Heterogeneous Product Differentiation: Evidence from the Differentiated Agglomeration of Hotel Chains Alliance in China[*†]

Jin Wang[‡] Tianli Xia[§] Yu Zhou[¶]

Jan, 2020

Abstract

This paper documents within and between group product differentiated agglomerations of hotel chains alliance in China, especially the within group agglomeration, which may raise the particular concern from the antitrust authority. We first provide a theoretical model to explain possible channels that might facilitate within-group and between-group competition and results in such agglomeration pattern. Using a unique data of hotel entries at China for international hotel groups, we exploit the exogenous variation of lodging industry market deregulation and identify the within-group and between group competition effects, especially the within-group same type competition effects. We find that the within group competition for the same type or across type are slightly larger than those of between group competition after taking into account of the unobserved market profitability. We will use our analysis to evaluate the impact of the recent merger of Marriott and Starwood in 2016. Our analysis provides theoretical and empirical framework to evaluate the competition effects when the presence of product differentiation of group chains and also shed the light on evaluating the effect of merger.

Keywords: Production Differentiation, Within-Group Competition, Between-Group Competition, Merger, Market Deregulation

Preliminary and Incomplete

---

[*]This paper is previously circulated with the title "Production Differentiation and Market Preemption: The Differentiated Agglomeration of Hotel Chains Alliance in China".

[†]We thank Jeremy Fox, Daniel Xu, Panle Jia and Frank Verboven for their valuable comments and suggestions. We also thank research assistants Yufan Tang and Haorui Wang for their excellent assitant. We also thank the conference participants in Econometric Society China Meeting 2018. All errors are ours.

[‡]Assistant Professor, Department of Economics, Kansas State University; email: jinw@ksu.edu

[§]Ph.D Candidate, Department of Economics, Cornell University email: tx58@cornell.edu

[¶]Corresponding Author, Assistant Professor, School of Economics, Fudan University; email: yuzhou@fudan.edu.cn

1 Introduction

Alliance or Group association becomes a popular business pattern in service industries such as lodging industry, especially when it comes to international markets. In international lodging industry at China, hotel chains become more common to associate with other hotel chains and affiliated in one giant alliance. Observably, in a local market, hotels chains in a group concentrate but differentiate themselves from the high-type to the low-type. Hotel chains from different groups are also observed to be spatially clustered at one place but distinguish themselves from others. We find such phenomenon prevalent in the international hotel industry at China and call it as differentiated agglomeration.

As one feature of such differentiated agglomeration, the agglomeration within hotel group is of great interest by antitrust authority, especially when it comes to merger and acquisition. Merger can potentially hurt non-merged hotel groups if such market agglomeration appears due to mitigated competition within group chains. It could be either due to the cost saving or possible collusion, the latter of which may raise the serious concern. In order to solve the policy debate, it is very important to document such phenomenon, separate out within-group and between group competition effects and decompose mechanisms under it. Conceptually, three distinct mechanisms can drive such within-group agglomeration: the less competition within group than that between groups, or selection on the observables, which means the observed characteristics of the markets attracts the hotel groups to enter, or selection on unobservable, which suggests that the hotel groups can pick up the most profitable markets to enter which is not observed by the researchers.

Using the game theoretic structure, we construct structural entry model to separate out three mechanisms theoretically. Exploiting the unique data with market deregulation of lodging industry in China, we propose a way to empirically separate out and identify within group competition effect from between-group competition effect. Our preliminary results shed light on these policy questions. First, we document that there indeed exists differentiated agglomeration in the lodging industry in China and the within group agglomeration is prevalent in the market. For example, controlling the local market condition, InterContinental Hotel Group (IHG), the largest hotel group in China, indeed has their chains surrounded in the same market in almost 50 percentage the market it enters. Second, our estimates suggest that within group agglomeration in China is mainly driven by the selection on the unobserved profitability rather than less-competition within-group after controlling for the observed market characteristics. Third, we plan to reevaluate the merger case between Starwood and Marriott in China market.

Early literature in lodging industry has provided both theoretical predictions and empirical evidence on the benefit of agglomeration. Chung and Kalnins (2001) suggest that the agglomeration can be treated as a safety signal for travel. Gan and Hernandez (2011) show that the agglomeration can facilitate tacit collusion. While we know quite a bit about the overall agglomeration, we know relatively little about why the strong differentiation along with agglomeration within the group and across the groups, except Mazzeo (2002a). In Mazzeo (2002a), he did discuss the lodging industry product differentiation but his data is abstracted away from alliance identity rather focus on the type of hotels.

Our paper contributes the literature in the following aspects. First, we extend the seminal paper by Mazzeo (2002a, 2002b) and existing studies to allow for heterogeneous product differentiation and distinguish within-group product differentiation and between-group product differentiation. As a caveat, although we can further allow for the heterogeneous product differentiation indexed by group identity, which would require large variations in group market entry, which we did not tempt to pursue in this paper. Second, using the unique feature of data, we propose a way of how to solve the endogeneity issues on the market entry and how to separately identify within group competition (on own type and different type) and between group competition. Such identification strategy and procedure can be also applied to other service industry, like airline alliance, which is also evident in differentiated agglomeration. Third, we use our analysis to address the policy question of merger event.

The rest of the paper is organized as follows: Section II contributes to literature review. Section III devoted to the data description and differentiated agglomeration pattern as well as the overview of lodging industry in China. Section IV is attributed to theoretical model to provide the possible sources of observed differentiated agglomeration. Section V presents the empirical model and the parameter identification. Section VI presents the reduced-form evidence. Section VII shows structural estimation results. Section VIII provides a set of counterfactual analysis on the merger of Marriott and Starwood. Section VIIII concludes.

2 Literature Review

This paper is related to two stands of research: the entry model with production differentiation and chain premium or agglomeration.

Entry games with product differentiation/chain premium: Early work on the entry model literature, including the seminal papers by Bresnahan and Reiss (1990, 1991a, b), Berry (1992), etc., consider the symmetric firms' entry decisions. Later, the literature extends the discussion to the entry with production differentiation. Mazzeo (2002 a, b) shows that homogeneous product differentiation in both market entry and price setting strategy in the lodging industry. Seim (2006) considers the market entry with endogenous product-type choice. Zhu et al. (2009) consider the firm and format differentiation heterogeneity in the retailing industry. Abstract away from the product differentiation, Jia (2008) finds that the retail chain can gain great chain premium by allocating together with their own stores in the neighborhood. Hellenbeck (2016) shows that hotel chains can gain excessive profit (20%) not through higher efficiency but rather information and reputation advantage.

Our paper differs from theirs in the threefold. First, to the best of our knowledge, we are the first one to differentiate within-group product differentiation from the between-group product differentiation. Second, we rationalize that within-group competition and between-group competition can explain the observed product differentiated agglomeration. Third, using the unique feature of shock in data, we can identify and estimate within-group competition and between-group competition separately even with the same type.

Agglomeration in the lodging industry: Geographical agglomeration has been long recognized in the hospitality literature. Baua and Haveman (1997) show that the geographical agglomeration leads to a large size difference and less price difference. Other papers (Kalmins and Chung 2004, 2006; Canina , 2005) find that in the agglomeration, large firms and chain contribute to positive externalities through their reputation while small or low- type hotels taking advantage of them and being willing to locate around big brands. Conlin and Kadiyali (2006) find that idle capacity can deter further market entry. Gan and Hernandez (2013) use dynamic entry model to find that hotel in the cluster has larger possibility to stay in a collusion regime than evenly distributed ones.

Our paper allows for the group heterogeneity in the observed agglomeration while the existing literature ignores the group heterogeneity in the observed agglomeration pattern. In addition, our paper contributes the agglomeration literature by showing that the agglomeration can be formed by the within-group synergy.

3 Data Source

3.1 Data Source

There are three data sources. The main data source on the hotel entry is scratched from a largest online travel agent company at China, the Ctrip.com. Our scratched data includes all the chains in the top six hotel groups until 2018: International Hotel Group, Wyndham, Starwood, Accor, Hilton and Marriott. This data set contains the most important information for hotels, including hotel full name, opening year, full address, rating score, chain name, group association as well as the current capacity. To collect chains in each group comparable, we consider the chains with the rating scores above 3.5 (rating system by Ctrip.com). Ideally, we would use the national star rating system to categorize the hotel type, but not every hotels have applied to national star rating. Instead, we use the rating system by Ctrip.com, which we will explain in the details.

The second dataset is the economic and demographic characteristics, drawn from the China Regional Economic Statistical Yearbook in 2000-2014 and Regional Annual Statistical Yearbook in 2015-2018. In this dataset, it includes the most important local economic and demographic information at the prefectural level city, including GDP, population, population density, per capita income, average salary, transportation, the presence of the airport, the total number of domestic tourists and foreign tourists. Using this data, we can capture the local market characteristics that drive the entry of the international hotels.

The third dataset is the geographic information of the hotels, their group headquarter and city centers. We manually search the address of hotels and their headquarters and city center in Google API to get the complete latitude and longitude information. Then we construct the geographic distance of the latitude and longitude to construct the distance between the city center and group headquarter, which is important to construct our instruments later.

Table 1 gives summary statistics for our scratched data and shows the top ten international hotel groups in China. Among the top ten international group in China, the six leading groups are IHG, Wyndham, Starwood, Accor, Hilton and Marriott. The market share of six leading group is approximately 90% of among the top ten international hotel groups.

3.2 Hotel Type Categorization

In order to measure local market competition and agglomeration level, it is important to categorize the chains into different types. Following the standard literature, we categorize hotel chains into high-type and low-type following the Ctrip.com standard rating system and international categorization. More specifically, this rating system is a comprehensive rating system, including the national star rating (if available), facility, service, amenity and brand awareness as well as reputation.[1][2]

Table 2 provides the detailed rating scores for each hotel on Ctrip.com in our dataset. It shows that even within the same chain, the rating score varies across different locations. To define the hotel type for a chain, we average their score and categorize the hotel as the high-type if its average score is above 3.5 and low-type otherwise. Such categorization is roughly matched with Ctrip's own category about luxury hotel chains. In addition, we compare this categorization with international categorization for the same chain in other countries. It also roughly matches the international categorization. From Table 2, it shows that IHG have the almost equal proportion on both high-type and low-type hotels. Wyndham and Accor are more specialized on the low-type hotels and Starwood, Hilton and Marriott focus on the high-type hotels.

[1] The rating score is categorized as 6-diamond level, which is based on a collective rating factors: including the national star rating (if available), facility, service, amenity and brand awareness as well as reputation. In particular, 6-diamond: almost equivalently represent the user-valued 5-star hotels; 5-diamond: almost equivalently represent the user-valued 5-star and some outstanding 4-star hotels; 4-diamond: almost equivalently represent the user-valued 4-star and some outstanding 3-star hotels; 3-diamond: almost equivalently represent the user-valued 3-star and some outstanding 2-star hotels; 2-diamond: almost equivalently represent the user-valued 2-star and some outstanding 1-star hotels; 1-diamond: almost equivalently represent the user-valued 1-star; half-diamond: represent the hotels between the two diamond level.

[2] Using this rating system score, we further compare it with the international categorization. On average, it matches the international categorization for each chain, and it corresponds to the upper-middle demand in lodging industry.

3.3 Market Definition

In this paper, we define a market as an administrative prefectural-level city, excluding Tibet, Xinjiang, Qinghai, Hong Kong, Macau and Taiwan. The market definition for us is important for our identification and estimation. To some extent, we should not define the market too small to have almost no between-group competition or agglomeration, while at the same time, we should not define the market too big to have to dense differentiated agglomeration. We choose the prefecture-level cities mainly because most hotels are located at the prefecture-level cities and we have complete economic and demographic information on the prefecture-level cities. Though it is natural to follow the administrative prefecture-level boundary, when it comes to the metropolitan areas, the hotels are all clustered in the nearby neighborhood. As a result, we use the district as a market for the metropolitan areas for four municipalities: Beijing, Shanghai, Tianjin and Chongqing. However, since group headquarters are located either in Beijing or Shanghai, we exclude these two cities as well in our final analysis. As a result, it results in 281 markets with 260 administrative prefectural-level city and 21 districts.

Table 3 presents the economic and demographic characteristics of the prefectural-level cities. It shows that the economic and demographic characteristics vary across the regions, which highly reflects the regional difference in China for the first tier of prefectural-level cities and the second and third tier of prefectural-level cities. It provides the rich market variation to allow us to identify the model parameter for the market characteristics.

4 Overview of Lodging Industry in China

4.1 History

Lodging industry is one of the most dynamic sectors of service industry in China. The development of the lodging industry can be divided into three key episodes: 1978-2001, 2002-2008, 2009-2018.

Chinese tourism and lodging industry were initiated since 1978. In the beginning of the development, the number of hotels is quite limited, and the hotel enterprises has been controlled by the government to that associated with political needs (Zhang, Pine & Lam, 2005). In 1988, a star rating system has been established, which allows for the comparison on the hotels across the nation. Until late 2001, international hotel chains host almost all the foreign tourists for political and safety reasons.

After December 2001 when China formally joined in WTO, as one type of foreign direct investments, China lowers the barrier of the entry for the international hotels in China. Within four years grace periods until 2006, China freely open the license market and make the entry much easier than before. International hotel groups can have direct investment in China rather than joint venture. Stimulated by the World Olympic hosted in China at 2008, there is a surge of international hotels. In 2002, Rujia, the first domestic budget hotel chain enters the market.

Later in 2009, largely due to the global economic recession, China experiences a slowdown on the new constructions. Then, these international hotel groups quickly find a new expansion on the domestic demand and expand towards the secondary cities in China.

The international hotels experience a sharp increase since 2002. At the same time, the total number of domestic hotels has also increased and exceed the total number of the international hotels. Along with the rapid increase on the number of hotels, around 2002 the international tourists and domestic tourists but late 00s, the domestic tourists surge and constitutes a large portion of the lodging industry demand.

4.2 Overview

The motivating observation of this paper is differentiated agglomeration, in particular, within-group agglomeration: agglomeration on the same type and different type. Below, we will provide the graphical illustration on the overall agglomeration pattern and the statistics on the within-group agglomeration and between-group agglomeration before we closely work on the model and parameter estimation.

4.3 Differentiated Agglomeration Pattern: Graphical Illustration

Figure 1 plots the top three international hotel groups market agglomeration at the province level. It is easy to see that in the province level, the density of hotels declines from the most developed regions east coast area to the inter-northwest regions. It largely reflects the expansion pattern of international hotels in China from the first-tier cities to the second-tier cities and then the third-tier cities.

Figure 2 present within-group and between-group product differentiation. It is difficult to illustrate across all six groups, so we will use the agglomeration pattern of top two leading groups, IHG and Wyndham as an illustration.

Within group agglomeration: Figure 2.1, represents within-group same-type and cross-type agglomeration within IHG. In addition, Figure 2.2 represents within-group same type and cross-type agglomeration. We find that no matter in what group, it is evident that the high-type hotels and low-type hotels within the group are agglomerated at some particular loci. More specifically, we find that (1) for some spots, there are more than two high-type hotels or two low-type hotels, which is the within-group same-type agglomeration, and in addition (2) for some spots, we find that a high-type hotel is clustered by a few low-type hotels, which is the within-group cross-type agglomeration.

Between Group Agglomerations: Figures 2.3-2.4 represent between-group same-type agglomeration. Figure 2.5-2.6 represent between-group cross-type agglomeration instead. Comparing Figure 2.3 to Figure 2.4, we find that between-group low-type agglomeration is more pronounced than the between-group high-type agglomeration. In addition, comparing Figure 2.5 to Figure 2.6, we find that between-group cross-type agglomeration are almost identical for both two groups and the cross-type agglomeration is clustered more than the same-type agglomeration.

4.3.1 Differentiated Agglomeration Pattern: Hotel Market Configuration

To further show the evidence of the agglomeration pattern, we will give the numerical evidence. Table 4 presents the market configuration of the different entry pattern for each group across the cities. In particular, we count the number of markets which has a particular entry pattern. For example, the number in the table of IHG and OHH represents the number of markets of IHG which has more than one high-type hotels. To give a second example, the number in table of IHG and OHL represents the number of market of IHG has high-type and low-type at the same location. It is easy to check that within group product differentiation, especially, own type agglomeration is quite pronounce besides the between-group product differentiation.

Table 5 shows how the entry pattern varies across local market structure. We separate the markets into four categories: >20 hotels, 20-11, 10-5, 5-2. It shows that the number of each market structure increases which indicates the local market profitability. We further found that the within group own-type agglomeration decreases as the local market profitability becomes thin but still exists in some markets.

4.3.2 Differentiated Agglomeration Pattern: Correlation Analysis

To further show the evidence of the agglomeration pattern, especially within-group agglomeration. We will give the numerical evidence on the correlation within-group and between-group. More specifically, we calculate the two correlation statistics based on the two exogenous shocks in 2001 and in 2008: (1) the correlation of the new entries between 2002 and 2008 with the existing stock of hotels in 2001; (2) the correlation of the new entries between 2009 and 2016 with the existing stock of hotels in 2008. Table 6 summarizes within group correlation for the top six international hotel groups. It shows that in six hotel groups, for the within group correlation, we find that the low-type new entries are more likely to the areas with high-type existing hotels (the correlation LH_own). The correlation is higher for the period of 2009-2016. In addition, the correlation on the low-type new entries are also more likely to the areas with low-type existing hotels (the correlation LL_own), except Hilton, as Hilton is targeting on the high-end hotels. Third, we find the mixture evidence on the high-type new entries on the areas with existing low-type stock of hotels. Fourth, we also find the mixture pattern for the high-end hotel entries on the areas with existing high-end hotels. It seems that there are some hotel group specific strategy. For IHG and Marriott, the high-end hotels have a relatively strong tendency to enter the areas with existing high-end hotels. But for Wyndham and Accor are not. For Starwood and Hilton, the evidence is not significant.

5 Reduced-Form Results

Table 7 lists the top twenty cities (except Tianjin and Chongqing) with largest number of hotels of six hotel groups. These cities represents the cities with large market potential. Our conjecture is that within group agglomeration will be more pronounced in the cities if within group agglomeration is driven by observed market potential. Table 8 lists top twenty cities with largest number of hotels for single hotel group, representing within group agglomeration in each city. Comparing Tables 7 and 8, the first thing to note is that the cities listed in both tables are almost the same. What is more interesting the rank of cities are almost the same in two tables. Tables 7 and 8 provide the evidence that the selection on observed market potential can be one source of within agglomeration. But can the selection of observed market potential explain all the within group agglomeration in our data?

To answer this question, we further consider a reduced form regression to control for the observed market potential. We pool all the market entry outcome for each group together and consider the following bivariate Probit model for high-type and low-type hotels. Table 7 summarizes the regression result from the bivariate Probit model. Noticablely, controlling for observed market characteristics in our data, the coefficients on competing products from own hotel group, or different/same type products from competitor groups are almost all positive and statistically significant at the 0.1 percent level. The results are similar for both high-type hotels and low-type hotels.

What can we conclude from these results? The basic fact in raw data is that the markets with more unobserved (by researchers) market profitability where a hotel group would like to enter are the markets other hotel groups would also like to enter. In other words, there is a strong selection on the unobserved market profitability. Without taking account of this fact, we may make a wrong inference that the competing products entry will induce the hotel group own entry. Now, a central empirical challenge of evaluating the competition effects of hotel entry is separating the true competition effects of hotel groups from correlation of unobserved hotel group heterogeneity within the market, possibly due to common market profitability. This suffices the necessity of the structural estimation.

6 Theoretical Model

7 Theoretical Insights

This section presents a theoretical model about the competition between hotel groups. We start with consumer utility function, deriving consumer demand for different types of hotels and the profit functions of hotel groups, then discuss the equilibrium entry and product-type decisions of hotel groups. Through the theoretical model, we show how a hotel group's profit depends on the types of hotels operated by itself and by competing hotel groups, which can help to justify the setup of the empirical model in the next section. Moreover, the theory sheds lights on impacts of the factors, such as the consumer preferences and firm operation costs, on the hotel groups' optimal entry and product-type decisions. This can help us understand the hotels' location observed in the data and interpret the empirical results.

Suppose there are N hotel groups, each of which offers high-end and low-end hotels. Firm $i \in \{1, 2, ..., N\}$ can choose to locate its high-end hotels (H) and low-end hotels (L) in a market. In the market, a representative consumer's preference for different types of hotels can be expressed as a quadratic-formed utility function:

$$U\left(q_{1H}, q_{1L}, q_{2H}, q_{2L}\right) = \alpha_H \sum_{i=1}^{N} q_{iH} + \alpha_L \sum_{i=1}^{N} q_{iL} - \frac{\beta_h}{2} \sum_{i=1}^{N} q_{iH}^2 - \frac{\beta_l}{2} \sum_{i=1}^{N} q_{iL}^2$$
$$- \delta \left(\frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} q_{iH} q_{jH} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} q_{iL} q_{jL} \right) - \delta\gamma \sum_{i=1}^{N} q_{iH} q_{iL} - \delta\eta \sum_{i=1}^{N} \sum_{j \neq i}^{N} q_{iH} q_{jL}$$

where q_{iH} and q_{iL} are the number of high-end and low-end hotels provided by firm i. The consumer perceives these products as substitutes with each other. Depending on the providers and qualities of hotels, we distinguish three types of cross-product substitutability: (1) For any two hotel products which have the same quality but are provided by different hotel groups,such as iH and jH $(i \neq j)$, or iL and jL $(i \neq j)$, their cross-product substitutability is measured by δ. (2) For any two products which are from the same hotel group but of different qualities, like iH and iL, their cross-product substitutability is measured by $\delta\gamma$. (3) For any two hotel products products which have different qualities and are from different groups, such as iH and jL, their substitutability is measured by $\delta\eta$. The parameters $\{\beta_l, \beta_h, \delta, \gamma, \eta\}$ satisfy Assumption 1.

Assumption 1. $\beta_l > \beta_h \geq \max\{\delta\gamma, \delta\} \geq \min\{\delta\gamma, \delta\} > \delta\eta > 0$.

Assumption 1 demonstrates the relative magnitudes of the three types of cross-product subsitutabilities. First, $\delta\eta$ is strictly less than $\delta\gamma$ and δ. It implies that from the consumer's perspective, the products of different qualities and from different hotel groups are least substitutable with each other. Second, note that Assumption 1 does not impose any restriction on the relative magnitude of δ and $\delta\gamma$ because the relative size of these two types of cross-product subsitutabilities could vary by consumer preferences. In reality, some consumers have strong preference on the quality than the brand of hotel group. They constantly choose hotels of good quality while do not care too much about whether the hotels are operated by IHG or Wyndham. For this type of consumers, the products with the same quality, such as iH and jH $(i \neq j)$, are more substitutable than the products within the same group, i.e., $\delta > \delta\gamma$. In the mean time, there could exist another type of consumers who have stronger preference on the hotel group brand. They prefer hotels from the same hotel group and care less about the quality. This could be partly due to the premium membership rewards program, a business strategy often used by large hotel groups like IHG. The consumers who frequently visit the chain hotels of the same group can accumulate rewards points and redeem them for free hotel nights in the future. In this circumstance, consumers perceive hotels provided the same group more substitutable: $\delta\gamma > \delta$. Last, Assumption 1 states that all of the three types of "cross-product" substitutability, δ, $\delta\gamma$ and $\delta\eta$, are smaller than "own-product substitutability", β_l and β_h, which is consistent with economic intuition.

The consumer faces a budget constraint $y = \sum_{i=1}^{N} p_{iH} q_{iH} + \sum_{i=1}^{N} p_{iL} q_{iL}$. Solving the utility maximization problem, we obtain the inverse demand functions for each type of hotel products offered by firm i:

$$\begin{array}{ll} (q_{iH}) & \alpha_H - \beta_h q_{iH} - \delta\gamma q_{iL} - \delta \sum_{j\neq i}^{N} q_{jH} - \delta\eta \sum_{j\neq i}^{N} q_{jL} = p_{iH} \\ (q_{iL}) & \alpha_L - \beta_l q_{iL} - \delta\gamma q_{iH} - \delta \sum_{j\neq i}^{N} q_{jL} - \delta\eta \sum_{j\neq i}^{N} q_{jL} = p_{iL} \end{array} \tag{1}$$

As for operation cost, it assumes the per-unit cost of hotel group i is c. When hotel group i provides different types of hotel products in the same region, it can benefit from the economics of scope and save its per-unit cost. For simplicity, we assume this within-group per-unit cost reduction is a linear function of the number of hotels of firm i : $s_h^o q_{iH} + s_l^o q_{iL}$. When different hotel groups enter the same market, it brings benefits of industry clustering to each firm, such as reducing consumer's costs of finding a hotel or switching between different brands of hotels, or generating a positive spillover effect of one firm's advertising on the business of other firms. We formulate the cross-group spillover effect by assuming the clustering of hotel groups reduces the per-unit operation cost of firm i by $s_h^c \sum_{j\neq i}^{N} q_{jH} + s_l^c \sum_{j\neq i}^{N} q_{jH}$. The presence of within-group and cross-group spillover effects attenuates the competition among different hotel products. To ensure competition effect still dominates, we impose the following restrictions on $\{s_h^o, s_l^o, s_h^c, s_l^c\}$.

Assumption 2. $\delta\eta > \max\{s_h^o, s_l^o, s_h^c, s_l^c\} > 0$

Firm i's profit on high-type and low-type hotel products can be written as

$$\begin{aligned} \pi_i^H &= \left(p_{iH} - c + s_h^o q_{iH} + s_l^o q_{iL} + s_h^c \sum_{j\neq i}^{N} q_{jH} + s_l^c \sum_{j\neq i}^{N} q_{jH} \right) q_{iH} \\ &= \left[\alpha_H - (\beta_h - s_h^o) q_{iH} - (\delta\gamma - s_l^o) q_{iL} - (\delta - s_h^c) \sum_{j\neq i}^{N} q_{jH} - (\delta\eta - s_l^c) \sum_{j\neq i}^{N} q_{jL} - c \right] q_{iH} \quad (2) \\ \pi_i^L &= \left(p_{iL} - c + s_h^o q_{iH} + s_l^o q_{iL} + s_h^c \sum_{j\neq i}^{N} q_{jH} + s_l^c \sum_{j\neq i}^{N} q_{jH} \right) q_{iL} \\ &= \left[\alpha_L - (\beta_l - s_l^o) q_{iL} - (\delta\gamma - s_h^o) q_{iH} - (\delta - s_l^c) \sum_{j\neq i}^{N} q_{jL} - (\delta\eta - s_h^c) \sum_{j\neq i}^{N} q_{jH} - c \right] q_{iL} \end{aligned}$$

It shows that for each type of hotel products, the profit function is a decreasing function of the number of competing hotel products which are either offered by the same hotel group or rival hotel groups. In equilibrium, firm i maximizes its total profits $\pi_i^H + \pi_i^L$ by choosing the two types of hotels, q_{iH} and q_{iL}. The optimal number of each type of hotels is determined by three sets of parameters: (1) $\{\alpha_H, \alpha_L\}$ which can be treated as a proxy of the market capacity for high-type and low-type hotels. (2) $\{\beta_h, \beta_l, \delta, \delta\gamma, \delta\eta\}$ which represents consumers' perceived subsitutability of different hotel products. (3)

$\{s_h^o, s_l^o, s_h^c, s_l^c\}$ which captures the cost-saving benefits caused by within-group and cross-group clustering.

We conclude this section by summarizing the implications of our theory and its relationship with the empirical study : first, the theory shows that the profit of a hotel brand is correlated with the number of its own hotels and the competing hotels in the same market. The competing hotels can be of three kinds: the hotels provided by the same hotel group but of different quality, and the hotels from different hotel groups with same or different qualities. In the following empirical study, we will follow this idea, formulating the profit of a hotel brand as a function of its own hotels and three kinds of competing hotels, and estimate their effects. As shown by (2), the estimates of these effects capture the substitutabilities between hotel products net of the cost-saving caused by within-group or cross-group clustering. If Assumption 2 holds, it can be expected that these estimates are all negative. That is, a hotel brand's profit decreases with the number of hotels in a market. As for the relative magnitudes of these four effects, it depends on the size of $\{\beta_h, \beta_l, \delta, \delta\gamma, \delta\eta\}$ versus $\{s_h^o, s_l^o, s_h^c, s_l^c\}$. Suppose $\{\beta_h, \beta_l, \delta, \delta\gamma, \delta\eta\}$ plays a dominant role, we will obtain a pattern as suggested by Assumption 1: the estimate of "own-product substitutability" is larger than any of "cross-product substitutability".

Second, the equilibrium entry and product type decisions are determined by three sets of parameters: the proxy of market capacity $\{\alpha_H, \alpha_L\}$, the cross-product substitutability $\{\beta_h, \beta_l, \delta, \delta\gamma, \delta\eta\}$ and the cost reduction $\{\theta_f, \theta_t\}$ generated by within-group or across-group clustering. In reality, the cross-product substitutability often varies by the consumers' preferences and demographic characteristics of a market, which can be either observable or unobservable to researchers. The levels of cost saving can vary by hotel groups and by markets, and more importantly, depend on the strategic location decision of hotel groups. In order to empirically study hotel groups entry and product-type decisions, it requires us to control market and firm level characteristics, especitally those unobserved common market shocks. The resulted identification problems and solutions will be discussed in the next section.

8 Structural Model

Following some insights derived from the theoretical model, in this section, we will move to the empirical model to quantify within-group and between-group competition associated with within-group and between-group product differentiation.

8.1 Potential Profit Function

We consider the potential profit function as follows. The profit for a high-type (low-type) hotel (H/L) of group i at market m, for $i = 1 \dots G$, $m = 1 \dots M$, is written as

$$\begin{aligned} \Pi^*_{im,H} &= X'_m\beta_H + \Gamma_{i,H}\left(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \boldsymbol{\gamma}\right) + \varepsilon_{im,H} \\ \Pi^*_{im,L} &= X'_m\beta_L + \Gamma_{i,L}\left(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \boldsymbol{\gamma}\right) + \varepsilon_{im,L} \end{aligned} \quad (3)$$

where X_m is the vector of observed market-specific characteristics and $(\varepsilon_{im,H}, \varepsilon_{im,L})$ is the vector of unobserved group-market specific characteristics, that are observed by the hotels but not by the researchers. We assume that the unobservable $(\varepsilon_{im,H}, \varepsilon_{im,L})$ are correlated across the group and across the market but independent of the market-specific characteristics. In particular, we decompose $\varepsilon_{im,H}$ as follows

$$\begin{aligned}\varepsilon_{im,H} &= r_1 \varsigma_m + \sqrt{1-r_1^2}\varsigma_{im,H} \\ \varepsilon_{im,L} &= r_1 \varsigma_m + \sqrt{1-r_1^2}\varsigma_{im,L}\end{aligned}$$

where ς_m is a common market shock; ς_{im} is a group-market specific shock and $\varsigma_{im,H}$ is a group-market-type specific shock. An example of a market-level shock is local market amenity. Group-market schoks are those shocks like the demand, cost and other logistic features, specific to a group in a particular market. To simplify the notation, we omit the index m without further explanations.

In this specification, we are particularly interested in the function $\Gamma_{i,H}$ and $\Gamma_{i,L}$ which captures the competition for a given differentiated agglomeration pattern. In each Γ function, $(\mathcal{N}_{m,H}, \mathcal{N}_{m,L})$ represents the market configuration of high-type and low-type chains for market m. We can further decompose $(\mathcal{N}_{m,H}, \mathcal{N}_{m,L})$ as

$$\mathcal{N}_{m,H} = (N_{im,H}, N_{-im,H}), \mathcal{N}_{m,L} = (N_{im,L}, N_{-im,L})$$

where $N_{i,,H}$ and $N_{i,,L}$, are the number of high-type or low-type chains for group i in market m; and in addition, $N_{-i,,H}$ and $N_{-i,,L}$ are the number of high-type or low-type chains for the other groups in market m. By doing, we can possibly separate out the competition effects within the group and the competition effects across the groups. In particular, we can write $\Gamma_{i,H}(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \boldsymbol{\gamma})$ and $\Gamma_{i,L}(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \boldsymbol{\gamma})$ as follows,

$$\begin{aligned}\Gamma_{i,H}(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \theta) &= \gamma^o_{hh} f^o_{hh}(N_{im,H}) + \gamma^o_{lh} f_{lh}(N_{im,L}) + \gamma^c_{hh} f^c_{hh}(N_{-im,H}) + \gamma^c_{lh} f^c_{lh}(N_{-im,L}), \\ \Gamma_{i,L}(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \theta) &= \gamma^o_{hl} f^o_{hl}(N_{im,H}) + \gamma^o_{ll} f_{ll}(N_{im,L}) + \gamma^c_{hl} f^c_{hl}(N_{-im,H}) + \gamma^c_{ll} f^c_{ll}(N_{-im,L}).\end{aligned}$$

In $\Gamma_{i,H}(\mathcal{N}_{m,H}, \mathcal{N}_{m,L}; \theta)$, the coefficients $(\gamma^o_{hh}, \gamma^o_{lh})$ associated with the first two terms present the competition effects for a H-type chain from H-type and L-type competitors *within* the group. The coefficients $(\gamma^c_{hh}, \gamma^c_{lh})$ associated with the last two terms present the competition effects for a H-type chain from H-type and L-type competitors *across* the rest groups. Our specification is consistent with the theoretical model prediction, that is, the profit function for each hotel depends on the within group own type competition, within group across type, between group own type and between group across type competition. Essentially, our specification resembles the specification in Mazzeo (2002a, 2002b), but we are able to distinguish the *within-group* competition effects from the *between-group* competition effects.

8.2 Conditions for Nash Equilibrium

In each market, we assume that the competitive groups make their entry and product-type decision simultaneously. Second, we assume that for each group, the decision proceeds in two substage. In the first substage, for each market m (for the notation convenience, we omit the subscript m in the following notation) fully anticipating other groups' entry and type decision $(N_{-i,H}, N_{-i,L})$, group i decides whether to enter with the maximum number of hotels $(N_{i,H}, N_{i,L})$ for which there is some $(N_{i,H}, N_{i,L})$. In the second substage, the chains within a group simultaneously decide to enter with their types, resulting with the configuration $(N_{i,H}, N_{i,L})$.

A pure strategy obtains, when holding the strategy of rival groups fixed, all entering hotels of two types in a group are profitable and all hotels in a group of two types except the profit below zeros. In particular, we chacterize a group's strategy with the following condition: given other group's entry and type strategy, each group operate the most profitable type hotel when it is profitable and its profitability yields a payoff exceeds the alternative type and will does not operate a hotel otherwise. Formally, given the potential profit function, a Nash equilibrium can be represented by a vector of $((N_{1,H}, N_{2,L}), \ldots, (N_{G,H}, N_{G,L}))$ for which the following conditions hold: for fixed $\left(N^*_{-i,H}, N^*_{-i,L}\right)$,

$$\Pi^*_H \left(N_{i,H} - 1, N_{i,L}, N^*_{-i,H}, N^*_{-i,L}\right) > 0 \text{ and } \Pi^*_L \left(N_{i,H}, N_{i,L} - 1, N^*_{-i,H}, N^*_{-i,L}\right) > 0;$$

and

$$\begin{aligned} \Pi^*_H \left(N_{i,H} - 1, N_{i,L}, N^*_{-i,H}, N^*_{-iL}\right) &> \Pi^*_L \left(N_{i,H} - 1, N_{i,L}, N^*_{-i,H}, N^*_{-i,L}\right); \\ \Pi^*_L \left(N_{i,H}, N_{i,L} - 1, N^*_{-i,H}, N^*_{-i,L}\right) &> \Pi^*_H \left(N_{i,H}, N_{i,L} - 1, N^*_{-i,H}, N^*_{-i,L}\right). \end{aligned}$$

It is easy to show that a Nash equilibrium exists in each market. But these threshold conditions are sufficiently weak. As a result, the certain value of unobserved shocks in a market can be consistent with multiple outcomes. In particular, we have two types of multiple equilibria.

The first type of multiple equilibria exists because such equilibrium condition only restricts on the relative magnitude. It is possible that for some market, that can support high-concentration market configuration can automatically sustain low-concentration market configuration.

The second type of multiple equilibria exists due to the fact that we assume symmetric hotel groups and symmetric potential product function. One example of multiple equilibrium can be constructed: for a given set of unobserved shocks, holding other groups' strategy, any two pairs of symmetric equilibrium for group i and group j can coexist.

For the first type of multiple equilibria, we will assume that the hotel group will seek all market profitability and the market is always saturated, so that we exclude the low-concentration market configuration.

For the second type of multiple equilibria, we will consider different equilibrium selection rules. The first equilibrium selection rule is we consider the average out of the equilibrium. The second equilibrium selection rule specification is to select the equilibrium that favors for the group with relatively higher market power. More specifically, if we choose two multiple equilibria are related with IHG and Wyndham, we will choose the one that favors for IHG and similar analogy can be applied to other groups.

8.3 Parameter Identification

In this subsection, we will briefly discuss the identification of model parameters. Our key parameter set includes: (1) within-group same type effects $(\gamma^o_{hh}, \gamma^o_{ll})$; (2) The within-group same type effects: $(\gamma^o_{hh}, \gamma^o_{ll})$; (2) The within-group across-type effects: $(\gamma^o_{hl}, \gamma^o_{lh})$; (3) The between-group same type effects: $(\gamma^c_{hh}, \gamma^c_{ll})$; (4) The between-group across-type effects: $(\gamma^c_{hl}, \gamma^c_{lh})$.

8.3.1 Berry's type of Identification

Consider a simple model with

$$\begin{aligned}
\Pi^*_{i,m,H} &= z_i\alpha + \gamma_{hh} f_{hh}\left(N_{i,H}\right) + \gamma_{lh} f_{lh}\left(N_{i,L}\right) + \delta_{hh} f_{hh}\left(N_{-i,H}\right) + \delta_{lh} f_{lh}\left(N_{-i,L}\right); \\
\Pi^*_{j,m,H} &= z_j\alpha + \gamma_{hh} f_{hh}\left(N_{j,H}\right) + \gamma_{lh} f_{lh}\left(N_{j,L}\right) + \delta_{hh} f_{hh}\left(N_{-j,H}\right) + \delta_{lh} f_{lh}\left(N_{-j,L}\right)
\end{aligned}$$

Berry (1992)'s identification strategy implies the identification below in our context. We will show the parameter identification of γ_{hh} explicitly and other terms identification can show in the similar fashion. To identify the parameter γ_{hh}, we observe that an additional hotel enters of group i in market m but not no additional hotel of group j enters, suggesting that $\Pi^*_{i,m,H} > \Pi^*_{j,m,H}$. Now, suppose that $N_{i,L} = N_{j,L}$, $N_{-i,H} = N_{-j,H}$ and $N_{-i,L} = N_{-j,L}$. It implies that

$$z_i\alpha + \gamma_{hh} f_{hh}\left(N_{i,H}\right) > z_j\alpha + \gamma_{hh} f_{hh}\left(N_{j,H}\right)$$

which means by simply exchange terms,

$$z_i\alpha - z_j\alpha > \gamma_{hh}\left(f_{hh}\left(N_{i,H}\right) - f_{hh}\left(N_{j,H}\right)\right).$$

Then γ_{hh} will be identified by the variation of z_j if $N_{i,H} \neq N_{j,H}$. Now by the same argument, we can identify γ_{lh}, δ_{hh} and δ_{lh} if there exists $z_{j'}$, $z_{j''}$and $z_{j'''}$ for group j', j'' and j''. This means that to make identification work, there exist (1) at least five groups to compete and (2) have large variation on the entry of hotels across the groups. Our data have six groups, which satisfies the first condition. However, the variation on the entry of hotels is listed in Table 3, which are not as rich as expected, except for a few markets, like the top 15 cities listed in the data.

8.3.2 Alternative identification

Instead only relying on the identification above, our data has a unique institutional feature of market deregulation. We propose to use the old equilibrium outcome before deregulation as an instrument to identify competition effect, especially own product competition effect. The idea is outlined as below. We observe that when the absence of exit (for simplicity), the following accounting identity holds

$$\mathcal{N}_t = \mathcal{N}_{t-1} + I_t;$$

where $\mathcal{N}_t$ is the new equilibrium of a single period of longer dynamic games and $\mathcal{N}_{t-1}$ is the old equilibrium outcome of a single period of a longer dynamic games. New investment I_t is resulted from emerging market profitbility after the old equlibrium is attained. Following this logic, we define the prediction error ϵ_t in new investment as

$$\begin{aligned} \epsilon_t &\equiv I_t - E\left(I_t | \mathcal{I}_{(t-1,t)}\right) \\ &= I_t - f\left(\mathcal{I}_{(t-1,t)}\right) \end{aligned}$$

where $\mathcal{I}_{(t-1,t)}$ are the information (observable characteristics) available betwen $t-1$ and t. Then it follows that

$$\mathcal{N}_t = \mathcal{N}_{t-1} + f\left(\mathcal{I}_{(t-1,t)}\right) + \epsilon_t$$

Note that we make an explicit assumption, rational expectation, to derive the equation above. That is, the market potential realized in time $t-1$ is fully obsorbed in $\mathcal{N}_{t-1}$ and will not persistent over times (i.e. no serial correlation).

Now, we ask what are the potentially valid instruments in our data and why they will satisfy the assumptions above. We propose two sets of instruments i,e old equilibrium drawn from our data. The first set of instrument is the number of hotels entered before 2002, represeting the old equilibrium in the regulated lodging industry. The second set of instrument is the number of hotels entered before 2009, representing the old equilibrium induced by WTO entry after 2002 and Beijing Olympic in 2008. Below, we discuss why these two instruments will more plausibly satisfy assumptions (independence and correlation) on the instruments in our content.

Recall that Figure 1 represent three phrases of lodging industry development in China for the 5-star and 4 star hotels entry in China. The turning points are shown around 2001 and 2008. We plot the similar patten in Figure 1 using the data in estimation. We find almost the same pattern to track each other over the years. In addition, the turning points points appear around 2001 and 2008. We argue that these turning points represent the old equilibria for the first two phrases. Furthermore, we argue that these two equilibria points are less likely to correlate with unobservable charateristics in the market that drives the new entry after 2002 and 2008. The total number of hotels in 2001, $\mathcal{N}_{2001}$, is driven by the political needs and assigned by the government authority. Since it is driven by non-profitable reason, it is less likely to correlated with the market potential that induces the new entries. The total number of hotels in 2008, $\mathcal{N}_{2008}$, might be another possible instrument, driven by market deregulation and Olympic game in Beijing. We argue that the entries carried in $\mathcal{N}_{2008}$ are highly possible driven by the local market profitability variation in that period. To be conservative, we use $\mathcal{N}_{2001}$ as an instrument in our papers.

Next, we show the correlation between $\mathcal{N}_{t-1}$ and $\mathcal{N}_t$. If the accounting identity holds, we will expect that in our data, the estimated coefficient on $\mathcal{N}_{t-1}$ is more or less close to one. We report our regression results using equation (**??**) in online appendix to save the space The dependent variable is the observed entry outcome of new equilibrium in 2016, $\mathcal{N}_{2016}$. The explatory variables are the observed entry outcome in old equilibrium $\mathcal{N}_{2001}$ and flexible polynomial function for $f\left(\mathcal{I}_{(t-1,t)}\right)$ using the changes in observable characteristics between time $t-1$ and t. We consider different polynomial specifications for f. It shows that most of estimated coefficients on $\mathcal{N}_{t-1}$ are significant at 1 percent level and robust across the different polynomial speficiation. Most results respond to our expectation. Two things are worth explaining. First, the estimates are slightly larger than one. It is not saying that our accounting identity is wrong. Rather it is largely because we use in sample observations for $\mathcal{N}_{t-1}$, which is smaller than actual observations at time $t-1$, due to exits of hotels before we are able to collect the data. Second, the estimates on $\mathcal{N}_{2001,L}$ is relatively larger than others. It is largely due to the fact that there is a much small number of low-type hotels in our data in 2001, resulting in a even inflated coefficient.

The last two things is to notice. First, until now, we discuss why the old equilibrium outcome $\mathcal{N}_{t-1}$ can be a valid instrument for observed $\mathcal{N}_t$. We haven't explained how $\mathcal{N}_{t-1}$ can be used for the simulated equilibrium $\hat{\mathcal{N}}_t$, which we will use in estimation. We will postpone it until the moment construction. Second, since each hotel group has its own old equilibrium outcome. Back to our equation, the old equilibrium of competing products can also be an instruments except the distance to the headquarter.

8.4 Moment Conditions and Objective Functions

The model does not yield the closed-form solution to their equilibrium. We need to use the simulated method. Although the maximum simulated likelihood (SML) and the method of simulated moments are candidates for our estimation, we choose the method of simulated moments in our estimation. In particular, it is because as the number of groups increases, the dimension of the market configuration increases. As a result, the joint probability of each market configuration might become less informative in SML.We choose SMM over SML. As the number of groups increases, the more moments can be created, additional variations added from the additional group will increases the precision of the estimates. We will proceed to construct the moment conditions used in this paper.

We first construct moment conditions for the coefficients of competing products and observable characteristics. Following Berry (1992) and Jia (2008), we define the prediction error as the difference between the observed equilibrium number of hotels in each market and the (conditional) expectation of the equilibrium number: In particular, we define the observed equilibrium market outcome $\mathcal{N}_m = (\mathcal{N}_{m,H}, \mathcal{N}_{m,L})$ and the true parameter $\boldsymbol{\theta}^o = (\beta_H, \beta_L, \boldsymbol{\gamma}_H^o, \boldsymbol{\gamma}_L^o, \boldsymbol{\gamma}_H^c, \boldsymbol{\gamma}_L^c)$

$$\boldsymbol{\nu}_{m,H}(\mathcal{N}_m, W_m; \boldsymbol{\theta}^o) = \mathcal{N}_{m,H}^* - \mathbb{E}[\mathcal{N}_{m,H}|W_m; \boldsymbol{\theta}^o]$$
$$\boldsymbol{\nu}_{m,L}(\mathcal{N}_m, W_m; \boldsymbol{\theta}^o) = \mathcal{N}_{m,L}^* - \mathbb{E}[\mathcal{N}_{m,L}|W_m; \boldsymbol{\theta}^o]$$

where $W_m = (X_m, \boldsymbol{Z}_m)$ with $\boldsymbol{Z}_m = \left(Z_{m1}^H, ...Z_{mG}^H, Z_{m1}^L, ...Z_{mG}^L\right)$. By definition, it follows that the prediction error vector $\boldsymbol{\nu}_m = (\boldsymbol{\nu}_{m,H}, \boldsymbol{\nu}_{m,L})$ is mean independent of X_m. In addition, define $\hat{\mathcal{N}}_{m,H}$ and $\hat{\mathcal{N}}_{m,L}$ as the simulution estimator for $\mathcal{N}_{m,H}$ and $\mathcal{N}_{m,H}$ for some fixed set of random draws $\hat{\mu}$ with

$$\hat{\mathcal{N}}_{m,H}(W_m, \hat{\mu}; \boldsymbol{\theta}^o) = \mathbb{E}\left[\hat{\mathcal{N}}_{m,H}|W_m; \boldsymbol{\theta}^o\right] + \eta_{m,H}$$
$$\hat{\mathcal{N}}_{m,L}(W_m, \hat{\mu}; \boldsymbol{\theta}^o) = \mathbb{E}\left[\hat{\mathcal{N}}_{m,L}|W_m; \boldsymbol{\theta}^o\right] + \eta_{m,L}$$

where we also have the simulation error $\boldsymbol{\eta}_m = \left(\boldsymbol{\eta}_{m,H}, \boldsymbol{\eta}_{m,L}\right)$ is mean independent of X_m. Now by definition, we have

$$\begin{aligned} \mathcal{N}_{m,H} &= \hat{\mathcal{N}}_{m,H}(W_m, \hat{\mu}; \boldsymbol{\theta}^o) + \boldsymbol{\nu}_{m,H} - \eta_{m,H}; \\ \mathcal{N}_{m,L} &= \hat{\mathcal{N}}_{m,L}(W_m, \hat{\mu}; \boldsymbol{\theta}^o) + \boldsymbol{\nu}_{m,L} - \eta_{m,L}; \end{aligned} \tag{4}$$

Write $\tilde{\boldsymbol{\nu}}_m = \boldsymbol{\nu}_m - \boldsymbol{\eta}_m$ and by definition, it satisfies that

$$\mathbb{E}[\tilde{\boldsymbol{\nu}}_m|W_m] = \mathbb{E}\left[\left(\begin{array}{c}\tilde{\boldsymbol{\nu}}_{m,H}\\ \tilde{\boldsymbol{\nu}}_{m,L}\end{array}\right)\middle| W_m\right] = \mathbf{0}_{2\times G}$$

We now construct the moment condition for own products. Recall that

$$\begin{aligned} \mathcal{N}_{m,H,t} &= \mathcal{N}_{m,H,t-1} + f_H\left(\mathcal{I}_{m,(t-1,t)}\right) + \epsilon_{m,H,t} \\ \mathcal{N}_{m,L,t} &= \mathcal{N}_{m,L,t-1} + f_L\left(\mathcal{I}_{m,(t-1,t)}\right) + \epsilon_{m,L,t} \end{aligned}$$

by construction, $\mathbb{E}[\epsilon_{m,H,t}|\mathcal{N}_{m,H,t-1}] = 0$ and $\mathbb{E}[\epsilon_{m,L,t}|\mathcal{N}_{m,L,t-1}] = 0$. Together with another identity presented above (4), then it follows that

$$\hat{\mathcal{N}}_{m,H}(W_m, \hat{\mu}; \boldsymbol{\theta}^o) = \mathcal{N}_{m,H,t-1} + f_H\left(\mathcal{I}_{m,(t-1,t)}\right) + \epsilon_{m,H,t} - \tilde{\boldsymbol{\nu}}_{m,H} \tag{5}$$
$$\hat{\mathcal{N}}_{m,L}(W_m, \hat{\mu}; \boldsymbol{\theta}^o) = \mathcal{N}_{m,L,t-1} + f_L\left(\mathcal{I}_{m,(t-1,t)}\right) + \epsilon_{m,L,t} - \tilde{\boldsymbol{\nu}}_{m,L} \tag{6}$$

by construction again

$$\begin{aligned} \mathbb{E}\left[\epsilon_{m,H,t} - \boldsymbol{\nu}_{m,H} + \eta_{m,H}|\mathcal{N}_{m,H,t-1}\right] &= 0; \\ \mathbb{E}\left[\epsilon_{m,L,t} - \boldsymbol{\nu}_{m,L} + \eta_{m,L}|\mathcal{N}_{m,L,t-1}\right] &= 0. \end{aligned}$$

where $\epsilon_{m,H,t} - \boldsymbol{\nu}_{m,H} + \eta_{m,H}$ can be estimated by usin the value of the predicted value $\hat{\mathcal{N}}_{m,H}(W_m, \hat{\mu}; \boldsymbol{\theta}^o)$ minus the predicted value on $\mathcal{N}_{m,H,t-1} + f_H\left(\mathcal{I}_{m,(t-1,t)}\right)$. Similar to $\epsilon_{m,L,t} - \boldsymbol{\nu}_{m,L} + \eta_{m,L}$. Given the equation above, naturally $\mathcal{N}_{m,t-1}$ is correlated with $\hat{\mathcal{N}}_m(W_m, \hat{\mu}; \boldsymbol{\theta}^o)$ and $\mathcal{N}_{m,t-1}$ is independent of structual market unobserved heterogeneity. Thus, $(\mathcal{N}_{m,H,t-1}, \mathcal{N}_{m,L,t-1})$ play a role of instrument for $\left(\hat{\mathcal{N}}_{m,H}(W_m, \hat{\mu}; \boldsymbol{\theta}^o), \hat{\mathcal{N}}_{m,L}(W_m, \hat{\mu}; \boldsymbol{\theta}^o)\right)$ and identify the own competition effect.

Combining two sets of moments, given the conditional expectation condition, we can construct our objective function. Define the simulated estimator as $\hat{\mathcal{N}}_m(W_m, \hat{\mu}; \boldsymbol{\theta}) = \frac{1}{R}\sum_{r=1}^{R} \hat{\mathcal{N}}_m^r(W_m, \hat{\mu}^r; \boldsymbol{\theta})$. Then, we have

$$\widehat{\vartheta_m} = \frac{\widehat{\tilde{\boldsymbol{\nu}}_m}}{\widehat{\epsilon_m - \tilde{\boldsymbol{\nu}}_m}} \quad \text{where} \quad \begin{array}{c} \widehat{\tilde{\boldsymbol{\nu}}_m} = \mathcal{N}_m - \hat{\mathcal{N}}_m(W_m, \hat{\mu}; \boldsymbol{\theta}) \\ \widehat{\epsilon_m - \tilde{\boldsymbol{\nu}}_m} = \hat{\mathcal{N}}_m(W_m, \hat{\mu}; \boldsymbol{\theta}) - \hat{\mathcal{N}}_{m,RF} \end{array}$$

where $\hat{\mathcal{N}}_{m,RF}$ is reduced form fitted value on $\mathcal{N}_m$ from accounting identity.Define $q_m\left(\tilde{W}_m, \boldsymbol{\theta}\right) = \frac{1}{M}\sum_{m=1}^{M} \widehat{\vartheta_m} \otimes \tilde{W}_m$ where $\tilde{W}_m$ is an augmented vector of instruments including the old equilibrium $\mathcal{N}_{t-1}$ with $\mathbb{E}[q_m(\boldsymbol{\theta}_0)] = 0$. The simulated sample objective function can be writen as

$$\begin{aligned} Q_M &= q_m' W^{-1} q_m = \left(\frac{1}{M}\sum_{m=1}^{M} \widehat{\vartheta_m} \otimes \tilde{W}_m\right)^T \Omega_m^{-1} \left(\frac{1}{M}\sum_{m=1}^{M} \widehat{\vartheta_m} \otimes \tilde{W}_m\right) \\ \text{and } \hat{\boldsymbol{\theta}} &= \arg\min_{\theta} Q_M \end{aligned}$$

where $\Omega_m = \mathbb{E}\left[q_m\left(\tilde{W}_m \boldsymbol{\theta}\right) q_m\left(\tilde{W}_m \boldsymbol{\theta}\right)'\right]$ for some preliminary estimator, $\boldsymbol{\theta}^{(1)}$. While it is straightforward to implement the standard two-step GMM estimator, we notice that

$$cov\left(q_{g,m}, q_{g',m}\right) \neq 0$$

if $g = g'$. The difficulty to obtain the consistent estimator in the standard two-step method is that the moments in our model setup can be correlated due to common market-group level shock. As discussed in Newey (2007), the correlation between the moments and their derivatives can be an important source of bias in two step optimal GMM estimator. Such bias can be removed by the continously updated estimator. So in our estimation algorithm, we will use the continously updated estimator.

The estimation procedure is as follows

Step 1: Guess the initial parameters and draw independently from the normal distribution the vector of idiosyncratic shocks: the market level errors $\{\varsigma_m\}_{m=1}^M$; the common market-group level vector $\{\varsigma_{i,m}\}_{m=1}^M$ and the market-group-type level vector $\{\varsigma_{i,m}^H\}_{m=1}^M$ and $\{\varsigma_{i,m}^L\}_{m=1}^M$ where $g = 1, \ldots, 6$.

Step 2: Obtain the simulated profit functions for each type $\left\{\Pi_{im,H}^*, \Pi_{im,H}^*\right\}_{m=1}^M$ for $g = 1, \ldots, 6$.

Step 3: Repeat Steps 1 and 2 R times and obtain $q_m\left(\tilde{W}_m, \boldsymbol{\theta}\right)$. Search for parameter values that minimize the objective function (4) using the identical matrix as weighting matrix, while using the same set of simulation draws for all values of $\boldsymbol{\theta}$.

Step 4: Presume the estimate using the identical matrix is $\boldsymbol{\theta}^{((1)}$, we can construct the estimator for the weighting matrix as

$$\hat{\Omega}_m^{(1)} = \frac{1}{m} \sum\nolimits_{m=1}^{M} q_m\left(\tilde{W}_m, \boldsymbol{\theta}\right) q_m\left(\tilde{W}_m, \boldsymbol{\theta}\right)'.$$

Step 5: Using the estimated weighting function $\hat{\Omega}_m^{(1)}$, repeat Steps 1-4 and obtain an estimate $\boldsymbol{\theta}^{(2)}$ and update $\hat{\Omega}_m^{(2)}$. Keep updating until the parameter $\boldsymbol{\theta}$ converges.

Following Jia (2008), we implement the Halton draws to get better convergence properties and smaller simulation variances. In our setting, we have 30 parameters to estimate. We construct the following set of moments: (a) the number of high-type hotels for each group (b) the number of low-type hotels together with seven control variables (including the constant). This creates sixty moment conditions in total.

8.5 Discussion

Discussion: Before proceeding to the estimation results of the model, we will discuss five simplifying model restrictions, which might potentially affect our results.

First, we do not consider a full dynamic entry model which is more flexible to capture the nature of competition due to lack of the exit data. Instead, we pursue a static and simultaneous game and interpret it as a single period of longer dynamic game. This assumption is like existing studies Seim (200?) and McDevitt and Roberts (2014).

Second, we assume complete information rather than incomplete information (e.g. Seim (2006) and McDevitt and Robert (2014)) for two reasons. First, due to the salient feature of lodging industry, each hotel may get access to the hotel online network system to know the competitor vacancy, capacity and price at any period. In this sense, the hotel group knows the full information of their competitor. It might be not realistic to assume the incomplete information game. Second, incomplete information is commonly assumed to reduce the endogeneity problem in estimation when lack of the proper instrument. In our paper, we have policy shock that can help to solve the endogeneity problem. It is not essential for us to assume incomplete information in order to simplify the estimation.

Third, the presence of multiple equilibrium may also complicate the estimation. As we discuss before, many equilibria can obtain in our setting largely due to the assumption of symmetric players. Several strategies have been used to guarantee the uniqueness of equilibria, like sequential assumption (Zhu, Singh and Manuszak (2009)) or arbitrarily equilibrium selection. We adopt a strategy in the middle, as we don't have any information on the order of entry, we assume to all the multiple equilibrium and simply average over all the equilibria. This simplification will substantially reduce the computation burden in the estimation.

Fourth, even though in this paper, we extend homogenous product differentiation to the heterogeneous product differentiation. But still we cannot allow for the full heterogeneity in competition parameters indexed by the group identity, because as the number of the group increases, the number of competition parameters proportional increases, which is commonly known as a curse of dimensionality issues.

Fifth, as one of key identification strategy, we use the within sample variation on the capacity to construct the instruments. This may raise the concern that there might exist some network effect of capacity among the cities within the same province. We argue that the lodging industry market is quite localized because it is less likely that travelers or business usage to move from one city to another. It might be possible that the unobserved local market profitability might correlated with the market profitability in the nearby cities. We argue that as long as the correlation in the prefecture-level cities exponentially declines with the distance after we control local economic statistics, our current construction on the instruments is valid.

Finally, one may concern about many moment bias, which might occur in our estimation as we have 242 moment restrictions with the sample size 342. We emphasize that though we have many moment restrictions, it is not the standard many moment issues. Note that many moment issues are caused by a large number of instruments for single equation (single errors) where the number of instruments is much larger than the number of endogenous variables (e.g. one endogenous variable, ?? instruments, Newey (2001)). In our setting, for each equation, we have 4 endogenous variables and we have 14 instruments (In fact, we can average these instruments and have the exact identification case). Many moment restrictions in our setting is due to an increasing number of error terms resulting from an increasing number of groups, which amplifies the number of moments. But by doing so, we may lose some important variation. Therefore, we adopt the non-averaging case.

Besides the modeling assumption, we also impose two additional assumptions on data construction, which also deserve discussion. First, we treat the districts in a direct administrative city as a market, since government administration at the district level is the same as the prefecture level city; and it might be not comparable in the agglomeration pattern if we treat a direct administrative city the same as the prefecture level cities.

Second, we exclude the observations in Beijing and Shanghai, largely because five hotel headquarters are located in Beijing and Shanghai. The resulting distance to headquarter for the hotels in Beijing and Shanghai, the key measure to construct the instrument, is much smaller than the distance we observed in other cities. These observations may not represent other observation, potentially leads outlier issues and render the weak instrument issue of the distance to headquarter per capacity in the reduced-form estimation and instable estimation in structural model. On the other hand, we are aware that by dropping the observations in Beijing and Shanghai, we may have loss of variation in the sample. By balancing the tradeoff between the sample size and the sample representative, we argue that dropping the observations in Beijing and Shanghai will help us to obtain the reliable estimates.

9 Structural Estimation Results

Table 10 displays SMM estimates of the coefficients on observed characteristics and competition effects. Most of the coefficients on the observable characteristics in the potential profit function are significant across the different specifications. On the whole, the results on observable characteristics correspond closely to our expectation.

One of the main questions that motivated this paper is whether within-group competition effect is smaller than between-group competition, which may result in the within-group agglomeration. We present two sets of results: one set of results are based on the Berry type identification without using the year 2001 stock as an instrument; the other set of results are based on identification strategy proposed in this paper using the year 2001 stock as an instrument.

Table 10 summarizes the results. Columns (1)-(2) show the first set of results and Columns (3)-(4) show the second set of results. It shows that with using the year 2001 stock, the parameters associated with competition effects are not identifiable including the domestic competitor coefficient and all other competition coefficients and the magnitude is much smaller than that in Columns (3) and (4).

The results in Columns (3) and (4) show that the same type competition is larger than the cross type competition (2) between groups, the same type competition is also larger than the cross-type competition (3) comparing the magnitude of within-group competition and between-group competition, the within-group competition effects are slightly larger than between-group competition for the same/different type(s).

The table suggests two key conclusions. The first is within group agglomeration is driven by the selection on market profitability not the mitigated competition within the firm. We estimate that within group competition for L-type from L-type competitor is 5.13, this effect is almost the same as between-group competition 5.16. In addition, the within group competition for H-type from H-type competitor is 4.63 and we notice that the within-group competition for H-type from L type is 4.92, which is slightly larger than HH type which suggests some business stealing from the low-type chains from H-type chains.

The second conclusion is that taking into account of selection on (observed and) unobserved heterogeneity is critical to obtaining accurate results. Failing to take into account of the correlated unobserved heterogeneity would lead us to conclude that hotel groups' entry are cooperative rather than competitive

10 Counterfactual Analysis

[To Be Complete]

11 Conclusion

This paper documents within and between group product differentiated agglomerations of hotel chains alliance in China, especially the within group agglomeration. We first propose a theoretical model of possible channels that might facilitate within-group and between-group competition to explain such agglomeration. Using a unique data of hotel entries at China for international hotel groups, we exploit the exogenous variation of lodging industry market deregulation and identify the within-group and between group competition effects. We find that the within group competition for the same type or across type are lsightly larger than those of between group competition after taking into account of the unobserved market profitability. We will use our analysis to evaluate the impact of the recent merger of Marriott and Starwood in 2016. Our analysis will provide theoretical and empirical framework to evaluate the competition effects when the presence of product differentiation of group chains and also shed the light on evaluating the effect of merger.

References

[1] A, Ren A. Ta Kosov and Francine Lafontaine (2012). "Much ado about chains: A research agenda," International Journal of Industrial Organization, 30(3), 303–308.

[2] Baum, Joel AC and Heather A. Haveman(1997). "Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990,"Administrative Science Quarterly, pp. 304–338.

[3] Bresnahan, Timothy F and Peter C Reiss (1991). "Entry and competition in concentrated markets,"Journal of Political Economy, 99(5), 977–1009.

[4] Canina, Linda, Cathy A. Enz, and Jeffrey S. Harrison (2005). "Agglomeration efects and strategic orientations: Evidence from the US lodging industry,"Academy of management journal, 48(4), 565–581.

[5] Chung, Wilbur and Arturs Kalnins (2001). "Agglomeration effects and performance: A test of the Texas lodging industry,"Strategic Management Journal, 22(10), 969–988.

[6] Ellickson, Paul B and Sanjog Misra (2008). "Supermarket pricing strategies,"Marketing science, 27(5), 811–828.

[7] Fischer, Jeffrey H. and Joseph E. Harrington Jr (1996), "Product variety and firm agglomeration,"The RAND Journal of Economics, pp. 281–309.

[8] Gan, Li and Manuel A. Hernandez (2013), "Making friends with your neighbors? Agglomeration and tacit collusion in the lodging industry,"Review of Economics and Statistics, 95(3), 1002–1017.

[9] Graitson, Dominique (1982) "Spatial competition in Hotelling: A selective survey," Journal of Industrial Economics, pp. 11–25

[10] Gu, Zheng (2003) "The Chinese lodging industry: problems and solutions,"International Journal of Contemporary Hospitality Management, 15(7), 386–392.

[11] Helsley, Robert W. and William C. Strange (1990)"Matching and agglomeration economies in a system of cities," Regional Science and Urban Economics, 20(2), 189–212.

[12] Hollenbeck, Brett (2014)"The spread of horizontal chains: Efficiency or market power," Working Paper in University of California, Berkeley.

[13] Hotelling, Harold (1929) "Stability in competition,"The Economic Journal, 39(153),41–57.

[14] Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, Jean Tirole, and Others, “The economics of tacit collusion,”Final report for DG competition, European Commission

[15] Jia, Panle (2008). “What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry,”Econometrica, 76(6), 1263–1316.

[16] Lafontaine, Francine and Kathryn L Shaw (2005) “Targeting managerial control: evidence from franchising,” RAND Journal of Economics, pp. 131–150

[17] Marco-Lajara, Bartolom E., Patrocinio Del Carmen Zaragoza-S A Ez, Enrique Claver-Cort E S, and Mercedes U Beda-Garc I A (2016)“Knowledge sources, agglomeration, and hotel performance,” Journal of Business Research, 69(11), 4856–4861.

[18] Mazzeo, Michael J.(2002a) “Competitive Outcomes in Product-Differentiated Oligopoly,” The Review of Economics and Statistics, 84(4), 716–728.
-(2002b), “Product choice and oligopoly market structure,”RAND Journal of Economics, pp. 221–242.
-(2004) “Retail contracting and organizational form: Alternatives to chain affiliation in the
motel industry” Journal of Economics & Management Strategy, 13(4), 599–615

A Appendix I: Table

Table 1: Hotel Group Description: Market Share

Hotel	**High End**	**Low End**	**Total**	**Percentage**
IHG	133	124	257	23.49%
Wyndham	74	142	216	19.74%
Starwood	130	41	171	15.63%
Accor	63	52	115	10.51%
Hilton	85	29	114	10.42%
Marriot	81	18	99	9.05%
Big 6	566	406	972	88.85%
Hyatt	49	0	49	4.48%
Ascott	0	35	35	3.20%
Peninsula	3	0	3	0.27%
Best Western	0	14	14	1.28%
Okura Nikko	11	0	11	1.01%
Carlson	8	2	10	0.91%
Total	637	457	1094	100.00%

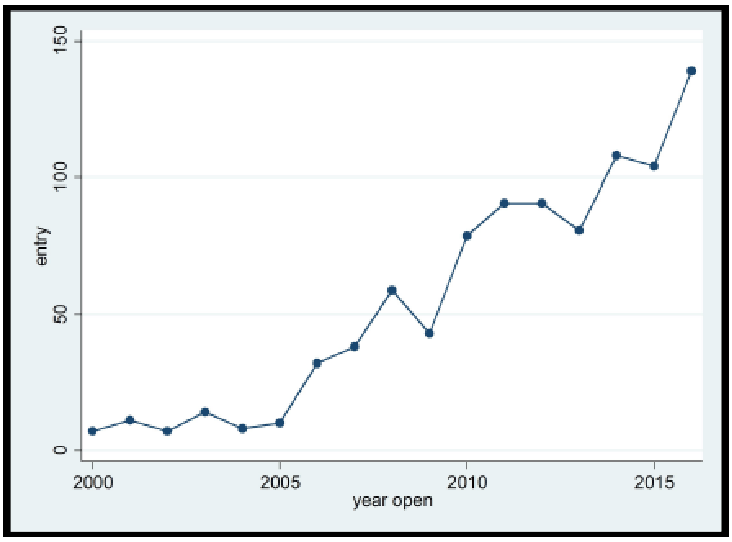

Figure 1: Figure 1: Entry by Year

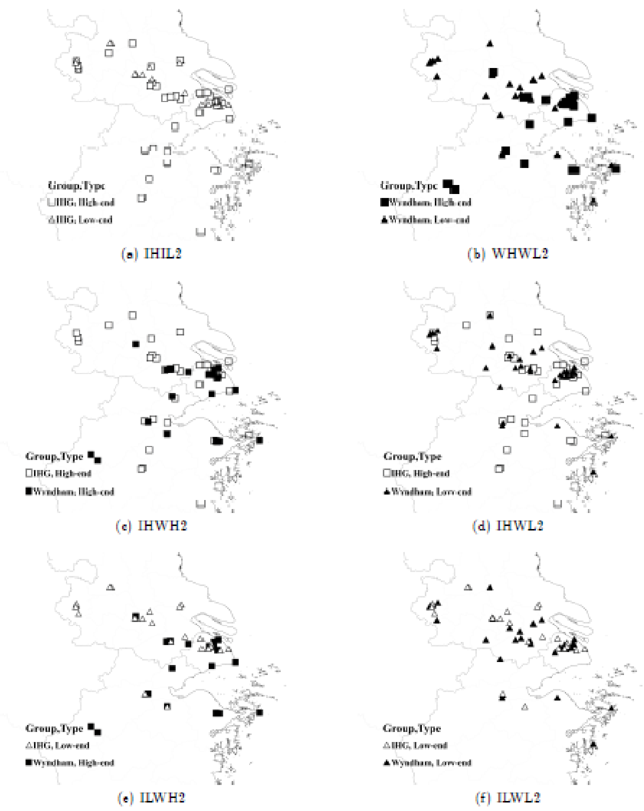

Figure 2: Figure 2: Market Structure

Table 2: Rating and Hotel Categorization

			2	2.5	3	3.5	4	4.5	5	Total	Average Score
IHG	Regent	High						0	1	1	5.00
	InterContinental							9	27	36	4.88
	Indigo							2	3	5	4.80
	Crowne Plaza							37	35	72	4.74
	HuaLuxe							2	2	4	4.75
	Holiday Inn	Low		0	0	14	30	19	8	71	4.15
	Holiday Inn Express			14	14	28	7	2	1	66	3.29
	Total			14	14	42	37	21	9	137	
Wyndham	Wyndham Grand Collection	High					0	3	9	12	4.88
	Wyndham						0	7	12	19	4.82
	Howard Johnson						5	24	19	48	4.65
	Ramada	Low		0	1	15	23	27	9	75	4.19
	Days Inn			1	2	33	16	8	4	64	3.81
	Total			1	3	48	39	35	13	139	
Starwood	St Regis	High						0	6	6	5.00
	Tribute Portfolio							0	1	1	5.00
	Westin							8	12	20	4.80
	Le Meridien							4	5	9	4.78
	Sheraton							33	35	68	4.76
	W Hotel							1	1	2	4.75
	The Luxury Collection							5	2	7	4.64
	Element	Low				0	0	1		1	4.50
	Four Points					3	4	17		24	4.29
	Aloft					5	4	3		12	3.92
	Total					8	8	72	62	144	
Accor	Raffles	High					0	0	1	1	5.00
	Sebel						0	0	1	1	5.00
	Fairmont						0	1	3	4	4.88
	Sofitel						1	3	15	19	4.87
	Swissotel						0	1	3	4	4.88
	Pullman						0	15	14	29	4.74
	MERCUREE	Low	0	1	0	0	2	5	2	10	4.30
	Novotel		0	0	0	2	9	5	1	17	4.15
	Mercure		1	0	0	5	5	2	1	14	3.82
	Ibis Styles		0	0	4	1	2	0	0	7	3.36
	ibis		1	32	18	0	1	0	0	52	2.69
	Total		2	33	22	8	20	32	41		
Hilton	Waldorf Astoria					0	0	0	2	2	5.00
	Conrad					0	0	1	3	4	4.88
	Hilton					1	1	22	20	44	4.69
	Doubletree by Hilton					0	1	16	7	24	4.63
	Hampton by Hilton					0	4	2	1	7	4.29
	Hilton Garden Inn					3	3	3	0	9	4.00
	Total					4	9	44	33	90	
Marriott	Ritz Carlton	High						0	9	9	5.00
	Edition							0	1	1	5.00
	JW Marriott							2	9	11	4.91
	Renaissance							4	14	18	4.89
	Executive Appartment							2	4	6	4.83
	Marriott							9	16	25	4.82
	Courtyard	Low					9	5		14	4.18
	Total					9	22	53			

Table 3: Summary Statistics on Controls

Variable	Mean	Std. Dev.	Min	Max
East	0.51	0.50	0.00	1.00
Airport (1 if Airport exists in prefecture city)	0.51	0.50	0.00	1.00
Distance to closest tourist site (1000 meters)	25.01	17.66	0.04	50.00
Foreign tourists (in 2014,million)	0.47	1.22	0.00	5.73
Urban ratio	0.54	0.21	0.07	1.00
Average salary (in 2014,thousand)	56.02	16.18	24.63	104.47
Population (in 2014,million)	1.18	1.00	0.15	6.37
Low-end independent hotels per area	0.07	0.28	0.00	3.25
High-end independent hotels per area	0.03	0.14	0.00	1.50

Table 4: Market Configuration (Hotel-Market Level)

	OLL	OHH	OLH	CLL	CHH	CLH	CHL
IHG	21	23	29	14	13	9	16
Wyndham	27	14	25	12	5	14	10
Startwood	6	23	15	5	8	4	12
Accor	21	6	16	10	12	7	10
Hilton	1	12	8	2	5	1	8
Marriott	2	8	4	0	5	0	4

Table 5: Market Configuration (Market Level)

	# of Market	OHH	OLL	OHL	CHH	CLL	CHL	CLH
>20	9	9	8	9	9	9	9	9
20–11	12	12	12	12	12	12	12	12
10–5	12	8	9	11	12	11	12	12
5–2	22	6	7	7	13	4	17	17

Table 6: Correlation Pattern Within the Group

IHG correlation matrix			Accor correlation matrix		
	2008	2016		2008	2016
HHOwn	-0.04	0.11	HHOwn	-0.14	-0.05
LLOwn	-0.07	0.63	LLOwn	-0.01	0.18
HLOwn	0.07	0.14	HLOwn	0.02	-0.02
LHOwn	0.05	0.64	LHOwn	0.18	0.02
Wyndham correlation matrix			**Hilton correlation matrix**		
	2008	2016		2008	2016
HHOwn	0.00	-0.14	HHOwn	-0.02	0.03
LLOwn	0.01	0.11	LLOwn	0.00	-0.03
HLOwn	0.11	-0.10	HLOwn	0.00	0.05
LHOwn	0.00	0.10	LH_Own	0.01	-0.04
Starwood correlation matrix			**Marriott correlation matrix**		
	2008	2016		2008	2016
HHOwn	-0.07	0.01	HHOwn	-0.04	0.43
LLOwn	0.00	-0.01	LLOwn	0.00	0.13
HLOwn	0.00	-0.01	HLOwn	0.00	0.09
LHOwn	0.00	0.03	LHOwn	0.03	0.06

Table 7: Top 20 Cities with Largest Number of Hotels of Six Hotel Groups

City Name	Total # of Hotels	Low-type	High-Type	Rank
Chengdu	46	26	20	1
Suzhou	34	21	13	2
Guangzhou	33	14	19	3
Sanya	31	4	27	4
Chuzhou	30	14	16	5
Xi'an	24	12	12	6
Shenzhen	23	8	15	7
Jinan	23	9	14	8
Nanjing	22	9	13	9
Taizhou	18	3	15	10
Wuhan	17	9	8	11
Hefei	17	10	7	12
Xiamen	15	5	10	13
Fuzhou	15	10	5	14
Wuxi	15	5	10	15
Zhengzhou	13	7	6	16
Fushan	13	6	7	17
Changzhou	12	7	5	18
Shenyang	12	6	6	19
Haikou	11	4	7	20

Table 8: Top 20 Cities with Largest Number of Hotel for Single Group

City Name	Total #	IHG	Wyndham	Starwood	Accor	Hilton	Marriott	Max #	Hotel Group Name	Market Share	Rank
Chengdu	46	20	8	5	5	4	4	20	IHG	0.43	1
Fuzhou	15	2	10	2	0	1	0	10	Wyndham	0.67	2
Suzhou	34	8	10	3	7	1	5	10	Wyndham	0.29	3
Nanjing	22	6	2	3	8	2	1	8	Accor	0.36	4
Jinan	23	5	4	8	4	2	0	8	Startwood	0.35	5
Xi'an	24	4	8	2	6	3	1	8	Wyndham	0.33	6
Sanya	31	5	5	8	4	4	5	8	Starwood	0.26	7
Guangzhou	33	8	7	7	2	6	3	8	IHG	0.24	8
Chuzhou	30	7	6	6	5	2	4	7	IHG	0.23	9
Hefei	17	6	3	3	3	2	0	6	IHG	0.35	10
Nanning	10	4	5	0	1	0	0	5	Wyndham	0.50	11
Shenyang	12	5	2	1	3	1	0	5	IHG	0.42	12
Changzhou	12	3	5	2	0	1	1	5	Wyndham	0.42	13
Shenzhen	23	5	3	5	2	3	5	5	IHG/St./Mar.	0.22	14
Taizhou	18	3	5	2	3	4	1	5	Wyndham	0.28	15
Wuxi	15	3	5	3	3	1	0	5	Wyndham	0.33	16
Wuhan	17	2	5	2	4	3	1	5	Wyndham	0.29	17
Nantong	8	4	2	0	2	0	0	4	IHG	0.50	18
SuiZhou	11	2	4	3	0	0	2	4	Wyndham	0.36	19
Fushan	13	3	4	2	1	2	1	4	Wyndham	0.31	20

Table 9: Reduced Form Estimates

Y_L=(Entry_L>0)	EST.	SE.	Y_H = I(Entry_H>0)	EST.	SE.
Own_HType	1.2041	0.1008	Own_LType	1.1892	0.0882
Com_LType	-0.0732	0.0618	Com_HType	0.1668	0.0585
Com_HType	0.1460	0.0624	Com_LType	0.1178	0.0577
DistLType	-1.0281	0.2008	DistHType	-0.3540	0.0577
Density	-0.0069	0.0120	Density	0.0074	0.0119
Income	0.0100	0.0108	Income	0.0200	0.0104
Airport	0.2068	0.1035	Airport	0.2357	0.1034
Foreign Tourist	-0.0543	0.0428	Foreign Tourist	-0.0693	0.0467
Domestic Tourist	0.0001	0.0008	Domestic Tourist	0.0014	0.0008
Capital	0.5367	0.1621	Capital	-0.0227	0.1676
Domestic Competitor Htype	0.0028	0.0048	Domestic Competitor Htype	-0.0048	0.0045
Domestic Competitor Ltype	-0.0012	0.0014	Domestic Competitor Ltype	0.0043	0.0015
Constant	-0.8222	0.3791	Constant	-2.0486	0.3701
Observation		1710	Log Likelihood		-760.4

Table 10: Structural Model Estimates with the Year 2001 Instruments

L-Type Hotel	Est.	SE.	High-Type Hotel	Esti.	SE.
Own_Ltype	-5.13	0.97	Own_Htype	-4.63	0.98
Own_HType	-3.76	0.98	Own_Ltype	-4.93	0.99
Com_HType	-5.16	0.99	Com_Htype	-3.11	0.99
Com_LType	-2.35	0.99	Com_Ltype	-2.35	1.00
DistLType	-0.61	1.00	DistHType	-1.89	0.99
Density	7.41	0.86	Density	5.33	0.93
Income	5.44	1.00	Income	3.15	0.90
Airport	7.82	0.93	Airport	6.62	0.98
Foreign Tourist	5.08	0.69	Foreign Tourist	3.89	0.52
Domestic Tourist	4.06	1.00	Domestic Tourist	5.69	1.00
Capital	6.21	0.04	Capital	9.01	0.05
Domestic Competitor Htype	-1.48	0.99	Domestic Competitor Htype	-1.57	0.99
Domestic Competitor Ltype	-3.48	1.00	Domestic Competitor Ltype	-0.99	1.00
Constant	4.37	0.8052	Constant	-2.06	0.99
			Rho	0.31	0.99
			Observations		296

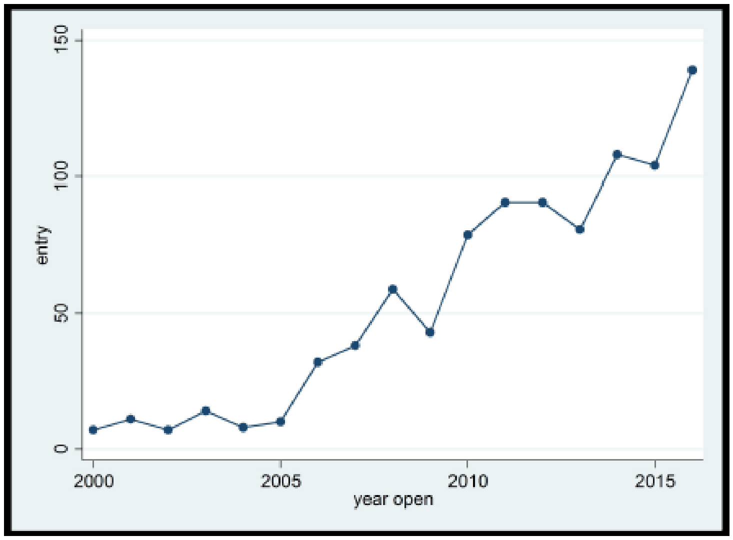

Figure 1: Entry by Year

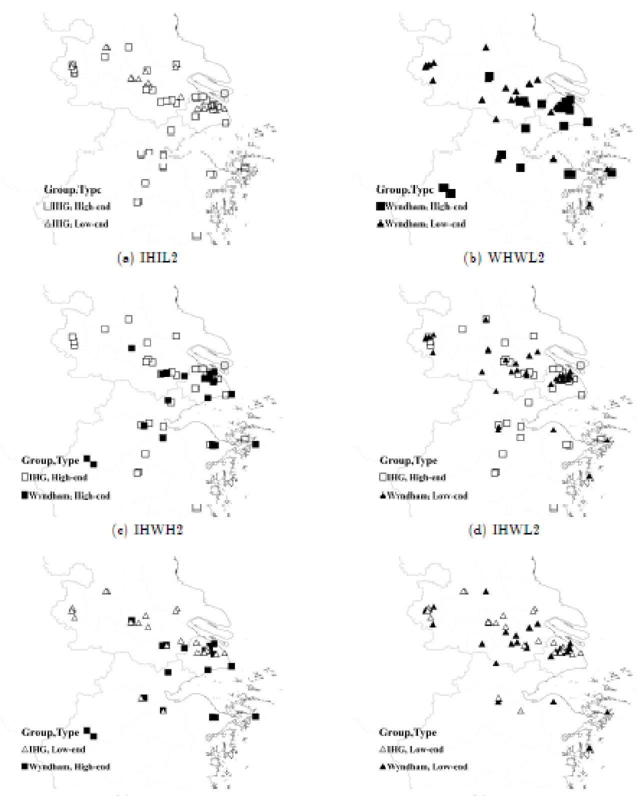

Figure 2: Market Structure