Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry*

Preliminary and incomplete.

Please do not circulate without the author's consent.

Maria Ana Vitorino

Graduate School of Business
University of Chicago

July 16, 2007

Abstract

This paper studies the joint entry decisions of stores in a particular form of retail cluster - the regional shopping center. I propose a strategic model of entry capable of quantifying the magnitude of inter-store spillovers and show how these effects can help in explaining the composition of a given market. The model is applied to a novel dataset containing information about the store configurations of all US regional shopping centers and estimated using a new econometric approach that is robust to multiple equilibria. I find evidence of significant spillovers across stores with sign and magnitude varying across store-types.

Keywords: Entry, Spillovers, Shopping Centers, Incomplete information game, Direct optimization approach
JEL Classification: C6, L0, L1, L2, L8, M3

*I thank my advisors Jean-Pierre Dubé, Pradeep Chintagunta, Kenneth Judd, Peter Rossi and Matthew Gentzkow for their constant support and guidance. I have also benefited from helpful discussions with Frederico Belo, Hongju Liu and Jeremy Fox. I also thank Che-Lin Su, Todd Plantenga, David M. Gay and Robert Fourer for providing support and licensing for the software AMPL and KNITRO and to Marcus Edvall for support with Tomlab. I would also like to thank Julie Cameron and Robert Michaels from General Growth Properties and Mike Tubridy from the International Council for Shopping Centers for sharing with me their insights on the Shopping Center Industry and to Bob Galvin from the National Research Bureau for making available part of the data used in this paper. I gratefully acknowledge the financial support from Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology), from the Calouste Gulbenkian Foundation and from the Kilts Center for Marketing at the University of Chicago. All errors are my own.

1 Introduction

A common characteristic of the United States retail market structure is the notion of a shopping hub. A shopping hub consists of a large cluster of retail stores in close geographic proximity to one-another. A surprising aspect of these geographic clusters of stores is the frequent presence of several highly-substitutible competitors. For instance, antique and jewelry stores are often found in close proximity to one another. The same phenomenon arises for automobile dealerships, often collectively termed as "auto malls". Firms locate near one another since doing so lowers consumers search costs and increases aggregate demand for these firms. Such positive spillovers offset the main cost from joining a cluster, which is that close proximity of firms selling similar goods intensifies competition.

In this paper, I study the determinants of market structure for a specific type of retail cluster – the regional shopping center. Shopping Centers have been the strongest area in real estate in recent years. Despite the importance of the industry, it has been largely overlooked by the Marketing Literature. At the same time, while there have been some (mostly) descriptive studies of shopping centers in the Real Estate literature, there as been little attempt to explain (structurally) the observed distribution of store brands in a shopping center.

The shopping center industry seems well suited for an analysis of spillovers in firms' entry decisions because a regional shopping center is a self-contained shopping hub where we observe the co-location of both complementary and substitutable retailers. In this paper, I focus on the entry decisions of anchor stores (traffic-generating stores) since these firms typically commit to a mall before smaller retailers. Most regional shopping centers have multiple anchor stores that are highly-substitutible competitors. Typically, these are competing department stores. A simple explanation for the coincidence of many highly competitive anchor stores might be that the profit-potential of the mall is sufficiently high to warrant head-to-head competition. However, interactions with mall planner practitioners indicate that another motive may arise from strategic complementarities: anchor stores generate traffic spillovers amongst themselves and to other retailers. Naturally, how much traffic each anchor store generates depends on its particular identity and also on the interaction with the other anchors in the mall.

Using a novel dataset containing information about the store configurations of all US regional shopping centers, I study the factors that influence store profitability in a mall. These factors can be separated into direct effects, such as observable demand characteristics, and strategic effects, caused by the impact of other firms' presence in the market. Hence, I formalize the observed configuration of anchor stores in a mall as the equilibrium outcome of a simultaneous moves entry game of incomplete information where each store's entry decision is affected by both the direct and the strategic effects. This framework is convenient for several reasons. First, by using a game theoretic model I can accommodate the simultaneity of firms' entry decisions. Second, the incomplete information nature of the game allows me to estimate a more realistic model with a relatively large number of players. Last, this approach does not require any price or quantity data since market structure alone is enough to obtain information on firms' profits.

While there is a well-established literature on empirical entry games, the extant work typically assumes away one of the types of spillovers I want to study. Studies in this literature usually impose that the entry of additional firms in a market has always a negative impact on firms' profits. This assumption is often made for econometric identification purposes as games with strategic complementarities are more prone to generate multiple equilibria. In turn, the existence of multiple equilibria brings difficulties to the estimation of strategic entry games. For full-information estimation methods, such as nested fixed-point approaches (e.g. Seim, 2006; Orhun, 2005 and Zhu and Singh, 2005), it is difficult or infeasible to construct a likelihood function in the presence of multiple equilibria. To circumvent this problem, researchers have used two-step methods (e.g. Bajari et al, 2006 and Ellickson and Misra, 2006) that are computationally simple. However, these approaches do not utilize all the information from the model. They also require a non-parametric first-stage which is infeasible in most empirical contexts.

In this paper, I develop a model of entry that allows for positive and negative spillovers among firms. I address the estimation difficulties that arise due to the presence of multiple equilibria by making use of an insight from Judd and Su (2006). I depart from the recent literature on the estimation of discrete choice static games by formulating the entry game as a Mathematical Problem with Equilibrium Constraints (MPEC). Specifically, I use a direct optimization approach that consists of maximizing the likelihood function subject to the constraint that the equilibrium conditions given by the economic model are satisfied. Using state-of-the-art constrained optimization programs, I obtain the true maximum likelihood estimates of the parameters of the model without having to restrict a priori the domain of the parameters while also allowing for the presence of multiple equilibria. While this paper constitutes the first attempt to use this direct optimization approach to address a specific empirical problem, this method can be used in a wide range of structural estimation problems other than entry models.

In addition to resolving the computational problems associated with multiple equilibria, I also tackle some of the standard econometric identification problems from games. I explore several plausible exclusion restrictions that help in resolving identification for the shopping mall entry game studied.

Two main conclusions may be drawn from the estimation of the entry model using the shopping center data. First, the empirical evidence from the Shopping Center Industry strongly supports the agglomeration and clustering theories that predict firms may have incentives to locate together despite potential business stealing effects. This result sustains the notion that, in some empirical settings, it is not realistic to assume that entry of additional firms always decreases the profits of the other firms in the market. Second, the empirical results demonstrate that the firms' negative and positive strategic effects (i.e. those effects that follow from the firms' interaction) help predict both how many firms can operate profitably in the market and the firm-types configurations. In some cases, the strategic effects can be large enough to outweigh the effect of the market demographics. The relative magnitude of such effects varies substantially across store-types. For example, I find that demographic market characteristics are the main determinant of long-run profits for discount department stores (e.g. Target). However, this is not true for more upscale department stores (e.g.

Macy's) for which the strategic competitive effects (from other anchors of the same type) and the positive spillovers from lower-scale department stores seem both to be the main determinants of long-run profits.

2 Related Literature

This paper draws on the entry, agglomeration and real estate/shopping center literatures.

Entry literature

Early empirical studies of entry (Bresnahan and Reiss 1991, Berry 1992) have recognized that a firm, when faced with the decision of entering a new market, must weigh the intensity of competition against the available market demand. Here, the increase in competition adversely affects profitability as more sellers share a common consumer base.

Berry (1992) allows for differences in fixed costs across firms and models the number of firms as a function of market and firm characteristics. Nevertheless, the differences in firms in his model do not give rise to varying intensities of competition between the dissimilar competitors.

Mazzeo (2002) and Seim (2006) were the first to allow for different degrees of product differentiation across firms. Mazzeo (1998) develops an equilibrium model of entry and quality-level choice and looks at highway-motels differentiation via quality levels. Seim (2006) studies firms' joint entry and location choices in retail markets where location choice constitutes a form of product differentiation, thus allowing for a larger number of product types than Mazzeo. She uses spatial differentiation as a representative form of product differentiation and looks at video stores' joint entry and location decisions.

By allowing for product differentiation, Mazzeo and Seim add a new dimension to the entry problem and explicitly incorporate the idea that firms may want to differentiate in order to soften competition.

One critical assumption made in all the above mentioned papers is that entry of an additional firm into the market always decreases incumbents' profits. Furthermore, this decrease will be larger if the new market participant is of the same type (or located close by) as the other firms in the market. While this assumption seems to be crucial for ruling out multiple equilibria, there are several applications in which it may not be very realistic. Neither Mazzeo nor Seim consider the positive externalities that may result from the close location (either in terms of attributes or physical location) of the players. In Seim's case it is not very likely that video-stores benefit from locating close to each other but in the highway-motel case, we could think of a story where the reason why so many motels locate near one another may not have exclusively to do with the fact that demand is higher in certain particular locations but also with the existence of positive firm spillovers. Consumers may find convenient to exit the highway only once in order to search for a motel to stay, either because they may want to compare motel quality/prices, or because it makes it easier to search next door if a given motel is fully booked.

[TO BE COMPLETED]

Agglomeration literature

Several papers (see Fujita and Thisse 1996 for a literature review) have studied the economics of firms' agglomeration and tried to understand the reasons behind the formation of economic clusters[1].

Krugman (1991) formally models what Myrdal (1957) called "circular causation": manufactures production will tend to concentrate where there is a large market, but the market will be large where manufactures production is concentrated.

Essentially, the agglomeration literature provides a rationale for why firms have incentives to co-locate: some possible reasons include the existence of externalities (either technology spillovers or demand externalities) or increasing returns (through a interaction of economies of scale with transportation costs) or natural advantages (related to the characteristics of the cluster's physical location).

Although the above mentioned papers refer mostly to the industrial setting, one type of agglomeration arises when retailers selling similar products are clustered within the same neighborhood of a city. Indeed, the phenomenon of store agglomeration has been examined in the theoretical work of Eaton and Lipsey (1979), Stahl (1982), Wolinsky (1983) and Dudey (1990). The main purpose of these articles is to derive conditions for a cluster of stores to exist in equilibrium. They show that firms selling similar or identical products may locate together even though this results in fiercer local competition (Dudey (1990)). This is the case when a firm which locates with a few of its competitors may be able to draw consumers (who are attracted to areas of high competition) away from other rivals.

[TO BE COMPLETED]

Real Estate and Shopping Center Literatures

[TO BE COMPLETED]

3 Industry background and description

Forty years ago, shopping centers barely dotted the retail landscape of the United States. Today, the number of shopping centers surpasses 48,000 and the shopping mall has become an integral part

[1] According to Porter (1998), clusters can be defined as "Geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions (for example, universities, and trade associations) in particular fields that compete but also co-operate." (Porter, 1998, p.197)

Also according to Porter, "A cluster is a form of network that occurs within a geographic location, in which the proximity of firms and institutions ensures certain forms of commonality and increases the frequency and impact of interactions." (Porter, 1998, p.226) of the economic and social fabric of the US. In 2005, about 190 million adults shopped at various US shopping centers each month. Also in 2005, shopping-center inclined sales [2] were estimated at $2.12 trillion and state sales tax revenue from shopping center-inclined sales generated $114 billion.

According to the 2005 CoStar/NRB[3] Shopping Center Census, there are more than 48,600 shopping centers[4] in the United States (2005). In 2005, approximately 60% percent of the US shopping centers had a Gross Leasable Area (GLA) of 100,000 square feet or less, 25% had a GLA between 100,000 and 200,000 square feet while 15% exceeded 200,000 square feet in size.

Shopping Center Definitions and Types in the United States

According to the most recent definition given by the International Council of Shopping Centers (ICSC), a shopping center is "A group of retail and other commercial establishments that is planned, developed, owned and managed as a single property, with on-site parking provided. The center's size and orientation are generally determined by the market characteristics of the trade area served by the center." The ICSC has subdivided shopping centers into eight types (see Table 1 in the Appendix for a brief definition of each of these types[5]).

The development process of a Shopping Center[6] and the importance of the Anchor Stores

The development of a shopping center requires certain essential stages and comprises a complex series of economic, financial, merchandising and design decisions. A range of technical experts is usually called upon to carry out development. They include a market analyst, land planner, architect, landscape architect, lawyer, engineering and construction specialist, accountant, financial adviser, leasing agent and shopping center manager. In some cases, several or all of these disciplines are represented on the developer's own staff; in others, the shopping center developer hires specialized consultants.

The development process of some major regional shopping centers has taken 15 years or longer. A feasibility study for construction of a shopping center is, in general, comprised by the following parts:

- Market analysis, including an evaluation of the existing competition, the potential future competition, the economic characteristics of the metropolitan area, the size, characteristics and demographics of the trade area, access and visibility, and factors affecting growth of the trade area;

[2] Shopping center-inclined sales, based directly on U.S. Commerce Department data, simply measure sales at stores that are likely to be at shopping centers. Shopping center-inclined sales include the following store types: General Merchandise, Apparel, Furniture, Electronic and "Other" (GAFO), Health and Personal Care, Food and Beverage, and Building Material & Garden Equipment & Supplies.

[3] National Research Bureau

[4] NRB defines a shopping center as three or more stores built as a unified structure.

[5] The definitions, and in particular the table in the appendix, are meant to be guidelines for understanding major differences between the basic types of shopping centers. Several categories shown in the table, such as size, number of anchors, and trade area, should be interpreted as "typical" for each center type. They are not meant to encompass the operating characteristics of every center. As a general rule, the main determinants in classifying a center are its merchandise orientation (types of goods/services sold) and its size.

[6] For more details on the shopping centers' development process please refer to the book published by the Urban Land Institute entitled "Shopping Center Development Handbook".

- Financial analysis;
- Site selection and evaluation;
- Commitments from key tenants;
- A leasing plan;
- Financial considerations;
- Zoning, subdivision, environmental and traffic impact, and other public approvals.

No precise "how-to" formula holds, because interlocking preliminary work precedes any final decision. Ultimate site acquisition, for example, depends on zoning approval. In turn, leasing depends on site approvals and on securing key tenants, while financing depends on leasing. As a rule, though, a regional shopping center will not be built until the developer has secured commitments from key tenants (also known as anchor stores). Because a developer is unlikely to begin construction of a new shopping center without a commitment from a construction lender, and construction financing is rarely available without a commitment from an anchor tenant, unwarranted construction of shopping centers is substantially reduced. Anchor tenants also ensure financial security because they lease a large block of space which constitutes a substantial portion of the shopping center for an extended period of time (typically more than 20 years). Moreover, once they locate in an area, they are not likely to dilute their market share by opening another location nearby.

For a proposed regional shopping center, the strength of the anchor tenants interested in the selected site determines the size, character and success of the center - even the price that can be paid for the land. *The Shopping Center Development Handbook* even goes to the extreme of characterizing the market study for a shopping center as a "chicken-and-egg conundrum": a potential key tenant will not be interested in a center until a market analysis has been completed, but a thorough analysis cannot be completed until it is known what kinds of key tenants the area will attract.

The anchor tenant contributes enormously to the center's cumulative attraction potential and is the generative source of retail activity. Because the anchor is the most important component of a shopping center, this tenant is able to extract favorable rent rates. An anchor tenant's gross leasable area rent can be one-third to one-half of the rate paid by the remaining tenants of the shopping center. Anchor tenants argue that their expenditures for advertising, which generates customer traffic to the center, plus their reduced rent provide financial benefits equal to those provided by tenants who pay higher rents but incur little or no advertising expense.

While anchors attract consumers to malls, not all anchor stores are alike. This heterogeneity implies that some anchors create greater externalities than others. The ability of anchor tenants to generate traffic externalities depends on their particular type/identity and also on their interaction with other specific anchors/tenants in the mall. Comparing and measuring these effects is part of what I propose to do in this paper.

4 Data

4.1 Data Sources and Market Definition

Shopping Centers have been the strongest area in real estate in recent years. Despite the importance of the industry, it has been largely overlooked by the Marketing Literature. The dearth of empirical studies of malls is partially explained by the absence of detailed and extensive mall store data. Developers are reluctant to give researchers access to confidential data such as sales, rent, size, store name, product category, lease length, year the lease began, and other contract provisions for the stores in the malls for which they are responsible.

In some cases, such data is available but only for relatively small mall samples (e.g. Gould et al, 2002 make use of detailed data for over 2,500 stores in around 35 large malls across the United States), which restricts the potential areas to be studied in this industry. In order to study the determinants of shopping center configuration and measure stores' strategic effects without any price or quantity data, empirical models of market structure are especially convenient since these are primarily based on the observed number and types of stores in a given market.

In this paper, a market is defined as a regional shopping center. A regional shopping center/mall is a shopping center that provides general merchandise and services in full depth and variety, serving a very large trading area of up to 15-20 miles. A typical regional center is usually enclosed with an inward orientation of the stores connected by a common walkway. Parking surrounds the outside perimeter. Typically, the total center store area ranges from 300,000 square feet to 850,000 square feet with 40 to 100 stores anchored by one or more department stores. For the purposes of analysis, I focus on the entry decisions of *anchor stores* (i.e. traffic generating stores), since, as explained in Section 3, these firms commit to a mall before smaller retailers.

To estimate the empirical model of entry described in Section 5, we need a dataset with information for a cross-section of regional shopping centers operating in the United States. There are two major sources of data that can be used to study the problem at hand. *The Shopping Center Directory* is a directory published by the National Research Bureau (NRB) and consists of a directory of shopping centers of all sizes. *The Directory of Major Malls*(DMM) is published by Directory of Major Malls, Inc. and, despite being a smaller directory that focuses specifically on centers with a gross leasable area of approximately 250,000 square footage and above, it tends to have more detailed information for this subset of data. According to the NRB data, there are around 2,400 regional centers in the United States but this number includes also some community and super-regional centers.[7] In fact, my primary data source, the DMM database, classifies only 830 centers as being regional in the United States. Since their definition of regional shopping center seems to be more consistent with the (somewhat broad) definition given by the International Council of Shopping Centers I restrict the dataset to these 830 malls. Key variables for each center include the identity of the anchor tenants, gross leasable area, the year the center opened, whether the mall is

[7]This is the case since NRB includes all centers from 300,000 up to 749,999 square feet in the regional center category (without taking other categorization criteria into consideration).

open or enclosed, the number of stores and the center's geographic coordinates. Also available are a few demographic variables such as average household income, number of households, population and median age statistics for 5, 10 and 20 mile radius zones around each shopping center. The data refer to January of 2006.

The data for the selected sample of malls were cross-checked for discrepancies with the NRB data. For the cases in which there were discrepancies between these two data sources, data entries were manually checked and corrected using several other sources such as the shopping centers' and mall developers' websites. After dropping miscategorized shopping centers (i.e. centers that are classified as regional but that do not satisfy the criteria for regional shopping centers) and centers with too many missing variables I ended up with a sample of 561 regional shopping centers.

Anchor stores for each mall were classified into several broad categories/types based upon their retail model and format. I examine the existence of inter-firm strategic effects within and across stores in these categories. To keep the model as parsimonious as possible, anchor types with small presence in regional shopping centers were dropped, thus restricting our attention to three main types of anchors which comprise approximately 80% of the total number of anchors in the sample. The three types of anchors studied are broadly defined as follows[8] :

Upscale Department Stores: Stores that generally sell designer merchandise above an average price level. When their items are on sale, their prices resemble those of average priced items at a lower scale department store. Upscale department stores typically provide checkout service and customer assistance within each department. Examples include Dillard's, Macy's and Nordstrom.

Midscale Department Stores: These stores sell brand names and non-brand names but don't sell upscale brand names. When compared with upscale department stores, midscale stores usually don't have perfumes and beauty supplies at the main entrance and don't have cosmetic specialists. Examples include JCPenney, Mervyn's and Kohl's.

Discount Department Stores: These stores encompass retail establishments selling a variety of merchandise for less than conventional prices. Target, Sears, Wal-mart and Kmart are examples of this store type. Most discount department stores offer wide assortments of goods; others specialize in such merchandise as jewelry, electronic equipment, or electrical appliances. Discount stores are not dollar stores, which sell goods at a dollar or less. Discount stores differ because they sell branded goods and prices vary widely between different products. When compared to midscale department stores, discounters sell fewer major brand names and offer a wider variety of products. Stores in the department-discount category typically have fewer sales workers, relying more on self-service features, and have centrally located cashiers.

[8]The classification of stores into these three categories agrees with classifications used in the Shopping Center's and Department Store's industries. It allows me to encompass over 80% of all anchor stores in the sample without having to create an unreasonable number of categories. I also experimented with other classifications but the level of homogeneity within categories was not as large as the one in the classification used; for example, grouping all upscale and midscale department stores into one single category would imply putting together stores with very different characteristics, costs and target segments.

4.2 Variables Description

Table 2 provides a summary of the variables[9] used to estimate the model. The demographic variables (Panel A in Table 2) include the purchasing power, median age and average household size of the population within a 20 mile radius of the mall and the percentage of single family houses in each mall's three-digit ZIP code[10]. All of these variables are seen by mall developers and potential tenants as key profit determinants.

Several shopping center-specific characteristics are also included in the model to control for other factors that can potentially influence stores' profits. Such variables are the number of parking spaces, the opening date and the developer of each shopping center. The number of parking spaces variable is used to control for the size of the shopping center since larger shopping centers are naturally expected to be able to accommodate more and bigger anchor stores. Note that the fact that a shopping center is larger does not necessarily have to do with the ability of that center to draw demand; in many cases the size of the shopping center is a reflection of factors such as the cost of the land in a given area, geographical constraints or even zoning restrictions (e.g. in response to environmental concerns). In order to reflect the fact that store location preferences (or costs) might vary per region, a dummy variable indicating shopping centers located in the west region of the United States is included.

I also include the date when the mall opened to control for the fact that more recently opened malls may have achieved a better match between the products sold by their stores and their current demand configuration and to control for any time trends in the industry. The categorical date variable used should also reflect the effects of a mall's physical deterioration since older malls often have lower sales than more recent malls.

Other variables that are important for the success of a shopping center have to do with the center's accessibility and its visual exposure. The fact that these are seen as essential requirements when planning larger shopping centers[11], allows me to exclude such characteristics from the estimation since they are, in general, shared by all the centers in my sample.

Lastly, several firm-specific variables (Panel B in Table 2) were also considered. Namely, the identity of the each anchor-store, the size (in square feet) of the store and the number of stores of the same chain in the same state.

These firm-specific variables, together with the *West* dummy variable, are of special importance for model identification purposes, as I discuss in section 5.3.2.

[9] All variables (except dummies) were rescaled since the optimization solver can be mislead when variables have very different orders of magnitude. Details on the transformation used can be found in Table 2's note.

[10] The data for this last variable were obtained from the 2000 Census for each mall's three-digit ZIP code. Three-digit ZIP codes average about 1000 square miles of overall land and water area. If they were circular, they would have mean radius of 15.1 miles. As mentioned before, this constitutes a reasonable radius for a regional shopping center trade area.

[11] The Shopping Center Development Handbook states: "Regional centers customarily are located on sites that are the most accessible...Interchange points between expressways and freeways in suburban areas are the most visible...The distance from an interchange to a regional center ideally is less than one-half mile to a mile, depending on local circumstances; but in all cases the center should be clearly visible."

5 A Model of Entry in a Cluster

5.1 Model Setup

I model the entry of firms (i.e. anchor stores[12]) in a market or retail cluster (i.e. a regional shopping center[13]) as a simultaneous move entry game of incomplete information. This assumption deserves some explanation. Although it is true that shopping centers are to same extent planned, as discussed before, it is also true that mall planners cannot freely determine which stores will enter in a given market. Thus, in order to specify how retailers enter the mall, two competing and extreme approaches could be taken. The first assumes that the mall planner has full control over the choice of stores; this assumption would allow us to model the composition of the mall as a discrete choice problem, where the mall planner chooses the tenant mix that maximizes the sum of expected profits of the retailers. However, given the large number of possible store combinations, this approach is not feasible in practice. The second approach follows the standard entry literature and assumes that potential retailers make their decisions to enter the market based on expected profits and taking into account the actions of the other retailers, with no intervention by the mall planner. I follow an intermediate approach.

In the model, I assume that a mall developer, having previously defined the location of the mall, approaches a set of potential entrants and provides them information about the characteristics (e.g. demographics) of the mall's geographical location. With this information, the potential entrants play an entry game of incomplete information. Since the game may have many possible equilibria, it is assumed that the mall planner chooses the equilibrium that maximizes some criteria. It should be emphasized, however, as it will become clear in Section 5.4, that this equilibrium selection rule does *not* play *any* explicit role in the estimation of the model. Discussions with mall planners indicate that the process leading to the observed configuration of anchor stores in a mall has indeed properties that are similar to those of a simultaneous entry model.

Consider a game where a finite number of potential stores indexed by $i = 1, .., n$, choose simultaneously whether or not to enter in a given market. Each potential entrant can be of one of the following three types $T = 1, ..., 3$, where $T = 1$ if it is an upscale department store, $T = 2$ if it is a midscale department store and $T = 3$ if it is a discount department store. These three store types were previously defined in Section 4.1.

Unlike Mazzeo (2002) and others, I take store types to be exogenously given. This is not an unreasonable assumption for many retail stores since usually anchor-store brands are already established and thus only face the decision to enter or not in a given market. The degree of store/product differentiation (either vertical or horizontal) that a market achieves thus depends on the simultaneous entry decision of many stores.

Let $a_i \in \{0, 1\}$ define the action of store i in which $a_i = 1$ if the store decides to enter the mall and $a_i = 0$ if it decides not to enter. Let $A = \{0, 1\}^n$ denote the vector of possible actions of all stores and let $a = (a_1, ..., a_n)$ denote a generic element of A. Finally, let $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$

[12]Henceforth, I will use "firms" and "stores" interchangeably.

[13]Henceforth, I will use "markets" and "regional shopping centers" interchangeably.

denote a vector of strategies for all stores excluding store i.

The vector of state variables for store i in market m is denoted by $x_{im} \in X_{im}$. In the empirical implementation of the model, these variables include market specific demographics and store specific variables. Define $X_m = \Pi_{i=1}^{n} X_{im}$ and let $x_m = (x_{1m}, ..., x_{nm}) \in X_m$ denote the vector of state variables of all stores in market m. In the model, I assume that x_m is common knowledge to all players in the game as well to the econometrician.

$\epsilon_i(a_i)$ is an idiosyncratic term which is private information to store i. It is assumed to be independently and identically distributed across stores and actions with a Type 1 extreme value distribution that is common knowledge. The density of $\epsilon_i(a_i)$ is denoted as $f(\epsilon_i(a_i))$ and the density of $\epsilon_i = (\epsilon_i(0), \epsilon_i(1))$ as $f(\epsilon_i)$.

The ex-post profits earned by store i are a function of $\epsilon_i = (\epsilon_i(0), \epsilon_i(1))$, x_m and the actions of the other stores, a_{-i}, and are given by

$$\tilde{\Pi}_i(a_i, a_{-i}, x_m, \epsilon_i; \theta, T) = \begin{cases} \Pi_i(a_{-i}, x_m, T; \theta) + \epsilon_i(1) & \text{if } a_i = 1 \\ \epsilon_i(0) & \text{if } a_i = 0 \end{cases} \tag{1}$$

where $\Pi_i(a_{-i}, x_m; \theta, T)$ is a known and deterministic function of the states and actions of player i of type T, and θ is a vector of parameters, which may be firm or type specific. In (1), the deterministic part of profits for the no-entry option has been normalized to zero.

Because a given store i does not observe the other stores ϵ_i's, that store must construct beliefs about the other stores' expected actions using all relevant, observable information. Given these beliefs, the store must make a decision, $a_i \in \{0, 1\}$, to maximize its expected profits. By the independence of private information, the expected profits for firm i associated with the decision a_i equal

$$E[\tilde{\Pi}_i(a_i, x_m, \epsilon_i, T; \theta)] = \sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta)\, \sigma_{-i}(a_{-i}|x) + \epsilon_i(a_i = 1) \tag{2}$$

where $\sigma_{-i}(a_{-i}|x_m) = \Pi_{j \neq i} \sigma_j(a_j|x_m)$. $\sigma_{-i}(a_{-i}|x_m)$ represents the beliefs that firm i has about the probability of entry of the other firms in the market. For each $j \neq i$, the probability (from firm's i perspective) that firm j enters the market is given by

$$\sigma(a_j = 1|x_m, \theta, T) = \Pr\{\Pi_j(a_{-j}, x_m, T; \theta) + \epsilon_j(1) \geq \epsilon_j(0)\} \tag{3}$$

In a Bayesian Nash Equilibrium, firms' beliefs are equal to their actual choice probabilities, so the Bayesian Nash equilibrium to the static entry game is a collection of beliefs $\sigma_i^*(a_i = 1|x_m, \theta, T)$ for each store $i = 1, .., n$.

The distribution of ϵ implies that the probability, prior to the realization of the ϵ's, that firm i chooses to enter market m is

$$\sigma^*(a_i = 1|x_m, \theta) = \frac{\exp\left(\sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta)\, \sigma^*_{-i}(a_{-i}|x_m)\right)}{1 + \exp\left(\sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta)\, \sigma^*_{-i}(a_{-i}|x_m)\right)} \text{ for } i = 1, .., n \tag{4}$$

The equilibrium probabilities σ^* solve the system of equations in (4) and clearly depend on the market characteristics x_m and on the structural parameters θ. For any (x_m, θ) the game has multiple equilibria if more than one value σ^* satisfies (4).

The properties of this model are well understood. In particular, Brower's fixed point theorem guarantees the existence of an equilibrium to this model for any finite x_m (see Section 5.3 for more details). In addition, provided that the distribution of $\epsilon_i(a_i)$ is atomless, the best reply function in (4) is unique, and thus we don't need to consider mixed strategies. Of course, there is no guarantee that the equilibrium is unique since multiple sets of beliefs $\sigma_i(a_i = 1|x_m)$ can satisfy the system (4) . I discuss this issue in more detail in Section 5.4.

5.2 Profit Functions

I specify the profit functions to be linear in the common state vector x_m and on the actions of the other stores. This parametric specification can be considered as a first order approximation to a variety of strategic models. Alternatively, to obtain the profit functions, one could start from primitive assumptions regarding supply and demand in the mall and derive the profit functions from the equilibrium conditions. However, without any price, quantity, or sales data, and with very limited information on store characteristics, this approach would be extremely demanding on the data and would rely heavily on the primitive assumptions.

Specifically, I assume that the deterministic component of the profit function of store i of type T is given by

$$\Pi_i\left(a_{-i}, x_m, T; \theta\right) = \alpha_i + \beta_i' x_m + \delta_{TT} \sum_{j \neq i} 1\left\{a_j = 1, T\right\} + \sum_j \delta_{\tilde{T}T} \cdot 1\left\{a_j = 1, \tilde{T} \neq T\right\} \tag{5}$$

where α_i is the store-specific mean profitability level, β_i measures the impact of the state variable x_{im} on store i's profits, δ_{TT} measures the impact of the total number of other anchors of the *same* type as store i and $\delta_{\tilde{T}T}$ measures the impact on profits from different type ($\tilde{T} \neq T$) stores.

A key parameter in the model is δ_{TT}. Standard oligopoly models predict that $\delta_{TT} < 0$ due to the competition (business stealing) effect, when demand is exogenous and there are no positive spillovers between stores. Most of the entry models take $\delta_{TT} < 0$ as given (thus imposing this restriction in the estimation) to conveniently reduce the multiple equilibria "problem". However, when demand is endogenous, as in a regional shopping center, firms may benefit from entering together, as discussed in the introduction, in which case this parameter does not have to be necessarily negative.

The mean profitability parameter α_i captures factors that are related to the efficiency of the stores as well as with other cost factors such as the level of rents, provided that the level of the rents for a given store is similar across regional shopping centers.

Under the assumption that the profitability shocks of entering a market are distributed Type 1 extreme value, the system of equations that define the equilibrium can be written as

$$\sigma^*(a_i = 1|x_m, T; \theta) = \frac{\exp\left(\alpha_i + \beta_i' x_m + \delta_{TT} \sum_{j\neq i} \sigma^*(a_j=1|x_m,T) + \sum_j \delta_{\tilde{T}T} \cdot \sigma^*(a_j=1|x_m,\tilde{T}\neq T)\right)}{1+\exp\left(\alpha_i + \beta_i' x_m + \delta_{TT} \sum_{j\neq i} \sigma^*(a_j=1|x_m,T) + \sum_j \delta_{\tilde{T}T} \cdot \sigma^*(a_j=1|x_m,\tilde{T}\neq T)\right)} \quad \text{for } i = 1,..,n. \tag{6}$$

5.3 Model Identification

In this section, I discuss the implications for the model's identification of different assumptions that are usually made in the entry literature.

One of the most important dimensions on which entry models usually differ has to do with the extent to which firms are assumed to be uncertain about the payoffs of other firms. Different assumptions about the information structure of the game will lead to different identification concerns which in turn will lead to different identification and estimation approaches. As discussed in the previous section, in this paper's empirical application, stores are assumed to play a game of incomplete information. However, and as to provide the necessary background, I will start by briefly reviewing the identification issues encountered in complete information games and the approaches that have been used to deal with them.

5.3.1 Identification Issues in Entry Games of Complete Information

If players possess full information about other players' payoffs and only pure strategies[14] are played, players' optimal strategies can be represented by a simultaneous discrete response system of the type studied by Heckman (1978). An example of such system can be the following model with two firms

$$\Pi_1 = x\beta_1 + \delta_{21} \cdot y_2 + \varepsilon_1$$
$$\Pi_2 = x\beta_2 + \delta_{12} \cdot y_1 + \varepsilon_2$$

$$y_i = \begin{cases} 1 & \text{if} \quad \Pi_i > 0 \\ & \qquad\qquad \text{for } i = 1, 2. \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where the set of endogenous variables Y consists of $Y = \{(0,0), (0,1), (1,0), (1,1)\}$, which in our case correspond to the entry decisions of the two firms.

In this type of system (i.e. a simultaneous-equation model with underlying discrete latent variables), a well-defined likelihood function for the set of observable outcomes exists if and only if the mapping from equilibria to inequalities on profits is a well-defined function. This means that,

[14] Aradillas-Lopez (2005) has shown that if mixed-strategies are allowed, a well-defined likelihood function exists for all the outcomes of the game under conditions much weaker than when only pure-strategies are allowed. Anyway, in this case, identification cannot still be achieved without any additional assumptions.

in order to have a well-defined likelihood function, the firms' equilibrium strategies (which depend on the error terms, the market observables and the model parameters) have to exist and be unique.

When a game has multiple equilibria, there is no longer a unique relation between players' observed strategies and those predicted by the model. As shown in Bresnahan and Reiss (1991), this always happens in simultaneous-move Nash models when the errors have sufficiently wide supports.

Amemiya (1974), Heckman (1978) and others have shown that the only way to get a well-defined likelihood function in this type of model is by imposing a "coherency condition". Such condition seems undesirable in an entry model since it does not preserve the simultaneous-interaction element of the model. In order to avoid the coherency condition, Bresnahan and Reiss (1990, 1991) impose the constraint that entry of an additional firm is always costly (i.e. it always reduces the other firms' profits) and redefine the space of outcomes (by aggregating nonunique equilibrium outcomes) in such a way that the game is transformed into one that predicts unique equilibria.[15]

Tamer (2003) proposes, instead, the use of a semiparametric estimation procedure based on the probability bounds for each outcome that results from the game's Nash Equilibrium conditions.

The disadvantage of these approaches is that they result in efficiency losses and limit the ability of the econometrician to make predictions over the entire set of observable outcomes.

Another way of dealing with the multiplicity of equilibria issue is to develop some theory of equilibrium selection. For example, one can make an assumption on the order in which the players move (as in Berry, 1992 and Mazzeo, 2002)[16] or model more formally (i.e. parametrically) an equilibrium selection mechanism (as in Bajari, Hong and Ryan, 2004). The use of an appropriate equilibrium selection rule assures the existence of a well-defined likelihood function for the entire space of observable outcomes . The problem with this approach is that consistency of the estimation depends critically on the validity of the assumed selection rule which is not testable.

5.3.2 Moving to Games of Incomplete Information

In this paper's empirical application, agents are assumed to play an incomplete information game (i.e. like the econometrician, potential entrants have incomplete information about each others' profits). This implies that the estimation approach used, which is described in section 5.4, must be based on a representation of equilibria in the space of players' choice probabilities. That is, equilibrium choice probabilities solve an equilibrium mapping that can be called "best response probability mapping".

The nature of equilibria in games with incomplete information is different from the complete information case. Since strategies are now mappings from players' types to actions where types are private to the players, rivals discrete actions are transformed into smooth predicted entry

[15] Note that if there is firm heterogeneity, even imposing these conditions, we can have multiple equilibria with different number of firms.

[16] Actually, imposing an order of entry may not be enough to guarantee uniqueness. For example, Mazzeo (2002) rules out payoffs increasing in the number of competitors and imposes that same-type competitors must affect profits at least as much as different-type competitors.

probabilities. The discrete action space is "transformed" into a continuous space.

Introducing private information does not necessarily eliminate the problems found in complete information games. Depending on the game primitives, the introduction of expectations about other players' profits may or may not ameliorate multiplicity and non-existence problems. For example, Seim (2006), in her incomplete information game with simultaneous entry decisions shows (using Monte-Carlo simulations) that, under certain parameter constraints, her model has unique equilibria.[17]

Conditions for identification of the model's structural parameters

A model is said to be identified if there exists a unique set of model primitives (θ) that can be inferred from a sufficiently rich dataset characterizing the firms' decision probabilities. Here, I'll study the nonparametric identification of the model described in Section 5. Although my estimation strategy relies on parametric methods, it is well known that a (parametric) estimation approach will be more valid if it does not rely exclusively on functional form assumptions.

I'll proceed by first establishing sufficient conditions for nonparametric identification of the deterministic part of the expected payoff functions $E[\Pi_i\left(a_i, x_m, \epsilon_i; \theta, T\right)]$ and then by looking at the identification of the deterministic part of profits, $\Pi_i\left(a_i, x_m, \epsilon_i; \theta, T\right)$, taking the expected profits as given.[18] The insights for this identification strategy are borrowed from the works of Hotz and Miller (1993) and Magnac and Thesmar (2002), in the context of dynamic games, and from Bajari et al (2006) in the static-game setting.

For purposes of analysis, the distribution of the players' private information will be taken as given since Rust (1994) and Magnac and Thesmar (2002), among others, have shown that in the absence of a distribution for the private information the model is clearly not identified. Furthermore, Bajari, Hong and Ryan (2004) also demonstrate that when one allows the deterministic part of profits to be nonparametric, an independence assumption is required for identification. Given the above discussion, I impose the following assumption:

Assumption 1 *The error terms ϵ are distributed i.i.d. Type-I extreme value across actions a_i and players i.*

Based on this assumption, the equilibrium in the model must satisfy:

$$E[\Pi_i\left(a_i = 1, a_{-i}, x_m, T; \theta\right)] + \epsilon_i\left(a_i = 1\right) \geq E[\Pi_i\left(a_i = 0, a_{-i}, x_m, T; \theta\right)] + \epsilon_i\left(a_i = 0\right) \quad \forall i \tag{8}$$

[17] The problem here is that, in order to estimate the model using real data, those constraints need to be incorporated in the estimation and thus are not testable. Actually, Seim does not impose any parameter constraints during the parameter search but this may result in inconsistent estimates since her fixed point problem needs to be solved at each iteration. Even if the final parameter estimates correspond to unique equilibria, it is possible that during the search through the parameter space, the parameters end up in regions where there are multiple equilibria.

[18] Non-parametric identification of the model is then studied by looking at the space of expected payoffs. An alternative and equivalent approach relies on the examination of the the system of equations with the choice probabilities in (13) (see Pesendorfer and Schmidt-Dengler, 2006).

which, in turn, implies the following expressions for the equilibrium choice probabilities:

$$\sigma^*(a_i = 1|x_m, \theta) = \frac{\exp\left(E[\Pi_i\left(a_i = 1, a_{-i}, x_m, T; \theta\right)] - E[\Pi_i\left(a_i = 0, a_{-i}, x_m, T; \theta\right)]\right)}{1 + \exp\left(E[\Pi_i\left(a_i = 1, a_{-i}, x_m, T; \theta\right)] - E[\Pi_i\left(a_i = 0, a_{-i}, x_m, T; \theta\right)]\right)} \quad \forall i \qquad (9)$$

Equation (9) maps $(E[\Pi_i\left(a_i = 1, \cdot\right)] - E[\Pi_i\left(a_i = 0, \cdot\right)])$ into $\sigma^*(a_i = 1|x_m, \theta)$. We can rewrite these equations as

$$(\sigma_i^*(a_i = 1|x_m, \theta_T) = \varphi\left(E[\Pi_i\left(a_i = 1, a_{-i}, x_m, T; \theta\right)] - E[\Pi_i\left(a_i = 0, a_{-i}, x_m, T; \theta\right)]\right) \qquad (10)$$

where φ denotes the mapping from choice specific profit functions to choice probabilities. Hotz and Miller (1993) show that equations in (10) can be inverted as

$$(E[\Pi_i\left(a_i = 1, a_{-i}, x_m, T; \theta\right)] - E[\Pi_i\left(a_i = 0, a_{-i}, x_m, T; \theta\right)]) = \varphi^{-1}\left(\sigma_i^*(a_i = 1|x_m, \theta\right) \qquad (11)$$

which, in practical terms, means that it is possible to recover $(E[\Pi_i\left(a_i = 1, \cdot\right)] - E[\Pi_i\left(a_i = 0, \cdot\right)])$ from choice probabilities σ_i^*.

Although the particular functional form of σ_i in (9) depends on the probability distribution chosen for ϵ, the above mapping holds for other error distributions than the Type 1 extreme value distribution as long as the chosen distribution has full support.

From (11), we can see that we will not be able to infer the absolute values $\Pi_i\left(a_i = 1\right)$ and $\Pi_i\left(a_i = 0\right)$, but only their difference. This means that a payoff normalization will be required for identification as payoffs are only determined relative to the payoff under action $a_i = 0$ (no entry decision). So, in order to allow for identification, I'll impose the following assumption

Assumption 2 $\Pi_i\left(a_i = 0, a_{-i}, x_m, T; \theta\right) = 0$

This is a standard identifying assumption in any multinomial choice model.

Now that I have laid out the assumptions needed in order to identify the expected profit functions $E[\Pi_i]$, we can turn to the identification of the profit functions Π_i given choice probabilities and expected profits.

Recall that the relation between the deterministic part of payoffs (i.e. Π_i) and the deterministic part of expected payoffs (i.e. $E[\Pi_i]$) can be written as

$$E[\Pi_i\left(a_i, x_m, T; \theta\right)] = \sum_{a_{-i}} \Pi_i\left(a_i = 1, a_{-i}, x_m, T; \theta\right) \sigma_{-i}\left(a_{-i}|x_m\right) \qquad (12)$$

Heuristically, the problem of non-parametric identification of the profit functions is equivalent to the problem of finding a solution to the system of linear equations in (12). Even after having established conditions for the identification of the values of $E[\Pi_i]$ and σ_{-i} in (12), it is clear that, without further restrictions, the system is not identified for Π_i. Holding x fixed, and making use of Assumption 2, this system has n equations[19] and $n \cdot 2^{n-1}$ unknowns.

[19]Here I am assuming that there is one player of each "type".

Note that the degree of under-identification in (12) increases with the number of agents. So, for example, if we have a game with two players and one additional player is added, the number of equations in the system increases only by 1 while the number of unknowns increases by 4.

In order to identify non-parametrically the payoffs Π_i in (12), we need to impose cross-equation restrictions (a.k.a. exclusion restrictions).[20] We can do this by including in the profit function firm-specific profit shifters. That is, covariates that influence directly the profit function of each one of the firms but not the others. Appendix 2 provides an example with two players where it is demonstrated how adding exclusion restrictions may help with identification.

I have shown that, in order to identify the profit function parameters, one has to rely on the mapping from the choice probabilities to the expected profit functions in (11). Since these choice probabilities represent the equilibria of an incomplete information game, we need to make sure that an equilibrium exists and make an assumption regarding the (possible) multiplicity of equilibria.

Existence and Multiplicity of Equilibria

Ensuring *existence* of equilibria in this setting corresponds to making sure that the system of equations given by (13) has at least one solution. From Assumption 1, firms' conjectures are monotonic, continuous, and strictly bounded inside the set (0,1), so existence of a solution to the system of equations follows immediately from Brower's Fixed Point Theorem (see McKelvey and Palfrey, 1995).

Since, in general, there will be values of the model primitives for which the model has *multiple* equilibria (i.e. there may be multiple solutions to the system in (13)), an additional identification assumption is in order.

Assumption 3 *Given a value for the primitives of the model $\Omega = \beta, \delta, X$, players (or nature) select only one equilibrium from the set of possible equilibria and they do not switch to other equilibria as long as Ω does not change.*

This assumption is a standard assumption in two-step methods used for estimating incomplete information games and is implicit when applying any instrumental variables technique (i.e. the instrumental variables technique assumes that there is a reduced-form functional relationship between endogenous variables and exogenous variables which is equivalent to a single equilibrium selection).

5.4 Estimation Method

Judd and Su (2006) propose a new optimization approach to estimate structural models using maximum likelihood. The proposed approach fits in a class of problems known as *Mathematical*

[20]Related identification results have been obtained in the single agent dynamics literature, see Rust(1994) and Magnac and Thesmar(2002), and in the multiple agent dynamic literature (see Pesendorfer and Schmidt-Dengler, 2006). Bajari et al (2006) also demonstrate the need for such exclusion restrictions in the context of static games.

Programs with Equilibrium Constraints (MPEC) and can be used to estimate a variety of economic models. Following Judd and Su's insight, I choose the structural parameters and endogenous economic variables in order to maximize the likelihood function subject to the constraint that the endogenous economic variables are consistent with the equilibrium defined by the structural parameters. More specifically, I proceed in only *one* step, attaching Lagrange multipliers to the constraints and choosing the parameters that maximize the following constrained likelihood function

$$L(\beta,\delta) \quad = \quad \prod_{m=1}^{M}\prod_{i=1}^{n}\left(\sigma^*(a_i=1|x_m,T;\theta)\right)^{1\{a_i=1,T\}}\cdot\left(1-\sigma^*(a_i=1|x_m,T;\theta)\right)^{1\{a_i=0,T\}}$$

subject to

$$\sigma^*(a_i=1|x_m,T;\theta)=\frac{\exp\left(\alpha_i+\beta_i'x_m+\delta_{TT}\sum_{j\neq i}\sigma^*(a_j=1|x_m,T)+\sum_j\delta_{\tilde{T}T}\cdot\sigma^*(a_j=1|x_m,\tilde{T}\neq T)\right)}{1+\exp\left(\alpha_i+\beta_i'x_m+\delta_{TT}\sum_{j\neq i}\sigma^*(a_j=1|x_m,T)+\sum_j\delta_{\tilde{T}T}\cdot\sigma^*(a_j=1|x_m,\tilde{T}\neq T)\right)} \tag{13}$$

Then, using one of the state-of-the-art constrained optimization programs,[21] I obtain the estimates of the parameters of the model.

Brief comparison with alternative estimation methods

Other methods of estimating discrete games of incomplete information have been proposed in the literature.

The standard nested fixed-point algorithms, like the ones used by Seim (2006), Orhun (2005) and Zhu and Singh (2005), require the repeated solution of the model for each trial value of the parameters to be estimated. Explicitly solving for the equilibrium may be computationally burdensome for complex models. Furthermore, the (possible) multiplicity of fixed points may render this method impractical.

To address some of the issues associated with nested-fixed point methods, a two-step pseudo maximum likelihood (PML) approach (see, for example, Aguirregabiria and Mira (2002), Bajari et al (2006) and Ellickson and Misra, 2006) has been proposed. In this approach, instead of solving the system of beliefs in (13) for each parameter vector, the researcher assumes that, in the data, only one of possible many equilibria is actually being played by the agents. So, by observing the entry behavior across a large number of markets, the econometrician can consistently estimate the conditional choice probabilities $\hat{\sigma}_i(a_i=1|x_m,T)$ that appear on the right hand side of equation (13), using non-parametric methods and looking at the firms' observed choices. These conditional choice probabilities are then plugged into a pseudo-likelihood function and the parameters are determined by standard maximization. If the pseudo-likelihood function is based on a consistent nonparametric estimator of the firms' beliefs, the two-step estimator is consistent and asymptotically normal.

The PML estimator is computationally simple but has several important limitations. First, this estimator requires a non-parametric first-stage which is infeasible in most empirical contexts. For

[21]To obtain the parameter estimates shown in Section 6,I submitted the estimation code, written in AMPL (A Modeling Language for Mathematical Programming), to the NEOS server and used the KNITRO and SNOPT optimization solvers.

instance, the shopping center game I consider has over half a dozen market characteristics. The first stage would then require estimating a non-parametric entry strategy function for each firm in over half a dozen dimensions. Second, the nonparametric estimator obtained in the first-step can be very imprecise in small samples generating serious biases in the estimated structural parameters. Last, the PML estimator is a limited information estimator. Thus, it is less efficient than a full information method where all the parameters are estimated simultaneously.

Another method is the nested pseudo likelihood (NPL) method proposed by Aguirregabiria and Mira (2007) which is basically a recursive extension of the two-step PML estimator.

The NPL estimator is more efficient than the two-step estimators and it does not require that one starts the algorithm with a consistent estimator of the choice probabilities. However, it has some important limitations. First, since the iterative procedure is performed using a method of successive approximations, this method is not sure to converge to the equilibrium that generated the data when this equilibrium is *unstable*. Also, to deal with multiple equilibria issues, the authors claim that the researcher should initiate the NPL algorithm with different guesses of the firms' choice probabilities such that, if different fixed-points are obtained, the one which maximizes the likelihood function is chosen. Unfortunately, this assumes that, if there are multiple fixed-points, the researcher is able to recover all of them, which may not be possible altogether.

Conceptually, the MPEC approach that I follow is not more complicated than the above methods. By incorporating the constraints directly in the maximization problem, the researcher can proceed in only one step which makes this method capable of producing efficient estimates (as opposed to the above mentioned two-step methods). The MPEC estimator can be shown to be consistent, asymptotically normal and efficient.

6 Results

Table 3, Panels A, B and C, reports the estimates of the structural parameters of the profit functions in (5). Panel A reports the estimates of the parameters associated with the continuous demographic market variables for each store, Panel B lists the parameter estimates associated with the dummy demographic variables and Panel C displays the parameter estimates for the strategic effects.

6.1 The Role of Demographics and Store Specific Characteristics

The intercept in Table 3, Panel A is negative for all stores. Thus, at the median level of the demographic variables, at the omitted categories of all dummy variables, none of the anchor stores has incentives to enter on its own.

The purchasing power of the area surrounding each shopping center (given by the variable $pop \times income$) has, in general, a positive effect for upscale stores and a negative effect for other stores like Sears and JCPenney. This reflects the fact that, all else equal, discount and midscale department stores usually cater to middle- to lower-income buyers who are more cost conscious and buy lesser-quality merchandise at a discount because they may not be able to afford better brands on a regular basis. Interestingly, the purchasing power estimated parameter for Target does not come out statistically significant, agreeing with the common perception that Target, when compared with other discounters, does not exhibit a clear preference for lower income customers.

The coefficient estimates associated with the variables age, average size of the household and percentage of single family houses, do not exhibit a clear pattern across nor within store types, except for a couple of particular cases. For example, department discount stores seem to prefer to enter in older markets and upscale department stores seem to prefer markets with smaller household sizes. Overall, the parameter estimates associated with these variables are statistically significant which suggests that these variables capture relevant information about the type of consumers of each store.

The variable *parking* is the number of parking spaces at each mall. This variable is used as a proxy for the overall size of the shopping center. Usually the number of parking spaces in a shopping center is determined by industry standards[22] as a function of the area available for construction of the mall. All stores, with the exception of Target and other discount department stores (which include stores like Walmart, K-Mart and Big Lots), seem to prefer to anchor larger centers. This could be because, since these stores are mostly one-stop destination stores (i.e. stores to which consumers make a special trip with the intent of shopping) they do not benefit as much as stores that are more likely to be seen as "impulse" or "backup"[23] stores from being located in larger malls.

The last two variables in Table 3, Panel A, are store specific variables: store square footage and number of stores under the same brand name in the same state. Focusing on the coefficients that are statistically significant, the results are intuitive. Stores make more profits the bigger they are. This makes sense since (a) larger stores tend to be more visible and hence attract more traffic; and (b) larger stores tend to have lower rents (per sqft)[24], implying that this fact together with other potential economies of scale can drive profits up. The # *State Stores* variable is, as expected, a useful predictor for most stores' entering decisions. The coefficient associated with this variable is positive for all stores suggesting that, coeteris paribus, stores enjoy reduced costs (e.g. logistics- and distribution-related costs) from locating in states where their brand has a more significant presence.

Table 3, Panel B, lists the estimated coefficients for the dummy demographic variables associated with the opening date of the mall, the mall developer and the mall location in the US. The customer draw of a shopping center can be affected by the physical aspects of the center. To control for the

[22]For more details on these standards please refer to the book published by the Urban Land Institute entitled "Parking Requirements for Shopping Centers".

[23]Conversations with mall managers indicate that consumers are likely to see a department store as a "backup" store. What is meant is that, while consumers are attracted to a large mall that allows for comparison shopping, they also like the guarantee that if they fail to find something that suits their taste in a smaller retailer, at least they can use the mall's department store as a "backup" store. Given the established brand names of some department stores, consumers know more about the prices and characteristics of the products sold in these stores than they do about products sold at smaller mall retailers. Konishi and Sandfort (2002) have an excellent discussion on these effects.

[24]Several studies support this assertion. For more details, refer, for example, to Wheaton (2000).

fact that older centers usually suffer physical neglect and have older facilities the categorical variable *Date* was included in the estimation. As expected, stores in centers constructed more recently seem to have, all else equal, larger profits than those located in older centers. From Panel B, we can also see that the impact of the variable *Date* varies across stores. This reflects the fact that some stores (e.g. JCPenney) have a longer history of presence in shopping centers than other stores (such as Target, for example). The mall developer variables are included to account for the fact that some developers seem to prefer different types of anchors. Larger and more successful developers such as General Growth Properties (GGP) and the Simon Property Group (SPG) seem to have a stronger preference for more upscale department stores and for historical brands like Sears and JCPenney. This reflects two characteristics of the industry: that some developers create long-term relationships with favored anchors; and that more upscale anchor stores prefer to enter in malls built and managed by bigger developers since this seems to lessen the risks of the developments and guarantee higher expected profits. The coefficient for the dummy variable *West*, indicating whether the market/mall is located in one of the western states, is, as expected, especially high and positive for the store Mervyn's which has been founded in California and tends to have a more significant presence on the West coast and thus a higher preference for malls there located.

6.2 Strategic Effects

By estimating a structural model of strategic entry, we can determine if stores are strategic complements or strategic substitutes. Table 3, Panel C reports the parameter estimates associated with the stores' strategic effects. The pattern of these estimates is complex. Considering the effects of the stores of a given type on a store of the same type (i.e. Panel C's main diagonal), we find that mid-scale department stores are strategic complements. The opposite is true for upscale department stores, i.e. upscale department stores seem to be strategic substitutes.

The pattern of the cross-effects reveals that upscale department stores appear to benefit considerably from the presence of mid-scale department stores but that the reverse is not true. This finding is consistent with practitioner beliefs that many shoppers in mid-scale department stores would be willing to trade-up to a higher end store. But the reverse is seldom true. This finding is reminiscent of the Blattberg and Wisniewski (1989) results for store brands whereby cross-price elasticities between store brands and national brands are found to be on price tiers where for example, cross-price elasticities between store brands and national brands are found to be asymmetric with brands in a higher tier (national brands) drawing sales from brands in lower price tiers (store brands) but not vice versa. Finally, the results also reveal that department discount stores do not appear to have positive effects in any of the other store types. Intuitively this makes sense since, as discussed before, discounters are most times one-stop destination stores which implies that they are not likely to generate much traffic for other stores. For the same reason, they are not likely to benefit from the traffic generated by other anchor stores either.

6.3 Profit Decomposition

The parameter estimates also allow us to determine the importance of each source of profits to the overall profits of each store. I decompose the profits of each store into three sources: (i) demographics; (ii) spillovers from stores of the same type; and (iii) spillovers from stores of other types. The results of this decomposition are presented in Table 4. The demographic variables, which represent the influence of exogenous (with respect to firms' entry) demand factors, have a first order effect on average store profits. In some cases, the effect of the market demographics can be large enough to outweigh the negative or positive effects across firms. The magnitude of the competitive effects from stores of the same type and of different types is also important. The negative competitive effects from stores of different types seems to be an important determinant of long-run profits for department discount stores and midscale department stores. In contrast, the negative effects from stores of the same type is a particularly important determinant of long-run profits for upscale department stores.

7 Conclusion

[TO BE COMPLETED]

References

[1] Aguirregabiria, V. and P. Mira (2002), "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70(4), 1519–1543.

[2] Aguirregabiria, V. and P. Mira (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75(1), 1–53.

[3] Amemiya, T. (1974),"Multivariate Regression and Simultaneous Equation Models when the Dependent Variables Are Truncated Normal," *Econometrica*, 42(6), 999–1012

[4] Aradillas-Lopez, A. (2005), "Semiparametric Estimation of a Simultaneous Game with Incomplete Information," manuscript, Princeton University.

[5] Bajari, P., H. Hong, J. Krainer and D. Nekipolov (2006),"Estimating Static Models of Strategic Interactions," working paper, Duke University.

[6] Berry, S. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889–917.

[7] Beyard, M. and W. O'Mara (1999), "Shopping Center Development Handbook," Urban Land Institute, 3rd edition.

[8] Blattberg, Robert C. and K. J. Wisniewski (1989),"Price-Induced Patterns of Competition," *Marketing Science*, 8 (Fall), 291–309.

[9] Bresnahan, T. and P. Reiss (1990),"Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

[10] Bresnahan, T. and P. Reiss (1991),"Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57–81.

[11] Bresnahan, T. and P. Reiss (1991),"Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009.

[12] Chintagunta, P.K., T. Erdem, P. Rossi and M. Wedel, "Structural Modeling in Marketing: Review and Assessment," *Marketing Science*, forthcoming.

[13] Czyzyk, J., M. Mesnier and J. Moré (1998) "The NEOS Server", *IEEE Journal on Computational Science and Engineering*, 5, 68–75.

[14] Dolan, E. (2001), "The NEOS Server 4.0 Administrative Guide, Technical Memorandum ANL/MCS-TM-250," Mathematics and Computer Science Division, Argonne National Laboratory.

[15] Dudey, M. (1990) "Competition by Choice: The Effect of Consumer Search on Firm Location Decisions," *American Economic Review*, American Economic Association, vol. 80(5), 1092–1104.

[16] Eaton, B.C. and R. G. Lipsey (1979) "Comparison Shopping and the Clustering of Homogeneous Firms", *Journal of Regional Science*, 19, 421–435.

[17] Eaton, B.C. and R. G. Lipsey (1982) "An Economic Theory of Central Places," *Economic Journal*, 92, 56–72.

[18] Ellickson, P. and S. Misra (2006) "Supermarket Pricing Strategies," working paper, Duke University.

[19] Ellison, G. and E. L. Glaeser (1997) "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105, 889–927.

[20] Fourer, R., D. M. Gay and B. W. Kernighan (2002) "AMPL: A Modeling Language for Mathematical Programming," Pacific Grove: Duxbury Press/Brooks/Cole Publishing Company.

[21] Fujita, M. and J. Thisse (1996) "Economics of agglomeration," *Journal of the Japan and International Economies*, 10, 339–378.

[22] Gentzkow, M. (2007) "Valuing New Goods in a Model with Complementarities: Online Newspapers," *American Economic Review*, 97, 713–742.

[23] Gould, E., B. Pashigian and C. Prendergast (2002) "Contracts, Externalities and Incentives in Shopping Malls," CEPR Discussion Papers 3598, C.E.P.R. Discussion Papers.

[24] Gropp, W. and J. Moré (1997) "Optimization Environments and the NEOS Server," *Approximation Theory and Optimization*, M. D. Buhmann and A. Iserles, eds., Cambridge University Press, 167–182.

[25] Heckman, J. (1978) "Dummy Endogenous Variables in a Simultaneous Equation System," *Econonometrica*, 46, 931–959.

[26] Judd, K. L. (1998) "Numerical Methods in Economics," Cambridge, MA: MIT Press.

[27] Judd, K. L. and C.L. Su (2006) "A New Optimization Approach to Maximum Likelihood Estimation of Structural Models," working paper, Hoover Institution.

[28] Konishi, H. and Sandfort, M. T. (2003) "Anchor stores," *Journal of Urban Economics*, 53(3), 413–435.

[29] Krugman, P. (1991) "Increasing Returns and Economic Geography," *Journal of Political Economy*, Vol. 99, no. 3: 483–499.

[30] Mazzeo, M. (2002) "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33(2), 1–21.

[31] McKelvey, R. and T. Palfrey (1995) "Games and Economic Behavior," 10, 6–38.

[32] Myrdal, G. (1957) "Economic Theory and Underdeveloped Regions," London: Duckworth.

[33] Orhun, Y. (2005) "Spatial Differentiation in the Supermarket Industry," working paper, University of California at Berkeley.

[34] Pashigan, B. and E. Gould (1998) "Internalizing Externalities: The Pricing of Space in Shopping Centers," *Journal of Law and Economics* XLI: 1:115–142.

[35] Porter, M. (1998) "Clusters and the New Economy," *Harvard Business Review.*

[36] Radner, R. and R.W. Rosenthal (1982) "Private information and pure-strategy equilibria," *Mathematics of Operations Research*, 7, 401–409.

[37] Seim, K. (2000) "Essays on Spatial Product Differentiation," Discussion paper, Yale University PhD Dissertation.

[38] Seim, K. (2006) "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, forthcoming.

[39] Spector, R. (2005) "Category Killers: The Retail Revolution and Its Impact on Consumer Culture," *Harvard Business School Press.*

[40] Stahl, K. (1982*a*) "Location and Spatial Pricing Theory with Nonconvex Transportation Cost Schedules," *The Bell Journal of Economics*, 13, 97–113.

[41] Stahl, K. (1982*b*) "Differentiated Products, Consumer Search, and Locational Oligopoly," *Journal of Industrial Economics*, 31, 97–113.

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

[43] Tamer, E. and S. Berry (2006) "Identification in Models of Oligopoly Entry," *Advances in Economics and Econometrics*, Blundell, Newey and Persson, eds., Ninth World Congress, Cambridge University Press, 46–85.

[44] Urban Land Institute (2000) "Parking Requirements for Shopping Centers: Summary Recommendations and Research Study Report," Washington, D.C.: Urban Land Institute.

[45] Vernor, J., and J. Rabianski (1993) "Shopping Center Appraisal and Analysis," Chicago, Il: Appraisal Institute.

[46] Wheaton, W.C. (2000) "Percentage rent in retail leasing: the alignment of landlord-tenant interests," *Real Estate Economics*, 28(2), 185–204.

[47] Wolinsky, A. (1983) "Retail Trade Concentration due to Consumers' Imperfect Information," *The Bell Journal of Economics*, Vol. 14, No. 1, 275–282.

[48] Zhu, T., V. Singh and M. Manusazk (2005) "Market Structure and Competition in the Retail Discount Industry," working paper, Carnegie Mellon University.

[49] Zhu, T. and V. Singh (2005) "Spatial Competition with Endogenous Location Choices: An Application to Discount Retailing," working paper, Carnegie Mellon University.

APPENDIX 1: Tables and Figures

Table 1
Classification of Shopping Centers by Type

ICSC SHOPPING CENTER DEFINITIONS-U.S.

TYPE OF SHOPPING CENTER	CONCEPT	SQUARE FEET (INCLUDING ANCHORS)	ACREAGE	TYPICAL ANCHOR(S)		ANCHOR RATIO*	PRIMARY TRADE AREA**
				NUMBER	TYPE		
MALLS							
Regional Center	General merchandise; fashion (mall, typically enclosed)	400,000-800,000	40-100	2 or more	Full-line department store; jr. department store; mass merchant; discount department store; fashion apparel	50-70%	5-15 miles
Superregional Center	Similar to regional center but has more variety and assortment	800,000+	60-120	3 or more	Full-line department store; jr. department store; mass merchant; fashion apparel	50-70%	5-25 miles
OPEN-AIR CENTERS							
Neighborhood Center	Convenience	30,000-150,000	3-15	1 or more	Supermarket	30-50%	3 miles
Community Center	General merchandise; convenience	100,000-350,000	10-40	2 or more	Discount department store; supermarket; drug; home improvement; large specialty/ discount apparel	40-60%	3-6 miles
Lifestyle Center	Upscale national chain specialty stores; dining and entertainment in outdoor setting.	Typically 150,000-500,000 but can be smaller or larger	10-40	0-2	Not usually anchored in the traditional sense but may include book store; other large-format specialty retailers; multi-plex cinema; small department store.	0-50%	8-12 miles
Power Center	Category-dominant anchors; few small tenants	250,000-600,000	25-80	3 or more	Category killer; home improvement; discount department store; warehouse club; off-price	75-90%	5-10 miles
Theme/Festival Center	Leisure; tourist-oriented; retail and service	80,000-250,000	5-20	N/A	Restaurants; entertainment	N/A	N/A
Outlet Center	Manufacturers' outlet stores	50,000-400,000	10-50	N/A	Manufacturers' outlet stores	N/A	25-75 miles

* The share of a center's total square footage that is attributable to its anchors; ** The area from which 60-80% of the center's sales originate.

Table 2
Descriptive Statistics

This table reports the descriptive statistics of all the variables in the sample of 561 enclosed regional shopping malls. **Panel A** reports the descriptive statistics of the market specific demographics. Pop×income is the product of the population within a 20 mile radius of the mall and the median income level of the same population. Age is the median age of the population within a 20 mile radius of the mall. Size HH is the average size of the household of the population within a 20 mile radius of the mall. Sing Family House is the percentage of houses that are single family house in the same three digit ZIP code of the mall. Parking is the number of parking spaces of the mall. Date 2 and 3 is a dummy variable reflecting the open date of the mall The omitted reference date group is Date 1 (not reported) which is a variable that takes the value of one if the mall was opened between 1952 – 1972. Date 2 is a dummy variable that takes the value of 1 if the mall was opened between 1973 – 1980, and Date 3 is a dummy variable that takes the value of 1 if the mall was opened between 1981 – 2006. Developer GGP, Developer SPG, Other Dev Medium and Other Dev Small are dummy variables reflecting the developer of a given mall. The omitted reference developer group is All Other Developers (not reported) which is a variable that takes the value of one if the developer of the mall is an independent or very small mall developer. Developer GGP takes the value of 1 if the mall developer is General Growth Properties. Developer SPG takes the value of 1 if the mall developer is Simon Property Group. Other Dev Medium takes the value of 1 if the mall developer is a general developer of medium scale. Other Dev Small takes the value of 1 if the mall developer is a general developer of small scale. **Panel B** reports the store specific descriptive statistics for each type of Store in the sample of 561 enclosed regional shopping malls. Sqft is the square foot size of the store in a given mall and # State Stores is the number of stores of the same store across all the malls in a given state. In the two panels, all continuous variables (not dummies) are transformed as follows: $X^{\text{transformed}} = \ln\left(\left(X^{\text{original}}/\text{mean}\left(X^{\text{original}}\right)\right)\right)$ so that each variable has a value of zero at its mean (across malls) value.

Panel A: Market Specific Demographics

Demographics	Mean	Standard Deviation	Minimum	Maximum
Pop×Income	−0.92	1.37	−4.17	2.33
Age	−0.00	0.08	−0.40	0.40
Size HH	−0.00	0.08	−0.33	0.32
Sing Family House	−0.02	0.21	−1.79	0.25
Parking	−0.04	0.30	−1.75	0.65
Date 2	0.32	0.47	0	1
Date 3	0.35	0.48	0	1
Developer GGP	0.17	0.37	0	1
Developer SPG	0.15	0.36	0	1
Other Dev Medium	0.26	0.44	0	1
Other Dev Small	0.19	0.40	0	1

Panel B: Store Specific Variables

Variable	Sears	Target	Other Disc	JCP	Mervyns	Other Mid	Dillards	Macys	Other Up
Number of Stores	370	60	32	334	54	229	155	88	230
Sqft									
mean	−0.05	−0.01	–	−0.06	−0.00	–	−0.03	−0.02	–
std	0.32	0.11	–	0.36	0.05	–	0.25	0.20	–
min	−0.89	−0.48	–	−1.12	−0.33	–	−1.20	−1.19	–
max	1.16	0.45	–	0.92	0.63	–	0.75	0.85	–
# State Stores									
mean	13.47	2.19	–	11.95	2.46	–	5.32	4.30	–
std	8.14	2.63	–	7.12	4.61	–	8.00	7.15	–
min	0	0	–	0	0	–	0	0	–
max	27	9	–	25	15	–	28	25	–

Table 3
Parameter Estimates of the Profit Functions of Each Store

This table reports the parameter estimates of the profit functions of each store using a sample of 561 enclosed regional shopping malls. **Panel A** reports the parameter estimates in the store profit function associated with the mall specific demographic variables and the store specific variables. Pop×income is the product of the population within a 20 miles radius of the mall and the median income level of the same population. Age is the median age of the population within a 20 mile radius of the mall. Size HH is the average size of the household of the population within a 20 mile radius of the mall. Sing Family House is the percentage of houses that are single family house in the same three digit ZIP code of the mall. Parking is the number of parking spaces of the mall. Sqft is the square foot size of the store in a given mall and # State Stores is the number of stores of the same store across all the malls in a given state. **Panel B** reports the parameter estimates in the store profit function associated with the dummy variables. Date 2 and 3 is a dummy variable reflecting the open date of the mall The omitted reference date group is Date 1 (not reported) which is a variable that takes the value of one if the mall was opened between $1952-1972$. Date 2 is a dummy variable that takes the value of 1 if the mall was opened between $1973-1980$, and Date 3 is a dummy variable that takes the value of 1 if the mall was opened between $1981-2006$. Developer GGP, Developer SPG, Other Dev Medium and Other Dev Small are dummy variables reflecting the developer of a given mall. The omitted reference developer group is Other Smaller Developers (not reported) which is a variable that takes the value of one if the developer of the mall is an independent or very small mall developer. Developer GGP takes the value of 1 if the mall developer is General Growth Properties. Developer SPG takes the value of 1 if the mall developer is Simon Property Group. Other Dev Medium takes the value of 1 if the mall developer is a general developer of medium scale. West takes the value of 1 if the Shopping Center is located in one of the Western Stares. **Panel C** reports the parameter estimates of the competitive effects. Each value in the table corresponds to the impact of row store type (i) on the column store type (j). All continuous variables (not dummies) are transformed as follows: $X^{\text{transformed}} = \ln\left(\left(X^{\text{original}}/\text{mean}\left(X^{\text{original}}\right)\right)\right)$ so that each variable has a value of zero at its mean (across malls) value. Values statistically significant at the 5% level are market with an asterisk

PANEL A: Profit Function Parameters Associated with Demographics

	Discount Dep. Stores			**Midscale Dep. Stores**			**Upscale Dep. Stores**		
Variable	Sears	Target	Other Disc	JCP	Mervyns	Other Mid	Dillards	Macys	Other Up
Intercept	−0.03	−2.07*	−1.97*	−1.25*	−3.87*	−0.53	−3.90*	−2.68*	−0.67*
	(.25)	(.42)	(.48)	(.28)	(.34)	(.27)	(.47)	(.41)	(.27)
Pop×Income	−1.01*	0.21	−0.13	−1.07*	0.15	−0.44*	−0.11	0.74*	0.62*
	(.14)	(.12)	(.15)	(.10)	(.11)	(.07)	(.11)	(.11)	(.09)
Age	3.49*	−4.00	8.26*	5.55*	−5.73*	−3.23*	−6.78*	6.28*	−1.12
	(1.0)	(2.1)	(2.2)	(1.1)	(1.5)	(.86)	(1.2)	(1.2)	(.75)
Size HH	6.38*	−5.16*	3.85	7.47*	−2.55	−3.54*	−11.87*	−1.49	−5.32*
	(1.5)	(1.9)	(2.6)	(1.5)	(1.7)	(.88)	(1.5)	(1.2)	(1.0)
Sing Fam House	−1.01*	4.07*	−0.31	−1.21*	5.53*	0.73	2.64*	−0.41	0.48
	(.33)	(.80)	(.78)	(.29)	(.97)	(.37)	(.59)	(.32)	(.25)
Parking	1.95*	0.78	0.10	1.19*	2.79*	−0.11	1.53*	1.51*	2.05*
	(.30)	(.53)	(.47)	(.31)	(.46)	(.25)	(.32)	(.33)	(.26)
Store Sqft	1.94*	−2.12	–	2.45*	−0.15	–	1.43*	−0.54	–
	(.29)	(1.1)	–	(.28)	(1.9)	–	(.34)	(.39)	–
# State Stores	0.03*	0.31*	–	0.01	0.17*	–	0.19*	0.09*	–
	(.01)	(.03)	–	(.01)	(.02)	–	(.02)	(.01)	–

PANEL B: Profit Function Parameters Associated with Dummy variables

	Discount Dep. Stores			Midscale Dep. Stores			Upscale Dep. Stores		
Variable	Sears	Target	Other Disc	JCP	Mervyns	Other Mid	Dillards	Macys	Other Up
Date 2	0.51*	0.30	0.38	0.81*	−0.09	0.06	0.74*	−0.11	0.32*
	(.17)	(.23)	(.36)	(.18)	(.24)	(.17)	(.21)	(.19)	(.13)
Date 3	0.93*	0.90*	0.31	1.22*	0.53*	0.11	1.73*	0.47*	0.72*
	(.20)	(.26)	(.44)	(.21)	(.26)	(.19)	(.20)	(.21)	(.15)
Developer GGP	1.44*	0.13	−0.09	1.28*	0.42	−0.51*	0.80*	1.76*	0.73*
	(.22)	(.26)	(.49)	(.23)	(.32)	(.21)	(.22)	(.25)	(.17)
Developer SPG	1.96*	0.05	−1.19	1.10*	−0.26	0.21	1.15*	2.36*	1.30*
	(.25)	(.40)	(2.8)	(.25)	(.38)	(.22)	(.21)	(.26)	(.17)
Other Dev Medium	1.68*	−0.14	−0.12	1.55*	−0.01	0.52*	0.58*	0.61*	0.85*
	(.23)	(.28)	(.46)	(.21)	(.24)	(.17)	(.20)	(.24)	(.15)
West	−0.21	0.33	−0.79	−0.49*	1.80*	−0.90*	1.34*	0.84*	0.35*
	(.19)	(.28)	(.55)	(.19)	(.26)	(.18)	(.19)	(.25)	(.13)

Panel C: Strategic Effects

	Discount	Mid-Scale	Upscale
Discount	0.68	−0.07	−0.93*
	(0.53)	(0.28)	(0.36)
Mid-Scale	−0.41	0.80*	1.18*
	(0.28)	(0.23)	(0.27)
Upscale	−1.56*	−0.98*	−1.33*
	(0.35)	(0.23)	(0.16)

Table 4
Decomposition of Average Total Profits of Each Store by Source of Profits

This table reports the decomposition of total average profits of each store by source of profits. Average Total Profits is the in sample average realized profits, Demographics is the realized average profits that can be attributed to demographics characteristics of the market, Competitive Same Type is the realized average profits that can be attributed to the entry of stores of the same type, and Competitive Other Type is the realized average profits that can be attributed to the entry of stores of other types.

	Average Total Profits	Demographics	Competitive Same Type	Competive Other Types
Sears	0.85	2.51	0.11	−1.77
Target	−2.47	−1.18	0.48	−1.77
Other Discount	−3.30	−2.05	0.52	−1.77
JCP	0.55	1.03	0.41	−0.89
Mervyns	−3.30	−3.22	0.81	−0.89
Oth Midscale	−0.47	−0.14	0.56	−0.89
Dillards	−1.53	−1.30	−0.75	0.53
Macys	−2.31	−1.93	−0.91	0.53
Oth Upscale	−0.39	−0.34	−0.58	0.53

APPENDIX 2: Identification: 2-Players example

In a market characterized by $X = x$ and with two potential entrants ($i = 1, 2$), let $\sigma_i = \sigma_i\left(a_i = 1|x\right)$. The system in (12) can then be rewritten as

$$\begin{cases} E[\Pi_1\left(a_1 = 1 \mid x, \theta\right)] = \Pi_1\left(a_1 = 1, a_2 = 1 \mid x, \theta\right) \cdot \sigma_2 + \Pi_1\left(a_1 = 1, a_2 = 0 \mid x; \theta\right) \cdot (1 - \sigma_2) \\ E[\Pi_1\left(a_1 = 0 \mid x, \theta\right)] = \Pi_1\left(a_1 = 0, a_2 = 1 \mid x, \theta\right) \cdot \sigma_2 + \Pi_1\left(a_1 = 0, a_2 = 0 \mid x, \theta\right) \cdot (1 - \sigma_2) \\ E[\Pi_2\left(a_2 = 1 \mid x, \theta\right)] = \Pi_2\left(a_1 = 1, a_2 = 1 \mid x, \theta\right) \cdot \sigma_1 + \Pi_1\left(a_1 = 0, a_2 = 1 \mid x, \theta\right) \cdot (1 - \sigma_1) \\ E[\Pi_2\left(a_2 = 0 \mid x, \theta\right)] = \Pi_2\left(a_1 = 1, a_2 = 0 \mid x, \theta\right) \cdot \sigma_1 + \Pi_1\left(a_1 = 0, a_2 = 0 \mid x, \theta\right) \cdot (1 - \sigma_1) \end{cases} \tag{1}$$

By Assumption (2), the second and last equations in (1) can be omitted, so we end up with a system with 2 equations and 4 unknowns. It is clear that, without any further restrictions, one cannot non-parametrically identify the $\Pi(\cdot)$ terms using exclusively the values for $E[\Pi_i]$ and σ_i.

Now suppose X is a variable that may take different values for different firms. . For example, assume that the variable X can take only one of two values $\{H, L\}$ (*high* or *low*). In this case, and making use of Assumption (2), the system in (12) can be written as

$$\begin{cases} E[\Pi_1\left(a_1 = 1 \mid X_1 = H, X_2 = H; \theta\right)] = & \Pi_1\left(a_1 = 1, a_2 = 1 \mid X_1 = H, X_2 = H; \theta\right) \cdot \sigma_2 \\ & +\Pi_1\left(a_1 = 1, a_2 = 0 \mid X_1 = H, X_2 = H; \theta\right) \cdot (1 - \sigma_2) \\ E[\Pi_1\left(a_1 = 1 \mid X_1 = H, X_2 = L; \theta\right)] = & \Pi_1\left(a_1 = 1, a_2 = 1 \mid X_1 = H, X_2 = L; \theta\right) \cdot \sigma_2 \\ & +\Pi_1\left(a_1 = 1, a_2 = 0 \mid X_1 = H, X_2 = L; \theta\right) \cdot (1 - \sigma_2) \\ \cdots & \cdots \\ E[\Pi_2\left(a_2 = 1 \mid X_1 = H, X_2 = H; \theta\right)] = & \Pi_2\left(a_1 = 1, a_2 = 1 \mid X_1 = H, X_2 = H; \theta\right) \cdot \sigma_1 \\ & +\Pi_1\left(a_1 = 0, a_2 = 1 \mid X_1 = H, X_2 = H; \theta\right) \cdot (1 - \sigma_1) \\ \cdots & \cdots \end{cases} \tag{2}$$

The system in (2) has 8 equations and 16 unknowns, so it is not identified. Basically, by allowing X to be not only market specific but also firm specific, both the number of equations and unknowns increased.

But, if we have (ex-ante) reason to believe that each firm's profits are not effected by the value that X takes for the other firm, then we can impose exclusion restrictions that will make the system identified. This is because, while the number of equations will remain the same as in (2), the number of unknowns will decrease considerably since now $\Pi_i\left(a_i, a_j \mid X_i = x_i, X_j = x_j; \theta\right) = \Pi_i\left(a_i, a_j \mid X_i = x_i, X_j = x_{j'}; \theta\right)\ \forall x_j \neq x_{j'}\ \forall i \neq j$. With 8 equations and 8 unknowns in the system we can now achieve non-parametric identification of the deterministic part of profits.

The reasoning used in this example can be easily generalized to accommodate more players and to allow for continuous X variables.