{it} \cdot 1\{S_{it} = 0\} + \theta S_{it} + V(e_{it}, n_{it}), \qquad (2)$$

where $1\{S_{it} = 0\}$ is an indicator function that has a value 1 if the firm has not entered the market before time t yet and 0 otherwise. Therefore K_{it} is the initial entry cost, which could change over time with, for example, the expansion of the chain's network. The second term is the operating cost of the existing outlets. The third term captures the opening and operating costs of new outlets, which could depend on the experience of chain i in the market, denoted by e_{it}.

Firm i's decision problem is to choose the optimal number of outlets to open (n_{it}^*) to maximize the profit conditional on her expectation on the rival's decision:

$$n_{it}^* = \arg\max_{n_{it}} (R_{it} - C_{it}). \qquad (3)$$

To simplify the problem, we assume that firm i uses the rival's new outlets opened in the last period (n_{jt-1}) to form her expectation of n_{jt}. The first-order condition of the profit maximization problem satisfies

$$\phi(\cdot) = [M_i^e(X_t, H_{it}, H_{jt}) - (a_1 n_{it} + a_2 S_{it}) - (b_1 n_{jt-1} + b_2 S_{jt})] - a_1(n_{it} + S_{it}) - \frac{\partial V}{\partial n_{it}} = 0. \qquad (4)$$

It suggests that the optimal number of new outlets would be such that the marginal revenue of a new store (in square brackets) equals the marginal "cost," which includes both the cannibalization effect and marginal store opening cost (the last two terms). We are particularly interested in how the rival's presence, as summarized by S_{jt}, would affect firm i's entry and expansion decision. Using the implicit function theorem, one can show that $\text{sgn}(\partial n_{it}^* / \partial S_{jt}) = \text{sgn}(\partial M_i^e / \partial S_{jt} - b_2)$ under some mild assumptions. It suggests that the net rival effect depends on the magnitude of the positive rival effect on the expected size of demand relative to the competition effect.

We specify the expected market size (M_i^e) as a function of market characteristics X_t, a firm's own past expansion summarized by $M_o^e(H_{it})$, and the rival's impact denoted by $M_r^e(H_{jt})$:

$$M_i^e(X_t, H_{it}, H_{jt}) = \beta X_t + M_o^e(H_{it}) + M_r^e(H_{jt}). \qquad (5)$$

One's own presence in a city may positively affect demand through consumer learning or brand loyalty. The presence of the rival may also have a positive effect for the following reasons: (1) the existence of the rival chain may cultivate local consumers' taste for the Western-style fast food, which expands the potential demand for the new product and benefits the focal chain (*demand expansion effect*);[9] and (2) the existence and expansion of the rival may signal the

[9] Agarwal and Bayus (2002) empirically show that new firm entry is a significant factor in explaining sales takeoff. Kuksov and Villas-Boas (2010) investigate consumers' search behavior when the product alternatives offered to consumers span the preference space.

size and growth of the market (*market learning effect*). Although there might be other sources of the positive rival effect, we focus on these two leading explanations that have appeared in the literature (e.g., Caplin and Leahy 1998, Toivanen and Waterson 2005). We now look at each explanation in greater detail.

Demand Expansion

Traditional Chinese food is distinct from Western food, and it takes time for local consumers who are used to oriental food to accept typical Western food such as hamburgers. The entry of the rival chain may help to cultivate local Chinese consumers' taste and facilitate consumer learning of the new product. We assume that both the scale of the Western fast-food chains and the time they have been in existence would matter in cultivating consumers' taste for Western fast food. Long-existing outlets (those outlets opened earlier) would have a stronger cumulative effect in consumer taste cultivation relative to the stores added more recently.[10] This implies the following: (1) A larger scale of rival presence suggests a higher degree of cultivation. (2) The time span since the rival's store opening has a positive effect on cultivating consumers' taste. (3) There is a positive interaction effect of the scale and length of outlets' existence. (4) The sequence of store opening matters. Conditional on the total number of outlets, more outlets opened in earlier years may convert to a higher stock of cultivation.

Market Learning

Another plausible explanation for why a rival's past actions may affect a firm's evaluation of market size is market learning or information externality (Caplin and Leahy 1998). When a multinational company enters a foreign market, an effective way of reducing uncertainties or information asymmetries is via observing competitors' past entry decisions in the same market.[11] Learning may continue in the expansion stage after initial entry given the dynamics and uncertainty inherent in emerging markets.

A firm's evaluation on market size could be decomposed into the assessment of the existing demand and the expectation of market growth. The rival's entry behavior may reveal information about both components.[12] The scale of a rival's presence (S_{jt}) can signal the size of the current market. A firm may further learn about demand growth from how fast its rival grows. We use the rival's average expansion rate (S_{jt}/e_{jt}) to capture the first-order rival's growth. Furthermore, we use the rival's net growth rate in the past L periods, which can be defined as $g_{jl} = n_{j(t-l)} - n_{j(t-l-1)}$ ($l = 1, 2, \ldots, L$), to capture the second-order rival's growth. If a rival's faster growth signals faster demand growth, it suggests that (1) conditional on the total number of a rival's outlets (S_{jt}), a higher average expansion rate (larger S_{jt}/e_{jt}) signals higher demand growth—that is, the same scale of a rival's expansion happening within a shorter time span suggests stronger demand growth; or (2) a rival's positive net growth ($g_{jl} > 0$) suggests faster demand growth—that is, the accelerating trend of a rival's expansion implies faster demand growth.

Because both the demand expansion effect and the market learning effect may contribute to a positive rival effect, we specify the following general function of M_r^e that can allow us to test which mechanism is more likely to explain the data pattern:

$$M_r^e(H_{jt}) = \eta_1 S_{jt} + \eta_2 e_{jt} + \eta_3 S_{jt} e_{jt} + \sum_{l=1}^{L} \nu_l g_{jl}, \qquad (6)$$

which includes the rival's stock of outlets, the rival's experience, the interaction effect between the two, and the lags of net growth (with L representing the number of lags to be used).

Conditional on the existence of a rival's positive spillover effect, we expect $\eta_1 > 0$ for either circumstance of demand expansion or market learning. However, the signs of the rest of the parameters in Equation (6) vary depending on which explanation is the driving force of the positive effect. For demand expansion, both longer experience and a larger scale of presence from the rival contribute to a higher stock of cultivation, and there is a positive interaction effect between experience and the scale of presence, so we expect $\eta_2 \geq 0$ and $\eta_3 \geq 0$. Furthermore, the sequence of store openings with more outlets opened in the earlier stage and fewer new outlets in the later stage contributes to a higher level of cultivation compared with the reverse sequence, i.e., fewer outlets opened in the early stage and more additions in the later stage. In other words, conditional on the same total number of outlets, a descending sequence of new store openings or negative net growth converts to a higher degree of consumer learning, all else being equal. Hence we expect $\nu_l \leq 0$.

For market learning, the predictions would be different. If a firm uses the growth pattern of the rival chain to predict demand growth, then positive net growth (g_{jl}) would imply faster demand growth. Therefore, one would expect $\nu_l \geq 0$. In addition, conditional on the (positive) effect of a rival's stock (S_{jt}), the interaction of a rival's stock and a rival's experience

[10] Put more formally, one can formulate the degree of customer cultivation by the rival chain as $\omega_{jt} = r\sum_{s=1}^{t-1}(t-s)n_{js}$, where n_{js} is the number of outlets that chain j opened in period s.

[11] For empirical research of observational learning on the consumer side, see, for example, Cai et al. (2009) and Zhang (2010).

[12] Yang (2013) structurally models the Bayesian learning mechanism in the context of fast-food chains' entry into the Canadian market. Although we are not explicit about the learning process, the idea is consistent with a Bayesian learning model that firms revise the expected market size upwardly given the entry of the rival.

is likely to be negative; that is, $\eta_3 \leq 0$. In other words, quicker expansion or achieving the same scale of expansion in a relatively shorter period of time suggests higher demand growth. We also conjecture that $\eta_2 \leq 0$ conditional on S_{jt}.

In summary, although both explanations of a rival's positive impact yield the prediction that $\eta_1 > 0$, the predicted signs of other parameters in Equation (6) diverge, which provides us with a set of testable hypotheses as to which effect is more likely to be the driving force when the rival's impact on a chain's entry decision is positive, summarized as follows:

- If *demand expansion* is the driving force of the rival's positive impact, conditional on the positive effect of a rival's stock of outlets, the rival's experience is likely to be positive (i.e., $\eta_2 > 0$), and the interaction of a rival's experience and a rival's stock is likely to be positive (i.e., $\eta_3 > 0$). In addition, the chain is more likely to expand in a city where the rival has decelerated (negative) net growth, all else being equal (i.e., $\nu_l < 0$).
- If *market learning* is the driving force of the rival's positive impact, conditional on the positive effect of a rival's stock of outlets, the rival's experience is likely to be negative (i.e., $\eta_2 < 0$), and the interaction of a rival's experience and a rival's stock is likely to be negative (i.e., $\eta_3 < 0$). In addition, the chain is more likely to expand in a city where the rival has accelerated (positive) net growth, all else being equal (i.e., $\nu_l > 0$).

4. Empirical Analysis

To examine whether a rival's presence has a net positive effect on a chain's entry decision, and to detect the potential mechanism based on the tests above, we estimate a model of chain's expansion decision using data from the entire expansion history of KFC and McD's in the China market.

4.1. Model

We assume that the number of new outlets firm i opened in market m at period t, n_{imt}, follows a Poisson distribution. A time period here is a year, because we believe it is reasonable to assume that entry decisions are made on a yearly basis (Toivanen and Waterson 2005). The empirical specification of the Poisson model closely follows the theoretical derivation of the optimal number of outlets to open:[13]

$$
\begin{aligned}
E(n_{imt}) &= \exp\Big[\beta_0 + \beta_x X_{mt} + \beta_{o1} S_{imt} + \beta_{o2} e_{imt} \\
&\quad + \beta_{o3} e_{imt}\, time + \beta_{r0} n_{jmt-1} + \beta_{r1} S_{jmt} \\
&\quad + \beta_{r2} e_{jmt} + \beta_{r3} S_{jmt} e_{jmt} + \beta_{r4} S_{jmt} e_{imt} \\
&\quad + \beta_{r5} e_{jmt}\, time + \sum_{l=1}^{3} \beta_{r6}^{l} g_{jml} + \sigma v_{im}\Big], \qquad (7)
\end{aligned}
$$

where X_{mt} includes city population density and local GDP at time t, as well as a full set of year dummies and province dummies to control for any unobserved time trend and unobserved regional characteristics.[14] The rest of the control variables can be summarized as follows: (1) The *own-chain effect* includes the chain's own stock of outlets in the city at the beginning of period t (S_{imt}). The empirical distribution of the cumulative number of a chain's outlets is highly skewed. Therefore, we use the log transformation of the stock of outlets in the estimation. The sign of β_{o1} captures the net effect of a chain's own stock on further expansion. We also control for its own experience in the market (e_{imt}), which is defined as the number of years since chain i made initial entry into city m. Longer experience may benefit additional entry through operating efficiency, for example. Yet such an effect may diminish over time. Therefore, we also include an interaction term of the chain's own local experience and time.[15] (2) To measure the *rival's impact*, we control for the rival's outlets opened in the last period (n_{jmt-1}); the rival's stock of outlets, which is again taken log transformation (S_{jmt}),[16] the rival's experience in the city (e_{jmt}); and the interaction between the rival's stock and rival's experience. In addition, we control for the rival's net growth rate in the last three years with $g_{jml} = n_{jm,\,t-l} - n_{jm,\,t-1-l}$. As discussed earlier, although the rival's outlets can steal business from the focal chain, they may have a positive effect on the expected demand. If the net effect from the rival is positive ($\beta_{r1} > 0$), the sign of the parameters β_{r2}, β_{r3}, and β_{r6}^{l} can help us to differentiate whether demand expansion or market learning is more likely to explain the positive rival effect. This set of parameters is the main interest of our analysis. We further control for the interaction of the rival's stock and the focal chain's own experience as well as the rival's experience and time, allowing the rival's impact to change with the chain's own local and national market experience.

[13] Please refer to the appendix for the detailed description of the link between theoretical derivation and empirical specification. We estimate the model separately for KFC and McD's. Therefore, all the parameters are chain specific.

[14] The set of year dummies could capture such events as the implementation of franchising policy and the entry of other notable fast-food chains. The province dummies could pick up regional taste differences or policy differences, for example.

[15] The variable *time* is set as 1 for the starting year of the sample. This variable can be interpreted as a proxy for the chains' years of experience in the China market.

[16] More precisely, it is the log of one plus the cumulative number of rival outlets up to period $t-1$. Therefore, if the rival has not entered yet, this variable takes the value of 0. Jia (2008) employs the same log transformation when estimating the competition effect of the number of competing stores in Walmart's expansion decision.

The last term, σv_{im}, captures the time-invariant unobserved factor that affects firm i's expansion decision in market m, where σ is the standard deviation of the unobserved factor and v_{im} follows standard normal distribution. This term introduces unobserved heterogeneity into the expected mean of the number of new outlets across cities.

Note that a chain's expansion in a given city is conditional on entering the city first. We model the initial entry decision into a city as the following binary choice:

$$y^*_{imt} = Z'_{imt}\alpha + \epsilon_{imt}, \tag{8}$$

$$y_{imt} = 1 \quad \text{if } y^*_{imt} > 0, \tag{9}$$

where y^*_{imt} is the latent variable that represents firm i's propensity to enter market m at period t, and the observed counterpart y_{imt} is 1 when y^*_{imt} exceeds the threshold (0). Once firm i had the initial presence in city m at period τ, we observe the firm's expansion decision n_{imt} for each period $t > \tau$.[17]

In the above equation, Z_{imt} is a vector of observed variables that affect a chain's initial entry decision. We include the distance between city m and the chain's headquarters, the distance between the focal city and the chain's closest existing market (pre-entered city), as well as the number of firm i's pre-entered cities in the same province at t to capture chain i's initial entry cost in city m at period t.[18] The distance measures are log-transformed because of the skewed distribution of these variables. We also control for the stock of the rival's outlets at the beginning of period t (S_{jmt}), city characteristics, as well as two important policy implementations during the data period that may affect the chains' market entry decision. In late 1992, the State Council issued a notice ("open policy") to further open up 5 cities along the Yangtze River and the capital cities of 11 inland provinces, supporting these cities in attracting foreign investment. Furthermore, in 2000, the central government announced a new policy ("promote-west" policy) on the large-scale development of the western region with the objective to speed up the economic development of the undeveloped inland, which includes 11 provinces and 1 municipality in our data set.[19] We introduce two dummy variables to capture the two policy implementations, where a policy dummy is set to 1 for the affected cities during the policy years and 0 otherwise.

The error term ϵ_{imt} captures the unobserved factors that affect a firm's initial entry decision. We further decompose the term into two components: $\epsilon_{imt} = u_{im} + \varepsilon_{imt}$, where u_{im} is the unobserved time-invariant market and chain-specific effect that follows normal distribution ($u_{im} \sim \mathrm{N}(0, \sigma_u^2)$), and ε_{imt} is the random error that follows i.i.d. standard normal distribution. Therefore, the initial entry decision is characterized by a random effects probit model. The existence of u_{im} suggests that ϵ_{imt} error terms are correlated within the city m.

The initial entry and expansion decisions by firm i in city m are potentially correlated. Therefore, we assume that u_{im} and v_{im} jointly follow bivariate normal distribution:

$$\begin{bmatrix} u_{im} \\ v_{im} \end{bmatrix} \sim \mathrm{N}_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u \\ \rho\sigma_u & 1 \end{bmatrix}\right), \tag{10}$$

where ρ is the parameter that captures the correlation of the two decision stages. Furthermore, the conditional distribution of v_{im} can be denoted as $f(v_{im} \mid u_{im}) = \mathrm{N}[(\rho/\sigma_u)u_{im}, (1-\rho^2)]$. Hence, v_{im} can be expressed as a function of u_{im} with $v_{im} = (\rho/\sigma_u)u_{im} + \sqrt{1-\rho^2}\eta_{im}$, where $\eta_{im} \sim \mathrm{N}(0, 1)$.

Let W_{imt} denote all the observed explanatory variables in Equation (7); then, conditional on the initial entry occurred at period τ, the expected number of new outlets opened in market m at t can be written as

$$\begin{aligned} &E(n_{imt} \mid W, Z) \\ &= E\left[\exp\left(W'_{imt}\beta + \rho\frac{\sigma}{\sigma_u}u_{im} + \sigma\sqrt{1-\rho^2}\eta_{im}\right)\middle| u_{im} > \right. \\ &\left. \quad -(Z'_{im\tau}\alpha + \varepsilon_{im\tau}) \text{ and } u_{im} < \min_{1\le l<\tau}\{-(Z'_{iml}\alpha + \varepsilon_{iml})\}\right]. \end{aligned} \tag{11}$$

The expansion decision is conditional on the initial entry the chain made at period τ (and not the periods before τ). Estimating the Poisson model without accounting for the initial entry condition would lead to sample selection bias. The expectations are taken over the distributions of u_{im}, η_{im}, and ε_{imt}.

Since the full information maximum likelihood approach is computationally burdensome with multiple integrals,[20] following Terza (1998), we estimate the model using a two-step estimation procedure. The idea can be traced back to Heckman (1979). To account for the unobserved heterogeneity that leads to the selection (initial entry condition being met)

[17] The initial entry equation is similar to a selection equation in Heckman's (1979) sample selection model.

[18] Another measure of entry cost is the distance to the nearest distribution center. Information on distribution centers is not available from the two chains' official websites. We therefore searched the establishment of KFC or McD's distribution centers from news archives. The results show that distribution centers were typically established after a chain has entered the region. For example, KFC's Beijing distribution center was reportedly constructed in 2004 (Alestron 2004).

[19] A brief introduction of the policies can be found on *Wikipedia* (http://en.wikipedia.org/wiki/Special_Economic_Zones_of_the_People's_Republic_of_China and http://en.wikipedia.org/wiki/China_Western_Development, accessed September 21, 2013).

[20] Greene (1997) and Winkelmann (2001) discuss the estimation of the cross-sectional count data model with sample selection using the full information maximum likelihood procedure.

in Equation (11), we can first derive estimates of the unobserved factor from the residuals of the random effects probit model in the initial entry stage and include them as additional explanatory variables in the estimation of the expansion decision stage. Although the simple Heckman-type correction method that applies to linear models does not directly apply to nonlinear models, Orme and Peters (2001) show that one can use polynomials of Heckman-type correction variables as an approximation in the second-stage estimation for count model. We extend the correction method proposed by Orme and Peters on cross-sectional count models to panel count data, noting that the unobserved time-invariant factor u_{im} is not only bounded below as $u_{im} > -(Z'_{im\tau}\alpha + \varepsilon_{im\tau})$ but also bounded above as $u_{im} < \min_{1\le l<\tau}\{-(Z'_{iml}\alpha + \varepsilon_{iml})\}$. The latter condition comes from the fact that initial entry threshold was not met prior to period τ. Letting $\pi^L_{im\tau} = -(Z'_{im\tau}\alpha + \varepsilon_{im\tau})/\sigma_u$ and $\pi^U_{im\tau} = \min_{1\le l<\tau}\{-(Z'_{iml}\alpha + \varepsilon_{iml})/\sigma_u\}$, the first-order correction term can be expressed as[21]

$$\kappa_{im} = \iint \frac{\phi(\pi^L_{im\tau}) - \phi(\pi^U_{im\tau})}{\Phi(\pi^U_{im\tau}) - \Phi(\pi^L_{im\tau})} \cdot f(\pi^L_{im\tau}) f(\pi^U_{im\tau})\, d\pi^L_{im\tau}\, d\pi^U_{im\tau}, \tag{12}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ represent the probability density function and the cumulative distribution function of the standard normal distribution, respectively. Note that the term within the integral reduces to the familiar inverse Mill's ratio if the entry decision is observed for only one period ($\tau = 1$). We use κ_{im} as the correction term to control the selection bias.

The two-step estimation proceeds as follows. In the first step, we estimate the random effects probit model. With the estimated parameter vector $\hat{\alpha}$ and $\hat{\sigma}_u$, we can derive the empirical correction term κ_{im}. In the second step, we estimate the Poisson model with the approximated conditional mean as follows:

$$E(n_{imt} \mid W, X) = \exp(W'_{imt}\gamma^* + \sigma\rho\kappa_{im}). \tag{13}$$

Note that since σ is nonnegative, the sign of the parameter in front of κ_{im} indicates the sign of ρ. To check the validity of this two-step estimation method, we conduct a simulation analysis and find that the key parameters can be correctly recovered.[22]

4.2. Endogeneity

An important issue to discuss before proceeding to estimate Equation (13) is endogeneity. Notice that by controlling for κ_{im} in this stage, we correct for the *sample selection bias*, i.e., the timing that a chain expands in a city (accumulates own experience) is not random. However, there is still concern about the endogeneity of other variables. Specifically, a rival's stock of outlets S_{jmt} could be an endogenous variable if there exist some unobserved market-specific factors such as the taste for a certain type of food that may affect both the focal and rival chains. In addition, the stock of a chain's own outlets S_{imt} is also potentially endogenous. For example, an unobserved chain-specific effect can affect a chain's past expansion as well as its current decision.

We address the issue by using the control function approach, which involves two steps (Wooldridge 2002). The first step involves ordinary least square regression of the endogenous variables on the exogenous and instrument variables. In the second step, we run the equation of interest with both the original endogenous variables and the first-stage residuals as additional variables. In essence, the second step uses the first-step residuals as a proxy for the unobserved variables to address the omitted variable issue and restore the consistency of the parameter estimates. Notice that the second step is different from the usual two-stage least square estimation that replaces the endogenous variables with their first-stage predicted values. Terza et al. (2008) show that the two-stage least square method that addresses endogeneity in linear models may produce inconsistent estimates in nonlinear models such as count data models, whereas the control function approach as described above is consistent.

We run the following instrumental variables (IV) regression of the endogenous variables S_{jmt} ($j \in \{\text{KFC}, \text{McD's}\}$) in the first step:

$$S_{jmt} = Z_{1jmt}\pi_1 + V_{jmt}\pi_2 + \nu_{jmt}, \tag{14}$$

where Z_{1jmt} includes the exogenous variables in Z, and V_{jmt} are the instrumental variables for S_{jmt}. The instruments that we use are the distance between city m and the closest city that chain j has entered up to period t and the total number of outlets that chain j owns up to t within a radius of 100 kilometers outside the focal city m. These variables are picked as instruments because the distance to the chain's network and the entry decision in nearby regions could be correlated with a chain's expansion decision in the focal city. For example, chain j may be more likely to open outlets in city m when the city is surrounded by more existing markets and stores, as the marginal operation and management cost is lower. Yet the variables are unlikely to be correlated with the unobserved characteristics of city m.[23]

The reduced-form errors ν_{jmt} are independent of the explanatory variables in Equation (14) by

[21] The derivation details are provided in the appendix.

[22] Please see the appendix for simulation details.

[23] One potential concern of using the number of outlets in the neighboring region as instrument is that the unobserved factors affecting the entry decision in the focal city could be correlated with

construction. By the nature of panel data, ν_{jmt} is decomposed into a time-invariant and a time-varying error, $\nu_{jmt} = \bar{\nu}_{jm} + \varsigma_{jmt}$. Both components can be recovered from within-panel least square regression. By including the estimated residuals $\bar{\nu}_{jm}$ and ς_{jmt} as additional regressors in the equation of interest, one corrects for the endogeneity of S_{jmt}, which might be caused by either time-invariant or time-varying unobserved factors.

Note that the main equation of interest (Equation (13)) involves two sets of endogenous variables: one is the rival's past expansion behavior in the focal city and the other is the focal chain's past expansion in the city. Therefore, we use two control functions to address the endogeneity concern: one focuses on the rival's stock of outlets S_{jmt} and the other focuses on one's own stock of outlets S_{imt}. We make an assumption that one control function suffices for each set of the endogenous variables.[24] The idea is that both the rival's stock of outlets S_{jmt} and other variables of a rival's expansion are potentially correlated with the same underlying factors. By controlling for the residuals from the control function on S_{jmt}—namely, $\bar{\nu}_{jm}$ and ς_{jmt}—one captures both the time-invariant and time-varying unobserved factors that may lead to the endogeneity of the set of variables if left uncontrolled.[25] The residuals from the two IV regressions are controlled in the Poisson model of a chain's expansion decision to address the endogeneity concern.

4.3. Empirical Results

The main results for KFC and McD's are summarized in Tables 4(a) and 4(b), respectively. In each table, the upper panel reports the results of a chains' expansion decision and the lower panel estimates the initial entry condition. First of all, we find evidence that a rival's presence has a positive effect on a chain's expansion decision in general. Second, we find that the positive effect seems to be driven by different mechanisms for KFC versus McD's based on the tests proposed in the theoretical framework.

Table 4(a) Model Estimates for KFC

Parameter	No endogeneity control	Endogeneity control
Intercept	−3.346***	−3.415
	(0.293)	(2.164)
Density	0.378	0.406
	(0.424)	(0.628)
GDP	−0.000	−0.001***
	(0.000)	(0.000)
KFC_stock	0.473***	0.882
	(0.131)	(1.113)
KFC_experience	−0.095	−0.176
	(0.069)	(0.115)
KFC_experience × *Time*	0.019***	0.023***
	(0.005)	(0.007)
McD's number of outlets opened last	0.059***	0.062*
	(0.022)	(0.032)
McD's stock	0.342**	0.479**
	(0.144)	(0.192)
McD's_experience	0.210**	0.298**
	(0.107)	(0.145)
McD's_experience × *Time*	−0.012***	−0.017***
	(0.005)	(0.007)
McD's_stock × *KFC_experience*	−0.064***	−0.089***
	(0.010)	(0.022)
McD's_stock × *McD's_experience*	0.028***	0.057***
	(0.011)	(0.017)
Growthlag1_McD's	−0.065***	−0.067***
	(0.02)	(0.024)
Growthlag2_McD's	−0.049***	−0.0531***
	(0.020)	(0.015)
Growthlag3_McD's	−0.013	−0.019
	(0.011)	(0.012)
Correction term	−0.244***	−0.355***
	(0.065)	(0.074)
Own-stock time-invariant residual		1.127
		(1.166)
Rival-stock time-invariant residual		−0.911***
		(0.246)
Time-variant residual		−0.889
		(1.112)
Year dummies included?	Yes	Yes
Province dummies included?	Yes	Yes
Number of observations	1,258	1,258
Log likelihood	−1,382.713	−1,354.405

4.3.1. Demand Learning vs. Market Expansion. Recall from the theoretical framework that the net effect from the rival on the focal chain's expansion decision depends on the magnitude of the positive rival effect on the expected size of the market demand relative to the competition effect. It is an empirical question as to which effect dominates in this setting. Thus, we first examine the net effect of rival's stock of outlets on a chain's expansion decision. In the upper panel of Tables 4(a) and 4(b), the first column reports the results before correcting for endogeneity. For both chains, the coefficient of a rival's

the unobservables affecting entry in the nearby regions. First note that the province dummies capture the regional similarity to some extent. In addition, we performed the Lagrange multiplier (LM) tests for both the spatially lagged dependent variable and spatial error correlation based on the method in Elhorst (2010). In these tests, the dependent variable is the binary decision of whether to enter or to add additional outlets, and the control variables are the ones discussed above. The location weighting matrix is generated based on the longitude and latitude of each city that assigns heavier weight to geographically closer cities. In both LM tests for KFC and McD's, we cannot reject the null hypothesis of no spatial lag or no spatial errors at the 10% level.

[24] Imbens and Wooldridge (2007) discuss a similar case that uses a single control function to deal with two endogenous variables.

[25] Empirically, using additional control functions on a rival's experience and controlling the residuals of both the rival's experience and rival's stock in the main equation result in a multicollinearity problem.

Table 4(a) (Cont'd.)

Initial entry condition	
Intercept	−4.225 (3.152)
Density	1.495*** (0.447)
GDP	0.005*** (0.001)
Capital city dummy	0.713* (0.424)
Distance to headquarters	−0.284 (0.458)
Number of pre-entered neighboring cities	0.222*** (0.041)
Distance to closest entered city	−0.063 (0.114)
Open policy dummy	0.200 (0.411)
Promote-west policy dummy	0.743** (0.356)
Rival stock	−0.151 (0.248)
Year dummies included?	Yes
Province dummies included?	Yes
Number of observations	3,177
Log likelihood	−496.778

Notes. Clustered standard errors are in parentheses. In the column with endogeneity controlled, the standard errors are computed using the bootstrapping method.

*Significant at the 0.1 level; **significant at the 0.05 level; ***significant at the 0.01 level.

Table 4(b) Model Estimates for McD's

Parameter	No endogeneity control	Endogeneity control
Intercept	−5.702*** (0.904)	−5.549 (4.949)
Density	−0.544 (0.527)	−1.131 (1.863)
GDP	0.001*** (0.000)	−0.001*** (0.000)
McD's_stock	0.009 (0.158)	−0.204 (0.140)
McD's_experience	0.307** (0.151)	0.462** (0.213)
McD's_experience × Time	0.005 (0.010)	−0.016 (0.012)
KFC's number of outlets opened last	−0.074*** (0.028)	−0.047** (0.024)
KFC stock	0.866*** (0.287)	2.752*** (0.975)
KFC_experience	−0.285* (0.173)	−0.222 (0.209)
KFC_experience × Time	0.023** (0.011)	0.018 (0.012)
KFC_stock × McD's_experience	−0.054** (0.027)	−0.043 (0.029)
KFC_stock × KFC_experience	−0.045** (0.022)	−0.049* (0.028)
Growthlag1_KFC	0.062*** (0.021)	0.048*** (0.018)
Growthlag2_KFC	0.080*** (0.025)	0.067*** (0.020)
Growthlag3_KFC	0.060*** (0.019)	0.055*** (0.017)
Correction term	−0.171** (0.070)	−0.218* (0.125)
Own-stock time-invariant residual		2.209*** (0.483)
Rival-stock time-invariant residual		−1.821 (1.283)
Time-variant residual		−2.048** (0.994)
Year dummies included?	Yes	Yes
Province dummies included?	Yes	Yes
Number of observations	787	787
Log likelihood	−796.113	−772.757

stock of outlets appears to be positive and significant. The second column addresses the endogeneity issue using the control functions approach as discussed above. The standard error of the parameters is obtained using bootstrapping; residuals from the first-stage regressions are controlled as observed variables.[26] In the IV regression of the rival's stock and a chain's own stock, the F-statistics are larger than the usual criterion of 10 (238.59 and 329.73, respectively), and the coefficients of the instrumental variables are highly significant, which lends support to the validity of the instruments.

Inspecting the coefficients of the included IV regression residuals, we find that for KFC, the rival's stock time-invariant residual is negative and significant, suggesting that the unobserved factor that encourages McD's past expansion could be negatively correlated with KFC's current expansion. Controlling for this factor leads to an increase of the coefficient in front of the rival's stock of outlets. In other words, one would significantly underestimate the positive effect of McD's stock of outlets on KFC's expansion decision if not controlling for the endogeneity of this variable. For McD's, we find that the own-stock time-invariant residual is positive and significant, whereas the time-varying residual is negative and significant. It implies that the unobserved factor that affects McD's past expansion is positively correlated with the chain's current expansion decision, which could be, for example, the chain's local reputation. The time-varying residual has a negative impact, suggesting that expansion in the past negatively affects the decision at the current period, which may be due to the chain's resource constraints in the expansion process.[27]

[26] Specifically, we use 100 replications of cluster bootstrapping in which each cluster unit is a city with a complete entry history. Using cluster bootstrapping helps preserve the unobserved correlation in the observations within the same city.

[27] Note that for each IV regression there are two residuals: one is time invariant and the other time varying. Therefore, in theory

Table 4(b) (Cont'd.)

	Initial entry condition
Intercept	−0.929 (1.222)
Density	0.006 (0.588)
GDP	0.006*** (0.001)
Capital city dummy	−0.207 (0.431)
Distance to headquarters	−0.337** (0.143)
Number of pre-entered neighboring cities	0.322*** (0.046)
Distance to closest entered city	−0.200* (0.105)
Open policy dummy	−0.063 (0.415)
Promote-west policy dummy	0.739* (0.380)
Rival stock	0.536*** (0.174)
Year dummies included?	Yes
Province dummies included?	Yes
Number of observations	3,648
Log likelihood	−376.297

Notes. Clustered standard errors are in parentheses. In the column with endogeneity controlled, the standard errors are computed using the bootstrapping method.

*Significant at the 0.1 level; **significant at the 0.05 level; ***significant at the 0.01 level.

Controlling for endogeneity, we discover that the previous finding of a rival's positive net effect remains robust.[28] According to the theoretical predictions presented in §3, although both demand learning and market expansion may contribute to positive rival effect, the two explanations generate different predictions on how the focal chain would react to its rival's expansion pattern. Recall that if the focal chain benefits from its rival's impact on cultivating local consumers, it is more likely to expand in cities where the rival has been existing longer with more outlets opened in the earlier years. In contrast, if the focal chain uses the growth of the rival chain to obtain information about market growth, then it is likely to add more outlets in cities where the rival expands more quickly with a positive net growth rate.

To test which explanation is more likely to drive the positive rival effect, we focus on the coefficients of the rival's experience, the interaction of the rival's experience and stock of outlets, and the rival's net growth rate in the last few periods. By comparing the results in Tables 4(a) and 4(b), we find that these coefficients are in the opposite direction for KFC versus McD's. For KFC, both McD's experience and the interaction of McD's stock with McD's experience are positive and significant, suggesting that KFC prefers markets with a longer presence of its rival. This is consistent with the demand expansion explanation. For McD's, by contrast, the interaction of KFC's stock and KFC's experience is significantly negative (conditional on the positive effect of the rivals' stock of outlets), indicating that McD's appreciates the markets where KFC opens more outlets in a shorter period of time. The coefficient of a rival's experience is also negative though insignificant after correcting for endogeneity. These results are in line with the market learning explanation.

Now look at the signs of the rival's lagged net growth rate. We find that the net growth of McD's has a negative effect on KFC's decision to expand. This suggests that KFC would open more outlets in a city where more McD's outlets were opened previously compared with a city where more outlets were recently added. A city with more rival outlets opened at an early stage has a higher stock of consumer cultivation than a city with fewer rival outlets in the beginning but more outlets newly added, conditional on the total number of a rival's stores. Hence the preference for the descending sequence of the rival's store expansion indicates that demand expansion is more likely to explain the positive effect of McD's presence on KFC.

In contrast, for McD's, we find that KFC's net growth in the last three periods has a positive and significant effect on its expansion decision. That is, everything else being equal, McD's is likely to add more outlets in a city where KFC continued to grow with a positive rate. This is consistent with the market learning explanation—that is, McD's is likely to use KFC's expansion as demand signals and revise its assessment on demand potential upwardly.

In summary, conditional on the positive effect of the rival's stock of outlets, we find that positive rival effect is likely to be driven by different underlying forces for KFC and McD's. The results are consistent with the explanation that the presence of KFC signals

we have four residuals to add to the Poisson model of interest. However, we find that adding both time-varying residuals to the model causes a multicollinearity problem. Therefore, empirically we add the two time-invariant residuals to control for the time-invariant unobserved factor of each chain and one time-varying residual to capture the changing unobserved factor.

[28] In an unreported analysis, we estimated a basic specification of chains' expansion decision, which only includes city characteristics, own-chain effect, and the rival's stock of outlets; it excludes other variables currently in the model. In a variety of model specifications, such as adding other city characteristics (e.g., school age population and local wage for the years when data are available, 1996–2007) and allowing for serial correlation in the error terms, we consistently find that the rival's stock of outlets has a positive effect on the focal chain's expansion decision. Results are available from the authors upon request.

market potential and growth to McD's, whereas the presence of McD's helps to cultivate consumer tastes that benefit KFC. The effect from a rival's stock of outlets and rival's experience tends to diminish as a chain's own local and global experience accumulates, which is reflected by the interaction terms of a rival's presence with the focal chain's own experience.

Finally, we briefly examine the role of a chain's own past expansion on one's current expansion decision. The coefficient of KFC's own stock of outlets tends to be positive, suggesting that the positive effect, such as customer learning and brand building from its own existing outlets, may outweigh the within-chain cannibalization effect. McD's tends to open more outlets in cities with longer local experience, which may be associated with lower operating cost.

4.3.2. Initial Entry Condition. Estimating the initial entry condition helps us to infer the value of the unobserved factor u_{im}, which is potentially correlated with the unobserved factor v_{im} in the chain's expansion decision. The sign of the parameter in front of the correction term κ_{im} in the expansion model indicates the direction of the correlation. The coefficient of κ_{im} in both chains' expansion models is negative and significant, implying $\rho < 0$. That is, the unobserved factors in a chain's expansion decision are negatively correlated with the unobserved factors in the initial entry decision. This result may seem counterintuitive at first. Deeper inspection, however, reveals that those cities with less attractive observed city characteristics but that witnessed the entry of KFC or McD's have bigger unobserved factors recovered from the first-stage estimation.[29] In other words, if KFC or McD's entered an economically less developed city, it suggests that there exists some unobserved positive factor that triggers the chain's entry decision. On the other hand, for economically more developed cities with attractive observables, the initial entry threshold is easily satisfied, which identifies a small unobserved factor in that stage. Conditional on initial entry, the economically more developed cities experience faster chain expansion than average, whereas we may observe fewer additions of outlets in the less developed cities that KFC or McD's entered possibly for strategic reasons. Therefore, the unobserved factor in the expansion stage is small for economically less developed cities, and the opposite is true for the more developed cities. Taken together, one finds a negative correlation of unobserved factors across the two stages.

We now look at the estimates of the random effects probit model for the initial entry condition reported at the lower panel of each table. First, GDP has a significant and positive effect on chains' initial entry decisions. KFC is also more likely to enter densely populated cities, whereas density is not a significant predictor for McD's initial entry. For both chains, the number of pre-entered neighboring cities in the same province significantly increases a chain's probability of entering the focal city, suggesting the existence of the economies of density (Holmes 2011). The distance to a chain's headquarters and the closest pre-entered market have a negative effect on entry, but the effect is significant only for McD's. It suggests that McD's is more likely to enter a city that is closer to its existing network or headquarters. In addition, the implementation of the promote-west policy increases both chains' entry probability in the affected regions, since the coefficient in front of this policy dummy is significantly positive for both KFC and McD's.

For KFC, the rival's stock of outlets seems to have an insignificant effect on its initial entry decision. However, we need to be careful in interpreting the result, as the insignificance of McD's stock on KFC's initial entry is largely dictated by the fact that there are limited observations where McD's makes the initial entry ahead of KFC. On the other hand, McD's seems more likely to enter a city with a larger stock of KFC's.[30]

4.4. Discussion

The empirical analysis of KFC and McD's expansion in the Chinese market suggests that the rival's net impact is positive during the period of time under study. Based on the difference in the chains' reaction to their rival's expansion pattern, we may conclude that market learning or information externality is more likely to be the driving force of KFC's positive effect on McD's, and demand expansion can explain the positive effect of McD's on KFC.

This result may be rationalized by the unique characteristics of each chain. It is likely that McD's has a stronger effect on cultivating consumer taste because it is more symbolic of Western food than is KFC. KFC, however, may have a better understanding of local Chinese consumers. For example, the founding leadership team of KFC's Great China division is formed mainly by Chinese born in Taiwan—people who have a better knowledge of the Chinese culture and market than the leader team in McD's formed by nonlocals (Liu 2008).

In addition to the current analysis, we also utilize supplemental data to test the proposed explanation.

[29] The simple correlation between the estimated unobserved factor and local GDP is negative for both chains.

[30] Note that we do not attempt to make a strong inference from the individual parameter estimates in the initial entry stage, as the main role of estimating the initial condition is to control for selection bias in the expansion model.

One check concerns market learning. A better indicator of market profitability than the total number of a rival chain's outlets in a market would be direct sales revenue. If a chain is learning about local demand and potential, then it should be more likely to expand to markets with higher observed sales revenue from the rival. Although we do not have data on the sales of KFC or McD's at the city level, we were able to collect aggregate information on fast-food restaurants' sales revenue at the province level from 2002 to 2007.[31] We then ran the analysis of a chain's expansion decision with the added variable of *restaurant sales revenue* and found it to be highly significant for McD's but not significant for KFC.[32] Although the variable is a crude indication of category demand that includes both Chinese and Western fast food, the significance of sales revenue on McD's expansion is consistent with its learning about the propensity of local fast-food consumption.

However, we want to point out that there might be other explanations that can rationalize the observed data pattern, and our analysis does not rule out alternative explanations. We also want to acknowledge the weakness of and a few potential issues in our data analysis. First, the observed city characteristics are limited. Many of the statistics are not available at the city level for early years. Although we have controlled for the most important economic city characteristics, missing some of the additional local characteristics may lead to overestimation of positive rival effect.

Second, although we have carefully dealt with selection and endogeneity issues, these concerns may not be fully alleviated. There are multiple potentially endogenous variables in the second-stage estimation of a chain's expansion decision. To address the issue, we use two control functions to correct the endogeneity arising from the variables related to the rival's past expansion (e.g., the rival's stock of outlets and rival's experience) and the variables related to one's own past actions, respectively. A crucial assumption is that the endogeneity of the set of rival (own)-related variables is caused by the same source, and therefore one control function suffices. If the assumption is not true, then there could be unresolved endogeneity, and the parameter estimates could be biased.

5. Variation of the Rival's Effect

One question is whether the rival's effect may evolve in different growth stages of a chain. For both KFC and McD's, 1999 marks their takeoff year (see Figures 1 and 2); more than 80% of new outlets were added to each company's territory map after 1999. To examine whether the positive effect from the rival differs in the conservative expansion era (pre-1999) and the aggressive expansion era (post-1999), we created a dummy named *pre-1999*, which takes the value of 1 for the years before 1999 and 0 otherwise. We include an interaction term of this dummy with the rival's stock of outlets in the expansion decision model. The results are shown in the left panel of Table 5. We find that for KFC, the rival's effect seems stronger in the post-1999 era when McD's has stronger presence, as suggested by the negative coefficient of the newly added interaction term. For McD's, the rival's impact does not differ significantly in the pre- and post-1999 stages. Note that the interaction term captures the overall change across time stages. For individual markets, the positive effect of the rival's presence decreases with one's own experience in the same market, as suggested by the negative interaction of the rival's stock of outlets and one's own local experience.

In addition, we examine whether the rival's impact is different in big versus small cities. Again, we add a simple interaction of the rival's stock of outlets and the dummy variable for big cities to capture the difference. The results are in the right panel of Table 5. We find that KFC's scale of presence has a positive effect on McD's expansion into both big and small cities, though the effect is stronger in big cities. However, McD's positive impact on KFC is mainly from the big cities.

Finally, we consider the case that the firms' actual decision window is longer than one year and check the rival's effect in this case. The idea is that even though one may observe the additions of new outlets from KFC or McD's in each year, firms may have their expansion plan made for longer periods and execute the plan during the time window. We therefore estimate the model (Equation (11)) assuming that the decision window for both chains are two years instead of one.[33] Now the dependent variable is the number of new outlets to open in the past two years. Accordingly, the lagged variables mean the change in the last decision period (instead of last year). The results are provided in Table 6. Although the aggregation of the information from yearly to biyearly reduces the number of observations for analysis and renders some of the estimates less accurate, the overall results are fairly close to the ones in Tables 4(a) and 4(b). We find the rival's net effect to be positive and significant.

[31] Data are collected from the 2004–2006 editions of the *Statistical Yearbook of China Restaurants* and the 2007 and 2008 editions of the *Statistical Yearbook of China Chain Stores and Retail Trades and Catering Service* (all editions compiled by the National Bureau of Statistics of China).

[32] The results are presented in Table A.1 in the appendix.

[33] The only difference is that we control for two lags of the rivals' net growth rate instead of three because one lag of a decision period is now two years.

Table 5 Variation of Rival Effects

Across time			Across big and small cities		
Variable	McD's	KFC	Variable	McD's	KFC
Intercept	−5.679	−3.383	*Intercept*	−4.480	−3.235
	(5.353)	(2.441)		(5.937)	(2.324)
Density	−1.065	0.413	*Density*	−1.231	0.397
	(2.190)	(0.563)		(1.899)	(0.511)
GDP	−0.001***	−0.001**	*GDP*	−0.001***	−0.001**
	(0.000)	(0.000)		(0.000)	(0.000)
Rival's number of outlets opened last	−0.044**	0.069**	*Rival's number of outlets opened last*	−0.041**	0.061**
	(0.020)	(0.030)		(0.020)	(0.030)
McD's stock	−0.183	0.642**	*McD's stock*	−0.271*	0.198
	(0.153)	(0.207)		(0.148)	(0.245)
McD's stock × Pre-1999 dummy		−0.300***	*McD's stock × Big city dummy*		0.274*
		(0.074)			(0.157)
McD's_experience	0.392	0.300**	*McD's_experience*	0.508**	0.302**
	(0.239)	(0.145)		(0.217)	(0.119)
McD's_experience × Time	−0.012	0.024**	*McD's_experience × Time*	−0.021*	−0.017***
	(0.014)	(0.007)		(0.012)	(0.006)
McD's_stock × KFC_experience		−0.102***	*McD's_stock × KFC_experience*		−0.087***
		(0.020)			(0.020)
McD's_stock × McD's_experience		0.052***	*McD's_stock × McD's_experience*		0.055***
		(0.018)			(0.016)
KFC_stock	2.729***	0.770	*KFC_stock*	1.768*	0.779
	(0.896)	(1.325)		(1.032)	(1.284)
KFC_stock × Pre-1999 dummy	−0.287		*KFC_stock × Big city dummy*	0.714***	
	(0.215)			(0.257)	
KFC_experience	−0.179	−0.174	*KFC_experience*	−0.124	−0.158
	(0.206)	(0.121)		(0.167)	(0.114)
KFC_experience × Time	0.017	0.024***	*KFC_experience × Time*	0.013	0.022***
	(0.013)	(0.007)		(0.010)	(0.007)
KFC_stock × McD's_experience	−0.048		*KFC_stock × McD's_experience*	−0.032	
	(0.030)			(0.028)	
KFC_stock × KFC_experience	−0.063**		*KFC_stock × KFC_experience*	−0.051**	
	(0.029)			(0.022)	
Growthlag1_McD's		−0.068**	*Growthlag1_McD's*		−0.066***
		(0.022)			(0.024)
Growthlag2_McD's		−0.049***	*Growthlag2_McD's*		−0.053***
		(0.014)			(0.015)
Growthlag3_McD's		−0.011	*Growthlag3_McD's*		−0.019
		(0.011)			(0.012)
Growthlag1_KFC	0.044***		*Growthlag1_KFC*	0.044***	
	(0.016)			(0.014)	
Growthlag2_KFC	0.064***		*Growthlag2_KFC*	0.063***	
	(0.018)			(0.018)	
Growthlag3_KFC	0.053***		*Growthlag3_KFC*	0.053***	
	(0.014)			(0.013)	
Correction term	−0.218*	−0.360***	*Correction term*	−0.218**	−0.351***
	(0.120)	(0.074)		(0.098)	(0.073)
Own-stock time-invariant residual	2.128***	1.219	*Own-stock time-invariant residual*	2.484***	1.118
	(0.476)	(1.366)		(0.390)	(1.299)
Rival-stock time-invariant residual	−1.674	−0.876***	*Rival-stock time-invariant residual*	−1.983*	−0.879***
	(1.195)	(0.229)		(1.143)	(0.210)
Time-variant residual	−1.892*	−0.837	*Time-variant residual*	−1.890**	−0.811
	(1.038)	(1.335)		(0.949)	(1.296)
Year dummies included?	Yes	Yes	Year dummies included?	Yes	Yes
Province dummies included?	Yes	Yes	Province dummies included?	Yes	Yes
Number of observations	787	1,258	Number of observations	787	1,258
Log likelihood	−771.554	−1,341.450	Log likelihood	−766.601	−1,348.232

Note. Bootstrapped standard errors are in parentheses.
*Significant at the 0.1 level; **significant at the 0.05 level; ***significant at the 0.01 level.

Table 6 Model Estimates with a Two-Year Decision Window

Variable	McD's	KFC
Intercept	−4.456	−2.192
	(4.112)	(1.987)
Density	−0.642	0.197
	(1.440)	(0.429)
GDP	−0.000	−0.001**
	(0.000)	(0.000)
Rival's number of outlets opened last	−0.082**	0.022
	(0.033)	(0.018)
McD's stock	−0.421***	0.718***
	(0.132)	(0.226)
McD's_experience	0.521**	0.142
	(0.263)	(0.138)
McD's_experience × Time	−0.017	−0.011
	(0.014)	(0.007)
McD's_stock × KFC_experience		−0.103***
		(0.019)
McD's_stock × McD's_experience		0.070***
		(0.015)
KFC_stock	2.480***	0.463
	(0.789)	(1.164)
KFC_experience	−0.529**	−0.136
	(0.239)	(0.121)
KFC_experience × Time	0.035**	0.024***
	(0.014)	(0.007)
KFC_stock × McD's_experience	−0.055	
	(0.041)	
KFC_stock × KFC_experience	−0.060*	
	(0.033)	
Growthlag1_McD's		−0.022**
		(0.010)
Growthlag2_McD's		0.004
		(0.006)
Growthlag1_KFC	0.071***	
	(0.023)	
Growthlag2_KFC	0.046**	
	(0.022)	
Correction term	−0.201	−0.367***
	(0.154)	(0.068)
Own-stock time-invariant residual	2.050***	1.297
	(0.419)	(1.182)
Rival-stock time-invariant residual	−1.456	−1.085***
	(0.973)	(0.245)
Time-variant residual	−1.350**	−0.517
	(0.667)	(1.229)
Year dummies included?	Yes	Yes
Province dummies included?	Yes	Yes
Number of observations	358	563
Log likelihood	−476.678	−855.035

Note. Bootstrapped standard errors are in parentheses.
*Significant at the 0.1 level; **significant at the 0.05 level; ***significant at the 0.01 level.

6. Concluding Remarks

Using an originally assembled data set on KFC and McD's expansion in China from their initial entry up to the year 2007, we investigate how the presence of a rival chain affects one's expansion decision, controlling for local market characteristics and own-chain characteristics. We find that the positive effects from the rival dominate the competition effect during the period under study, which leads to an overall positive rival impact on a chain's market expansion.

Our focus is to identify the driving forces of the positive effect from the rival. The empirical evidence suggests that market learning is likely to explain the positive influence of KFC on McD's expansion decision. That is, McD's uses KFC's entry and expansion patterns as signals of market size and growth. On the other hand, demand expansion is more plausible with McD's positive spillover on KFC. KFC benefits from McD's role of cultivating local consumers' taste to embrace Western fast food and is more likely to expand in cities with a long and large presence of McD's. However, we also want to acknowledge that there could be alternative explanations that rationalize the chain expansion pattern one observes. The bottom line is that we find the two chains respond differently to the expansion pattern of the rival chain, and such differences together with other empirical and institutional evidence are consistent with the explanations above.

This paper sheds light on how multinational firms make entry and expansion decisions in an emerging market. It empirically shows that under certain market environments, the rivals' presence may have a positive effect on a firm's expansion in the market. There are a number of avenues that future research can pursue. First, future research can be extended to structurally measure the multiple effects within and across chains.[34] Our analysis can only identify the net effect from the rival's or own stock of outlets. A structural model is required to separate the positive and competition or cannibalization effect. Second, we study the positive impact from the rival within the product category. It would also be interesting to study the spillover effect across categories and industries in the context of entry and expansion in emerging markets. Last, but not least, examining the rivals' impact on entry in other industries or emerging markets may offer additional insights on different explanations of a rival's effect.

Acknowledgments
The authors are listed alphabetically and contributed equally to the paper. They are grateful to the editor, the associate editor, and three anonymous reviewers for their constructive comments and helpful suggestions. The

[34] Holmes (2011) can be seen as an initial step for such an attempt, which allows one's previous entry to have a spillover effect on one's additional entry in nearby locations through economies of scale. A model with multiple spillover effects and competition would be much more challenging.

authors thank David Bell, Eric Bradlow, Tat Chan, Yuxin Chen, Ivan Png, Jagmohan Raju, Anita Rao, K. Sudhir, Raphael Thomadsen, Miguel Villas-Boas, John Zhang, and Juanjuan Zhang for helpful discussions and feedback on the earlier draft of the paper. The authors also thank seminar participants at the National University of Singapore, University of California at Berkeley, 2011 Marketing Science Conference, 2012 Frank M. Bass UT Dallas FORMS Conference, 2012 China India Insights Conference, and 2012 Marketing Science Emerging Markets Conference for their comments. The authors gratefully acknowledge financial support from the Global Initiatives Research Program of the Wharton School and the Singapore Ministry of Education Research [Project R-316-000-078-112].

Appendix

A.1. Mapping from the Theoretical Framework to the Empirical Model

In our empirical analysis, we specify the cost component as $\partial V/\partial n_{it} = c_0 + c_1 e_{it} + c_2 e_{it}\,time$ (in Equation (2)). We expect the marginal cost of new store opening and operation to decrease with a chain's experience in the focal market as the first-order effect and that such cost savings may become smaller and smaller as time goes by or as the chain ages.

In Equation (5), we specify a chain's own effect on expected market size as

$$M_o^e(H_{it}) = \lambda_1 S_{it} + \lambda_2 e_{it} + \lambda_3 e_{it}\,time, \tag{15}$$

which is a function of own stock of outlets and own experience. For rival's impact on the expected market size, we further allow for the rival's impact to be attenuated by a chain's own experience in the local market as well as the overall age of the chain; therefore,

$$M_r^e(H_{jt}) = \eta_1 S_{jt} + \eta_2 e_{jt} + \eta_3 S_{jt} e_{jt} + \eta_4 S_{jt} e_{it} + \eta_5 e_{jt}\,time + \sum_{l=1}^{L} \nu_l g_{jl}. \tag{16}$$

We expect η_4 to have the opposite sign of η_1 and η_5 to have the opposite sign of η_2. The rival's effect, either positive or negative, is likely to diminish with the accumulation of one's own local or global experience. Substitute the specification of $M_o^e(H_{it})$ and $M_r^e(H_{jt})$ back to Equation (5); we now have the expression for the expected market size: $M_i^e(X_t, H_{it}, H_{jt})$. Plugging this into Equation (4) and collecting terms, we can derive the explicit solution for the optimal number of outlets to open:

$$\begin{aligned} n_{it}^* = {} & \beta_0 + \tilde{\beta} X_t + \tilde{\lambda}_1 S_{imt} + \tilde{\lambda}_2 e_{it} + \tilde{\lambda}_3 e_{it}\,time - \tilde{b}_1 n_{jt-1} \\ & + \tilde{\eta}_1 S_{jt} + \tilde{\eta}_2 e_{jt} + \tilde{\eta}_3 S_{jt} e_{jt} + \tilde{\eta}_4 S_{jt} e_{it} + \tilde{\eta}_5 e_{jt}\,time \\ & + \sum_{l=1}^{L} \tilde{\nu}_l g_{jl}. \end{aligned} \tag{17}$$

Note that $\tilde{\eta}_1 = (\eta_1 - b_2)/(2a_1)$ captures the net effect from the rival. One can not separately identify the positive and negative effects in the reduced-form analysis. The coefficients of the rest of the terms in the second line of the equation are simply scaled by a positive parameter, and the earlier predictions of η and ν under two explanations preserve.[35]

A.2. Correction Term for Sample Selection

Given Z and W and the assumption that η_{im} follows a standard normal distribution that is independent of u_{im}, Equation (11) can be written as

$$\begin{aligned} E(n_{imt} \mid W, Z) = {} & \exp\left[W'_{imt}\beta + \frac{\sigma^2(1-\rho^2)}{2}\right] \\ & \cdot E\left[\exp\left(\rho\frac{\sigma}{\sigma_u}u_{im}\right)\,\middle|\, u_{im} > -(Z'_{im\tau}\alpha + \varepsilon_{im\tau}) \text{ and } u_{im} < \min_{1\le l<\tau}\{-(Z'_{iml}\alpha + \varepsilon_{iml})\}\right]. \end{aligned} \tag{18}$$

We now focus on the second part in Equation (18). Let $\delta = \rho\sigma$ and $\tilde{u}_{im} = u_{im}/\sigma_u$ so that $\tilde{u}_{im}$ follows a standard normal distribution. Furthermore, let $\pi_{im\tau}^L = -(Z'_{im\tau}\alpha + \varepsilon_{im\tau})/\sigma_u$ and $\pi_{im\tau}^U = \min_{1\le l<\tau}\{-(Z'_{iml}\alpha + \varepsilon_{iml})/\sigma_u\}$, which are the lower and upper bounds of u_{im}, respectively. Then we have[36]

$$\begin{aligned} & E\left[\exp(\rho\sigma\tilde{u}_{im} \mid \pi_{im\tau}^L < \tilde{u}_{im} < \pi_{im\tau}^U)\right] \\ & = \exp\left(\frac{\delta^2}{2}\right)\iint \frac{\Phi(\pi_{im\tau}^U - \delta) - \Phi(\pi_{im\tau}^L - \delta)}{\Phi(\pi_{im\tau}^U) - \Phi(\pi_{im\tau}^L)} \cdot f(\pi_{im\tau}^L) f(\pi_{im\tau}^U)\, d\pi_{im\tau}^L\, d\pi_{im\tau}^U. \end{aligned} \tag{19}$$

The double integration reflects that $\pi_{im\tau}^L$ and $\pi_{im\tau}^U$ are random variables as a function of the random error ε_{imt}. Let

$$R_{im}(\delta) = \iint \frac{\Phi(\pi_{im\tau}^U - \delta) - \Phi(\pi_{im\tau}^L - \delta)}{\Phi(\pi_{im\tau}^U) - \Phi(\pi_{im\tau}^L)} f(\pi_{im\tau}^L) \cdot f(\pi_{im\tau}^U)\, d\pi_{im\tau}^L\, d\pi_{im\tau}^U. \tag{20}$$

Equation (18) can be written as

$$E(y_{imt} \mid W, Z) = \exp\left(W'_{imt}\gamma^* + \frac{\delta^2}{2} + \ln R_{im}(\delta)\right). \tag{21}$$

Note that γ^* absorbs the constant $(\sigma^2(1-\rho^2))/2$. Following Orme and Peters (2001), one can approximate $k_{im}(\delta) = \delta^2/2 + \ln R_{im}(\delta)$ by Taylor expansion at $\delta = 0$:

$$k_{im}(\delta) = \sum_{j=1}^{r} \frac{\delta^j}{j!}\kappa_{imj} + o(|\delta|^r). \tag{22}$$

Since δ is finite, the approximation error goes to 0 when we increase the order r. For $r = 1$,

$$\kappa_{im1} = \iint \frac{\phi(\pi_{im\tau}^L) - \phi(\pi_{im\tau}^U)}{\Phi(\pi_{im\tau}^U) - \Phi(\pi_{im\tau}^L)} f(\pi_{im\tau}^L) f(\pi_{im\tau}^U)\, d\pi_{im\tau}^L\, d\pi_{im\tau}^U. \tag{23}$$

Note that π_{im}^L follows normal distribution with mean $-(Z'_{im\tau}\alpha)/\sigma_u$ and standard error $1/\sigma_u$. However, the distribution of π_{im}^U, which is the minimum of $\tau - 1$ random variables, is difficult to derive. We use a simulation method to compute the double integration and obtain estimates of κ_{im1}. Given κ_{im1}, the approximated conditional mean of n_{imt} is

$$E(n_{imt} \mid W, X) = \exp(W'_{imt}\gamma^* + \sigma\rho\kappa_{im1}). \tag{24}$$

[35] Specifically, $\tilde{\eta}_k = \eta_k/(2a_1)$ $(k = 2, 3, 4, 5)$ and $\tilde{\nu}_l = \nu_l/(2a_1)$, where $a_1 > 0$.

[36] It can be verified that, for a random variable η that follows a standard normal distribution, $E[\exp(a\eta_{im})] = \exp(a^2/2)$.

A.3. Simulation

We use simulation to check whether the parameters in the chain expansion model conditional on initial entry can be correctly recovered using the two-step method introduced in §4.1.

We first generate N replications of the data. To make the simulated data comparable to the original data, in each replication, the total number of markets is set at $M = 246$, and the total number of simulation time periods is $T = 20$, which matches the actual data. We also use the same initial condition observed in the data.

For the exogenous variables such as city characteristics and the distance between the focal city and the chain's headquarters, we take the value from the actual data. All the endogenous variables, including each firm's stock of outlets and experience in each period at each city, are generated by the model with a set of prespecified parameters. Since a firm's expansion decision is affected by the rival's expansion, both firms' decisions are generated in the same simulation loop.

The simulation proceeds as follows to generate one sample:

(a) Set the true values for the parameters in the model.

(b) Take the random draws of the error terms. Specifically, for each firm i, we draw (1) M random draws from the normal distribution, $u_{im} \sim \mathrm{N}(0, \sigma_{iu}^2)$, corresponding to the M markets; (2) $M * T$ random draws from the standard normal distribution, $\varepsilon_{imt} \sim \mathrm{N}(0, 1)$; the realized random variables u_{im} and ε_{imt} affect the initial entry decision (Equation (8)); and (3) M random draws from the normal distribution, $\eta_{im} \sim \mathrm{N}(0, 1)$. The unobserved factor v_{im} in the expansion stage is determined by $v_{im} = (\rho/\sigma_u)u_{im} + \sqrt{1-\rho^2}\eta_{im}$.

(c) Given the initial condition, we simulate both firms' entry and expansion decisions forward for each time period in each market. In the simulation loop, for time t of market m:

(i) We check whether chain i has entered city m before t. If not, given the parameters and the drawn unobserved factors u_{im} and ε_{imt}, compute the latent entry variable y^*_{imt} (Equation (8)). The observed entry variable is set to be 1 if $y^*_{imt} > 0$ and 0 otherwise.

(ii) If chain i has already entered, we compute the expected number of new outlets to open $E(n_{imt})$ according to

$$E(n_{imt} \mid W, X) = \exp\left(W'_{imt}\beta + \sigma\frac{\rho}{\sigma_u}u_{im} + \sigma\sqrt{1-\rho^2}\eta_{im}\right), \quad (25)$$

where W_{imt} are the explanatory variables and η_{im} is the drawn random variable. Then draw a realization of n_{imt} from the Poisson distribution with mean $E(n_{imt})$.

(iii) We save the entry and expansion decisions for each chain and update all the endogenous variables (including interaction terms) related to one's own stock, one's experience, its rival's stock, and its rival's experience in market m, which are important input factors for time $t+1$ decisions.

After generating N samples using the procedure sketched above, we estimate each sample using the two-step method illustrated in the model part. In this simulation exercise, we include the following variables in the main model of chain expansion: (1) the variables that are directly related to our tests of the two explanations—rival stock, rival experience, the interaction of rival stock and rival experience, and the growth terms; (2) all the other rival-related variables, including the interaction of rival stock and own experience, rival experience, and time and the number of stores rival opened in time $t-1$; and (3) a variable to control one's own-chain effect and a variable for exogenous city characteristics. Although none of the own-chain-related variables is significant for both chains in the results, we control for own-chain experience for the completeness of model setup.[37] For the initial entry condition, we control for a set of exogenous city

Table A.1 Impact of Restaurant Sales Revenue on Expansion

Variable	McD's	KFC
Intercept	−4.640***	−2.115
	(0.441)	(0.230)
Density	−0.075	0.707
	(0.666)	(0.501)
GDP	0.0005***	0.0002**
	(0.0001)	(0.0001)
Rival's number of outlets opened last	0.017	−0.029***
	(0.011)	(0.009)
McD's stock	0.121	0.682***
	(0.184)	(0.142)
McD's_experience	0.186	0.534***
	(0.200)	(0.132)
McD's_experience × Time	−0.003	−0.032***
	(0.012)	(0.006)
McD's_stock × KFC_experience		−0.042***
		(0.012)
McD's_stock × McD's_experience		0.021***
		(0.012)
KFC_stock	1.093***	0.335***
	(0.242)	(0.112)
KFC_experience	−0.012	−0.156
	(0.178)	(0.110)
KFC_experience × Time	0.009	0.013*
	(0.011)	(0.007)
KFC_stock × McD's_experience	−0.055**	
	(0.022)	
KFC_stock × KFC_experience	−0.046**	
	(0.023)	
Restaurant sales revenue	0.012***	0.002
	(0.003)	(0.002)
Year dummies included?	Yes	Yes
Province dummies included?	Yes	Yes
Number of observations	1,216	1,216
Log likelihood	−683.463	−1,204.386

Notes. Results are from the Poisson model analysis of a firm's store opening decision on the above control variables. Rival's net growth terms are omitted because of the short span of years (2002–2007). Clustered standard errors are in parentheses.

*Significant at the 0.1 level; **significant at the 0.05 level; ***significant at the 0.01 level.

[37] Note that in our original model of Equation (7), controlling for only one's own-chain experience and not one's own stock of outlets yields qualitatively the same results and does not change the main conclusions.

Table A.2 Simulation Results

Variable	True value	Estimate	SD
GDP	0.0005	0.0005	0.0000
Own experience	0.17	0.1758	0.0208
Rival's_number_of_outlets_opened_last	0.09	0.0824	0.0562
Rival_stock	0.40	0.4186	0.1868
Rival_experience	0.15	0.1575	0.0362
Rival_experience × *Time*	−0.004	−0.0042	0.0015
Rival_stock × *Rival_experience*	0.03	0.0271	0.0135
Rival_stock × *Own_experience*	−0.02	−0.0213	0.0116
Rival_growthlag1	−0.2	−0.1936	0.0484
Rival_growthlag2	−0.1	−0.0897	0.0382
Rival_growthlag3	−0.04	−0.0270	0.0272
Correction_term	0.24	0.1629	0.0708
Initial entry condition			
GDP	0.05	0.0376	0.0033
Capital_city	0.68	0.5520	0.4307
Distance_to_headquarters	−0.26	−0.1792	0.1979
Rival_stock	0.40	0.2153	0.1049

characteristics and a rival's stock of outlets. We omit several nonessential variables (e.g., density) to make the simulation clean and clear. Yet all the main variables, regardless of their significance level from our estimation, are kept in the simulation.

Table A.2 reports the means and standard deviations of parameter estimates based on 100 simulated samples, along with the true value of the parameters.[38] The simulation results of the main model are presented in the upper panel. Note that all the key variables for our tests, which are rival stock, rival experience, the interaction of rival stock and rival experience, and rival's net growth in the last few periods, are well recovered. The parameter in front of the correction term (κ_{im}), although not being accurately recovered, still correctly reflects the positive correlation between the unobserved factors in the initial entry and expansion decision as we set the correlation to be positive ($\rho = 0.6$) in the simulation.[39] In addition, the term is able to control for the initial entry condition and leave the rest of the parameters in the expansion model being correctly inferred. The correction term is derived from the estimation of the initial entry decision. Although we are not trying to make an inference from the parameters in the initial entry condition, we also report the corresponding simulation results in the lower panel of the table. Overall, the parameters are close to their true value, but the coefficient in front of rival's stock tends to be underestimated.

In summary, the simulation exercise indicates that the key variables in the main model used to test the two explanations of positive rival effect can be well recovered using the proposed two-step estimation procedure. We also want to point out that the ability to identify the impact of each variable depends on whether the chosen set of parameters can generate enough variation in the simulated chains' expansion paths. Different sets of parameters can lead to different degrees of variation in chains' expansion decisions, resulting in better or worse identification of some variables.

[38] As we apply the same model to both chains' expansion decisions, we report the results for one chain to show whether the model parameters can be correctly identified. The simulation results for another chain are similar.

[39] As derived in Equation (13), the coefficient in front of the correction term κ_{im} is $\rho\sigma$. As ρ is not separately identifiable in the estimation, we only intend to infer the sign of the correlation parameter.

References

Agarwal R, Bayus BL (2002) The market evolution and sales takeoff of product innovation. *Management Sci.* 48(8):1024–1041.

Alestron (2004) Yum! Brands China puts its distribution center into service. (October 25) http://business.highbeam.com/436093/article-1G1-123558586/yum-brands-china-puts-its-distribution-center-into.

Borensztein E, De Gregorio J, Lee J-W (1998) How does foreign direct investment affect economic growth? *J. Internat. Econom.* 45(1):115–135.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Amer. Econom. Rev.* 99(3):864–882.

Caplin A, Leahy J (1998) Miracle on Sixth Avenue: Information externalities and search. *Econom. J.* 108(446):60–74.

Elhorst JP (2010) Spatial panel data model. Fischer MM, Getis A, eds. *Handbook of Applied Spatial Analysis* (Springer, Berlin), 377–407.

Greene WH (1997) FIML estimation of sample selection models for count data. Working paper, Stern School of Business, New York University, New York.

Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47(1):153–161.

Henisz WJ, Delios A (2001) Uncertainty, imitation and plant location: Japanese multinational corporations, 1990–96. *Admin. Sci. Quart.* 46(3):443–475.

Holmes TJ (2011) The diffusion of Wal-Mart and economies of density. *Econometrica* 79(1):253–302.

Imbens G, Wooldridge J (2007) Control function and related methods. Lecture Note 6, NBER Summer Institute, July 31, National Bureau of Economic Research, Cambridge, MA.

Isobe M, Makino S, Montgoery D (2000) Resource commitment, entry timing, and market performance of foreign direct investments in emerging economies: The case of Japanese international joint ventures in China. *Acad. Management J.* 43(3):468–484.

Javorcik BS (2004) Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkage. *Amer. Econom. Rev.* 94(3):605–627.

Jia P (2008) What happens when Wal-Mart comes to town: An empirical analysis of the discount retail industry. *Econometrica* 76(6):1263–1316.

Kuksov D, Villas-Boas JM (2010) When more alternatives lead to less choice. *Marketing Sci.* 29(3):507–524.

Kwok D (2010) McDonald's aims to expand franchise trial in China. *Reuters* (May 6) http://www.reuters.com/article/2010/05/06/us-mcdonalds-china-idUSTRE6451W420100506.

Liu W (2008) *KFC in China: Secret Recipe for Success* (John Wiley & Sons (Asia), Singapore).

Mariotti S, Piscitello L, Elia S (2010) Spatial agglomeration of multinational enterprises: The role of information externalities and knowledge spillovers. *J. Econom. Geography* 10(4):519–538.

Orme CD, Peters SA (2001) A simple two-step estimator for count data models with sample selection. Working paper, University of Manchester, Manchester, UK.

Sault J, Toivanen O, Waterson M (2006) Learning and location. Working paper, University of Warwick, Warwick, UK.

Shanghai Daily (2006) KFC makes it easier to buy franchise. *China View* (April 21) http://news.xinhuanet.com/english/2006-04/21/content_4456737.htm.

Terza JV (1998) Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *J. Econometrics* 84(1):129–154.

Terza JV, Basu A, Rathouz PJ (2008) Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *J. Health Econom.* 27(3):531–543.

Thomadsen R (2007) Product positioning and competition: The role of location in the fast food industry. *Marketing Sci.* 26(6): 792–804.

Toivanen O, Waterson M (2005) Market structure and entry: Where's the beef? *RAND J. Econom.* 36(3):680–699.

Vitorino M-A (2012) Empirical entry games with complementarities: An application to the shopping center industry. *J. Marketing Res.* 49(2):175–191.

Winkelmann R (2001) *Econometric Analysis of Count Data* (Springer, Berlin).

Wooldridge J (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).

Yang N (2013) March of the chains: Herding in restaurant locations. Working paper, Yale University, New Haven, CT.

Zhang J (2010) The sound of silence: Observational learning in the U.S. kidney market. *Marketing Sci.* 29(2):315–355.