

MARIA ANA VITORINO*

The author proposes a strategic model of entry that allows for positive and negative spillovers among firms. The model is applied to a novel data set containing information about the store configurations of all U.S. regional shopping centers and is used to quantify the magnitude of interstore spillovers. The author addresses the estimation difficulties that arise due to the presence of multiple equilibria by formulating the entry game as a mathematical problem with equilibrium constraints (MPEC). Although this study constitutes the first attempt to use this direct optimization approach to address a specific empirical problem, the method can be used in a wide range of structural estimation problems. The empirical results support the agglomeration and clustering theories that predict that firms may have incentives to collocate despite potential business stealing effects. The author shows that the firms' negative and positive strategic effects help predict both how many firms can operate profitably in a given market and the firm-type configurations. The relative magnitude of such effects varies substantially across store types.

Keywords: entry, spillovers, shopping centers, incomplete information game, direct optimization approach, mathematical programming with equilibrium constraints (MPEC)

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

*Maria Ana Vitorino is an Assistant Professor, The Wharton School, University of Pennsylvania (e-mail: vitorino@wharton.upenn.edu). This article is based on the author's doctoral dissertation at the University of Chicago. The author is indebted to dissertation advisors Jean-Pierre Dubé, Pradeep Chintagunta, Kenneth Judd, Peter Rossi, and Matthew Gentzkow for their guidance and support. She has greatly benefited from discussions with Frederico Belo, Hongju Liu, Jonah Berger, Alon Eizenberg, Jeremy Fox, Yeşim Orhun, Katja Seim, Upender Subramanian, and Ting Zhu. For their helpful comments, the author thanks Peter Reiss; Günter Hitsch; Oleg Urminsky; seminar participants at the University of Chicago, Emory University, CMU, University of Rochester, University of California, Berkeley, University of California, Los Angeles, New York University, University of Pennsylvania, University of North Carolina—Chapel Hill, University of Iowa, University of Toronto, HKUST, University of Texas at Dallas, and QME 2007 Conference; and the anonymous *JMR* reviewers. The author also thanks Che-Lin Su, Todd Plantenga, David Gay, and Robert Fourer for providing help with AMPL and KNITRO and Hector Castro for excellent research assistance. Julie Cameron and Robert Michaels from General Growth Properties and Mike Tubridy from the ICSC shared their insights on the shopping center industry, and Bob Galvin from the NRB made available part of the shopping center data. The author gratefully acknowledges the financial support from the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology), the Calouste Gulbenkian Foundation, and the Kilts Center for Marketing at the University of Chicago. Yuxin Chen served as associate editor for this article.

A common characteristic of the U.S. 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 substitutable competitors. For example, antique and jewelry stores are often found in close proximity. The same phenomenon arises for automobile dealerships, often collectively termed "auto malls."

Firms locate near one another because doing so lowers consumer search costs and increases aggregate demand for these firms. Such positive spillovers offset the main cost from joining a cluster, which is that the close proximity of firms selling similar goods intensifies competition (for a review of the literature on the economics of firms' agglomeration, see Fujita and Thisse 1996).

This article examines 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. The marketing literature has largely overlooked this industry, despite its importance. At the same

time, while there have been some (mostly) descriptive studies of shopping centers in the real estate literature, there has 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 the collocation of both complementary and substitutable retailers can be observed. The current study focuses on the entry decisions of anchor stores (traffic-generating stores) because these firms typically commit to a mall before smaller retailers. Most regional shopping centers have multiple anchor stores that are highly substitutable competitors. Typically, these are competing department stores. A simple possible explanation for the coincidence of many highly competitive anchor stores is that the profit potential of the mall is sufficiently high to warrant head-to-head competition. However, interactions with mall developers indicate that another motive may arise from strategic complementarities: Anchor stores generate traffic spillovers among themselves and to other retailers. Naturally, how much traffic each anchor store generates might depend on its particular identity and also on the interaction with the other anchors in the mall.

Using a novel data set containing information about the store configurations of all U.S. regional shopping centers, this study investigates 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.

Thus, I formalize the observed configuration of anchor stores in a mall as the equilibrium outcome of a simultaneous moves entry game of incomplete information in which each store's entry decision is affected by both the direct and the strategic effects. This framework is convenient for several reasons. First, using a game theoretic model, I can accommodate the simultaneity of firms' entry decisions. Second, the incomplete information nature of the game allows estimation of a more realistic model with a relatively large number of players. Last, this approach does not require any price or quantity data, because market structure alone is enough to extract relevant information about firms' profits.

While there is a well-established body of literature on empirical entry games, the extant work typically assumes away one of the types of spillovers studied here. Studies in this literature usually impose the restriction that the entry of an additional firm into the market always decreases incumbents' profits (e.g., Berry 1992; Bresnahan and Reiss 1990; Mazzeo 2002; Seim 2006). Furthermore, this decrease is assumed to be larger if the new market participant is of the same type as (or located close by) 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 realistic.

This negative-spillover assumption is often made for econometric identification purposes because 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., Orhun 2005; Seim 2006; Zhu and Singh 2009), it is difficult or infeasible to construct a likelihood function in the presence of multiple equilibria. To circumvent this problem, researchers have tried to guarantee equilibrium uniqueness by either imposing additional structure on their models (e.g., Mazzeo [2002] assumes firms make sequential entry decisions) and/or by adding restrictions to the model parameters (e.g., Seim 2006). Alternatively, others have used two-step methods (e.g., Bajari et al. 2010; Ellickson and Misra 2008) that are computationally simple. However, these latter approaches do not utilize all the information from the model. They also require a nonparametric first stage, which is infeasible in most empirical contexts.

This research develops 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 Su and Judd (2008). 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 estimate the model using maximum likelihood without having to restrict a priori the domain of the parameters while also allowing for the presence of multiple equilibria. Although this study is 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.¹

In addition to resolving the computational problems associated with multiple equilibria, I also address some of the standard econometric identification problems from games. Several plausible exclusion restrictions are explored that help resolve identification for the shopping mall entry game studied.

Two main conclusions can 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 that firms may have incentives to collocate 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 help predict both how many firms can operate profitably in a given market and the firm-type configurations. In some cases, the strategic effects can be large enough to outweigh the effect of market demographics. The relative magnitude of such effects varies substantially across store types. For example, the data reveal that market- and store-specific characteristics are the main

¹In subsequent work to this research, Dubé, Fox, and Su (2009) show that it is possible to improve the numerical performance of Berry, Levinsohn, and Pakes' random coefficients demand estimation by recasting it as an MPEC. In addition, Musalem and Woochoel (2009) use MPEC to estimate a structural model of product line competition.

determinant of long-term profits for discount department stores (e.g., Target). However, this is not true for midscale department stores (e.g., JCPenney) for which the negative strategic competitive effects (from other anchors of different type) and the positive spillovers from same-type stores can be the main determinants of long-term profits.

DATA

Data Sources and Market Definition

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 are available but only for relatively small mall samples (e.g., Gould, Pashigian, and Prendergast [2005] make use of detailed data for more than 2500 stores in approximately 35 large malls across the United States), which restricts the potential areas to be studied in this industry.

To study the determinants of shopping center configuration and measure stores' strategic effects without any price or quantity data, empirical models of entry and market structure are especially convenient, because these are primarily based on the observed number and types of stores in a given market, thus requiring few data.

In this article, a market is defined as a regional shopping center. This is a natural definition because a regional shopping center is a self-contained shopping hub that provides general merchandise and services in full depth and variety, serving a large trading area of 15–20 miles. Other types of shopping centers were not included in the analysis (for a brief description of each type of shopping center as defined by the International Council of Shopping Centers, see Table 1) to make the markets in the sample studied relatively homogeneous with respect to the types of variables that influence the stores' entry decisions and to the unobservable market characteristics. In turn, this allows isolation and measurement of the spillover effects across stores.

I focus on the entry decisions of anchor stores in such markets. Anchor stores are usually major chain stores that occupy the largest spaces in a center (in total, approximately 50%–70% of the gross rentable floor area in a regional shopping center) and serve as the primary traffic generators. Anchor stores commit to a mall before smaller retailers, and as a rule, a regional shopping center is not built until the developer has signed leases (of length typically greater than 20 years) from the anchor stores.

Data from the 2006 *Directory of Major Malls* provided information for a cross-section of regional shopping centers operating in the United States. This directory includes, for each of 647 regional shopping centers (as defined previously), not only the identity of the anchor tenants but also several store and mall characteristics. Additional demographic data were obtained from the 2000 census for each mall's three-digit zip code. Three-digit zip codes average approximately 1000 square miles of overall land and water area. If they were circular, they would have mean radius of 15.1 miles. As mentioned previously, this constitutes a reasonable radius for a regional shopping center trade area. In

addition, I collected information on the location of each anchor's headquarters and distribution centers and computed the distance of each shopping center to each anchor's headquarters site and its nearest distribution center.

The data for this selected sample of malls were cross-checked and complemented with data from another directory, *The Shopping Center Directory* (National Research Bureau). When 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 centers with too many missing variables were eliminated, the final sample consisted of 561 regional shopping centers. Anchor stores for each mall were classified into several broad categories/types according to their retail model and format. I examine the existence of interfirm 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 eliminated, thus restricting attention to three main types of anchors, which constitute approximately 80% of the total number of anchors in the sample. The three types of anchors studied are broadly defined as follows:

- *Upscale department stores*: These stores 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 in each department. Examples include Dillard's, Macy's, and Nordstrom.
- *Midscale department stores*: These stores sell brand names and non-brand names but do not sell upscale brand names. Compared with upscale department stores, midscale stores usually do not have perfumes and beauty supplies at the main entrance and do not 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. Most discount department stores offer wide assortments of goods; others specialize in merchandise such 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 among products. Compared with midscale department stores, discounters sell fewer major brand names and offer a wider variety of products. Stores in the discount department store category typically have fewer sales workers, relying more on self-service features, and have centrally located cashiers.

Preliminary Data Analysis: Colocation Patterns

Preliminary evidence from the raw data shows that clustering seems to play a role in the configuration of shopping centers. According to Figure 1, the number of anchor stores varies considerably across malls. Most regional shopping centers in the sample have three or four anchors, but there are malls in the data that have up to eight anchors. Figure 2 indicates that midscale department stores tend to cluster more than other types of department stores. Approximately 50% of the malls in the sample that have this type of department store have at least two midscale department stores, whereas upscale department stores, for example, seem to avoid colocating with other stores of the same type.

Table 1
CLASSIFICATION OF SHOPPING CENTERS BY TYPE

Type of Shopping Center	Concept	Square Feet (Including Anchors)	Acreage	Typical Anchors		Anchor Ratio ^a	Primary Trade Area ^b
				Number	Type		
<i>Malls</i>							
Regional center	General merchandise, fashion (mall, typically enclosed)	400,000–800,000	40–100	Two or more	Full-line department store, junior 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	Three or more	Full-line department store, junior department store, mass merchant, fashion apparel	50%–70%	5–25 miles
<i>Open-Air Centers</i>							
Neighborhood center	Convenience	30,000–150,000	3–15	One or more	Supermarket	30%–50%	3 miles
Community center	General merchandise, convenience	100,000–350,000	10–40	Two 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	Zero to two	Not usually anchored in the traditional sense but may include book stores, other large-format specialty retailers, multiplex cinemas, small department stores.	0%–50%	8–12 miles
Power center	Category-dominant anchors, few small tenants	250,000–600,000	25–80	Three 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

^aThe share of a center's total square footage attributable to its anchors.

^bThe area from which 60%–80% of the center's sales originate.

Notes: 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. N.A. = not applicable.

Figure 1
NUMBER OF ANCHOR STORES ACROSS MALLS

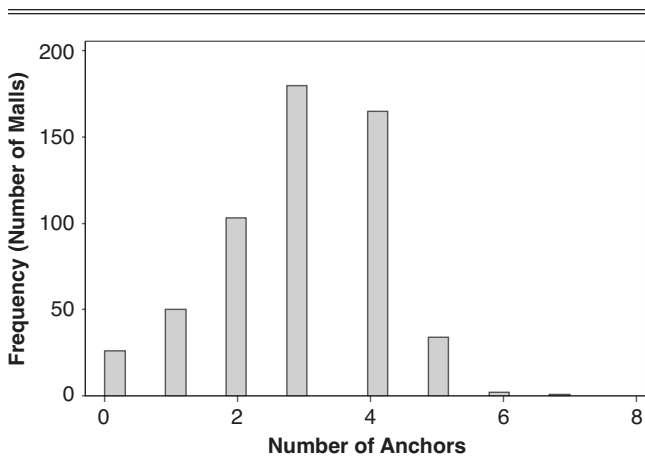


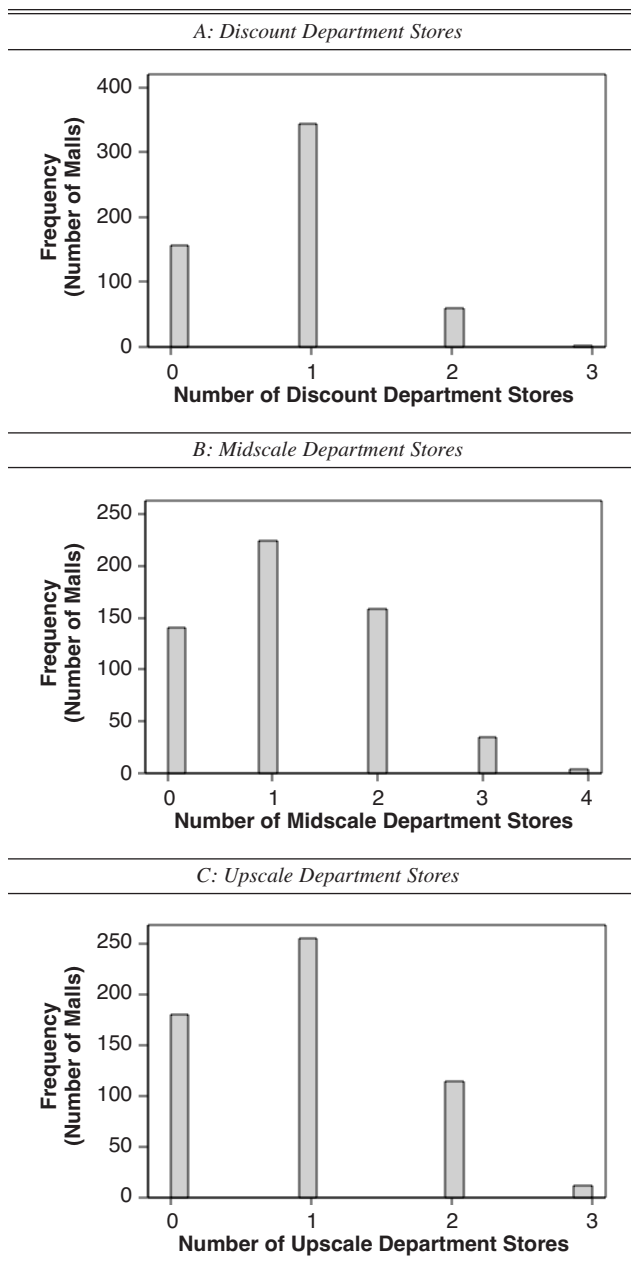
Table 2 reports the most observed anchor configurations across the malls in the sample. The most frequent configuration is seen in almost 9% of the malls in the sample and consists of a Sears, a JCPenney, and an additional midscale department store (other than Sears or JCPenney). As the table shows, it is not easy to identify clear patterns in the observed mall configurations because they vary substantially across malls. One striking feature, though, is that approximately 50% of the malls in the sample include a Sears and a JCPenney. There may be several reasons these two stores tend to collocate. They may benefit from the same market characteristics, or there may be positive spillovers between these two stores.

There are several mechanisms that explain why positive spillovers may exist. First, consider demand-driven explanations. Shopping involves consumer search costs: Consumers go to a mall because they know they can sample similar or complementary goods from several stores before making a purchase. This may be the case whether consumers are searching over product attributes (Stahl 1982) or low prices (Dudey 1990). Agglomeration of certain types of stores can then create heightened demand: The number of consumers will increase beyond the sum that the stores might have otherwise attracted when located separately because consumers are more likely to find the products they prefer at the prices they are willing to pay.

Second, consider supply-side reasons stores benefit from collocating. Shopping center tenants share their obligations in the provision of high-quality public services and facilities. By sharing such costs, tenants obtain the collective benefits of higher-quantity and -quality services and facilities to draw more customers to the shopping center. Closeness to competitors can also be an important source of up-to-date information about competitors and consumers because it makes it easier to monitor competitors' demand, sales, product assortment, and so on.

Last, market-level intensity of competition may be reduced if the number of markets in which firms compete is significantly large. In the specific context of anchor stores, the simultaneous presence of stores across different markets

Figure 2
NUMBER OF ANCHOR STORES PER TYPE ACROSS MALLS



increases the opportunity for retaliation against competitive attacks in the same way that the threat of retaliation in future interactions in a same market works. If indeed there is reduced rivalry between some stores due to expected long-term interactions (because mall anchors sign long-term contracts typically with duration of 25 or more years), this will also be reflected in the estimated strategic effects.

Although the preceding descriptives are useful as a first cut at the data, to understand the stores' collocation patterns, a model is necessary to quantify the store spillovers and investigate their direction while controlling for market characteristics (e.g., demographics). This is the purpose of the model described in the next section.

Table 2
MOST FREQUENTLY OBSERVED ANCHOR COMBINATIONS

Discount Department Stores			Midscale Department Stores			Upscale Department Stores			% of Malls
Sears	Target	Other	JCPenney	Mervyn's	Other	Dillard's	Macy's	Other	
1	0	0	1	0	1	0	0	0	8.91
1	0	0	1	0	0	0	0	1	5.7
1	0	0	1	0	1	0	0	1	4.99
0	0	0	0	0	0	0	0	0	4.63
1	0	0	1	0	1	1	0	0	4.63
1	0	0	1	0	0	1	0	0	4.46
...	<4.45

A MODEL OF ENTRY IN A CLUSTER

Model Setup

I model the entry of firms (i.e., anchor stores) in a market or retail cluster (i.e., a regional shopping center) as a simultaneous-move entry game of incomplete information. The assumptions behind this type of game deserve some explanation. First, although it is true that shopping centers are to some extent planned, it is also true that mall planners cannot freely determine which stores will enter in a given market. Thus, to specify how retailers enter the mall, two competing and extreme approaches can be taken. The first assumes that the mall planner has full control over the choice of stores; in this case, the composition of the mall can be modeled as a discrete choice problem, in which the mall planner chooses the tenant mix that maximizes the sum of expected profits of the retailers. However, it is not realistic to assume that the mall planner can choose which anchor stores to have in a mall: Anchor stores have significant power when deciding which malls to enter. Furthermore, given the large number of possible store combinations, this approach is not feasible in practice. The second extreme approach follows the standard entry literature and assumes that potential retailers make their decisions to enter the market solely on the basis of expected profits and taking into account the actions of the other retailers, with no intervention by the mall planner. Similarly, this is also unrealistic, because the mall developer has influence on the shopping center's configuration. The current study follows an intermediate approach, which is more reflective of reality.

The model assumes that a mall developer, having previously defined the location of the mall, approaches a set of potential entrants and provides 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. Because the game may have many possible equilibria, it is assumed that the mall planner chooses the equilibrium that maximizes some criteria; that is, the mall developer plays the role of a "focal arbitrator" (Schelling 1960) in the selection of an equilibrium. It should be emphasized, however, as the "Estimation Method" subsection elucidates, that this equilibrium selection rule does not play any explicit role in the estimation of the model. Beyard and O'Mara (1999) and discussions with mall planners indicate that the process leading to the observed configuration of anchor stores in a mall indeed has properties similar to those of a simultaneous-move entry game. The Web Appendix (www.marketingpower.com/jmr_webappendix) provides a more detailed discussion on the

role of the mall developer and shows how the current model is consistent with a more complex process of bargaining between the anchor stores and the mall developer.

The incomplete information model provides a tractable framework for studying the complex interactions of entry decisions in which a trade-off exists between increased competition and the desirability of certain market characteristics. The model translates the discrete actions of competitors into smooth entry decision probabilities that represent the likelihood of a firm's entry into a market rather than actual realizations. In addition, allowing players to have private information greatly simplifies the computational burden of estimation. By avoiding the complicated regions of integration that arise in the complete information case, the model can accommodate a much larger number of players than what is typically feasible in games of complete information.

Formally, consider a game in which a finite number of potential stores indexed by $i = 1, \dots, n$ choose simultaneously whether to enter in a given market. Each potential entrant can be one of three types $T = 1, \dots, 3$, 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.

Unlike Mazzeo (2002) and others, this study takes store types to be exogenously given. This is a reasonable assumption for many retail stores that already have a well-established brand (e.g., department stores) and thus only face the decision whether to enter 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 $a = (a_1, \dots, a_n)$ denote a generic element of A . Finally, let $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$ 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- and store-specific variables. Define $X_m = \prod_{i=1}^n X_{im}$, and let $x_m = (x_{1m}, \dots, x_{nm}) \in X_m$ denote the vector of state variables of all stores in market m . It is assumed that x_m is common knowledge to all players in the game as well as to the econometrician.

The random variable $\varepsilon_i(a_i)$ is an idiosyncratic term representing private information to store i . The density of $\varepsilon_i = [\varepsilon_i(a_i = 0), \varepsilon_i(a_i = 1)]$ is denoted as $f(\varepsilon_i)$ and is common knowledge to all stores.

The *ex post* profits earned by store i are given by

$$(1) \quad \tilde{\Pi}_i(a_i, a_{-i}, x_m, \varepsilon_i, T; \theta) = \Pi_i(a_i, a_{-i}, x_m, T; \theta) + \varepsilon_i(a_i)$$

where $\Pi_i(a_i, a_{-i}, x_m, T; \theta)$ is a known and deterministic function of the states x_m and actions a_i of player i of type T and of the actions of the other stores, a_{-i} , and θ is a vector of parameters, which may be firm or type specific.

At this point, two assumptions for model identification purposes are made (the importance of these assumptions is discussed in the “Model Identification and Estimation Method” section):

Assumption 1: The error terms ε_i are distributed i.i.d. Type I extreme value across actions a_i and players i .

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

Because a given store i does not observe the other stores’ ε_i , 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 (see Assumption 1), the expected profits for firm i associated with the decision of entering (i.e., $a_i = 1$) equal

$$(2) \quad \begin{aligned} & E[\tilde{\Pi}_i(a_i = 1, a_{-i}, x_m, \varepsilon_i(1), T; \theta)] \\ &= \sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta) \Pr_{-i}(a_{-i} | x_m, \theta) + \varepsilon_i(1), \end{aligned}$$

where $\Pr_{-i}(a_{-i} | x_m, \theta) = \prod_{j \neq i} p_j(a_j | x_m, \theta)$ represents the beliefs that firm i has about the probability of each possible combination of entry and no-entry decisions taken by the other firms in the market. For each $j \neq i$, the probability (from firm i ’s perspective) that firm j enters the market is given by

$$(3) \quad \begin{aligned} & p_j(a_j = 1 | x_m, \theta, T) \\ &= \Pr\{E[\tilde{\Pi}_j(a_j = 1, a_{-j}, x_m, \varepsilon_j(1), T; \theta)] \geq \varepsilon_j(0)\}, \end{aligned}$$

where the deterministic part of profits for the no-entry option has been normalized to zero (see Assumption 2).

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 $p_i^*(a_i = 1 | x_m, \theta, T)$ for each store $i = 1, \dots, n$.

The distribution of ε (see Assumption 1) implies that the probability, prior to the realization of the ε ’s, that firm i chooses to enter market m is

$$(4) \quad \begin{aligned} & p_i^*(a_i = 1 | x_m, \theta, T) \\ &= \frac{\exp\left[\sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta) \Pr_{-i}^*(a_{-i} | x_m, \theta)\right]}{1 + \exp\left[\sum_{a_{-i}} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta) \Pr_{-i}^*(a_{-i} | x_m, \theta)\right]} \\ & \quad \text{for } i = 1, \dots, n, \end{aligned}$$

where $\Pr_{-i}^*(a_{-i} | x_m, \theta) = \prod_{j \neq i} p_j^*(a_j | x_m, \theta)$.

The equilibrium probabilities p^* solve the system of equations in Equation 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 p^* satisfies Equation 4. The “Model Identification and Estimation Method” section discusses the issues of existence and multiplicity of equilibria.

Profit Functions

The deterministic component of the profit function of store i of type T is specified to be linear in the common state vector x_m and on the actions of the other stores:

$$(5) \quad \begin{aligned} \Pi_i(a_i = 1, a_{-i}, x_m, T; \theta) &= \alpha_i + \beta_i' x_m + \delta_{TT} \sum_{j \neq i} 1\{a_j = 1, T\} \\ &+ \sum_j \delta_{\tilde{T}T} \times 1\{a_j = 1, \tilde{T} \neq T\}, \end{aligned}$$

where α_i is the store-specific mean profitability level, β_i measures the impact of the state variable x_m 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 ($T \neq \tilde{T}$) stores. The parametric linear specification in Equation 5 can be considered a first-order approximation to a variety of strategic models and is widely used in the entry models literature.

The key parameters in the model are the strategic effects parameters δ_{TT} and $\delta_{\tilde{T}T}$. Standard oligopoly models predict that due to the competition (business stealing) effect and when demand is exogenous, there are no positive spillovers between stores. Most of the entry models take $\delta < 0$ as given (thus imposing this restriction in the estimation) to conveniently reduce the possibility of multiple equilibria. However, when demand is endogenous, as in a regional shopping center, firms may benefit from entering together, as discussed previously, in which case this parameter does not need to be necessarily negative. In this article, I do not restrict the sign of the strategic effects parameters but rather let the data speak for themselves.

The mean profitability parameter α_i captures factors related to the efficiency of the stores as well as with other cost factors such as the level of rents, provided the level of the rents for a given store is similar across regional shopping centers. Under Assumptions 1 and 2, the system of equations that defines the equilibrium can be written as follows:

$$(6) \quad \begin{aligned} & p_i^*(a_i = 1 | x_m, \theta, T) \\ &= \frac{\exp\left[\alpha_i + \beta_i' x_m + \delta_{TT} \sum_{j \neq i} p_j^*(a_j = 1 | x_m, \theta, T) + \sum_j \delta_{\tilde{T}T} \times p_j^*(a_j = 1 | x_m, \theta, \tilde{T} \neq T)\right]}{1 + \exp\left[\alpha_i + \beta_i' x_m + \delta_{TT} \sum_{j \neq i} p_j^*(a_j = 1 | x_m, \theta, T) + \sum_j \delta_{\tilde{T}T} \times p_j^*(a_j = 1 | x_m, \theta, \tilde{T} \neq T)\right]} \\ & \quad \text{for } i = 1, \dots, n \text{ and } j = 1, \dots, n. \end{aligned}$$

Exogenous Determinants of Profits

Tables 3 and 4 provide summary statistics of the state variables X_m used to estimate the model. The demographic variables in Table 3 include population size, median age,

and average household size of the population within a 20-mile radius of the mall and the percentage of single-family houses and the median house value in each mall's three-digit zip code. Mall developers and potential tenants see all these demographic variables as key profit determinants.

Table 3 also describes several shopping center-specific characteristics that are included in the model to control for other factors that can potentially influence stores' profits.

Table 3
DESCRIPTIVE STATISTICS: MARKET-SPECIFIC VARIABLES

Variable	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Pop	1,205,753.06	1,825,253.05	34,698	12,272,646
Age	37.02	3.14	24.90	55.30
Size HH	2.57	.22	1.84	3.53
Sing Fam House	.67	.11	.11	.86
House Med Val	123,376.29	66,925.04	50,800	53,800
Parking	3746.63	1024.26	650	7177
Date 2	.32	.47	0	1
Date 3	.35	.48	0	1
Developer GGP	.17	.37	0	1
Developer SPG	.15	.36	0	1
Other Dev Medium	.26	.44	0	1

Notes: Pop is the total population within a 20-mile radius of the mall. 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 Fam House is the percentage of houses that are single-family houses in the same three-digit zip code of the mall, and House Med Val is the median value of the houses located also in the same three-digit zip code of the mall. Parking is the number of parking spaces of the mall. Date 2 and Date 3 are dummy variables reflecting the opening date of the mall. The omitted reference date group is Date 1 (not reported), which is a variable that takes the value of 1 if the mall was opened between 1952 and 1972. Date 2 is a dummy variable that takes the value of 1 if the mall was opened between 1973 and 1980, and Date 3 is a dummy variable that takes the value of 1 if the mall was opened between 1981 and 2006. Developer GGP, Developer SPG and Other Dev Medium 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.

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 because larger shopping centers are naturally expected to be able to accommodate more and bigger anchor stores. Note that a shopping center's size does not necessarily reflect its ability to draw demand; in many cases, shopping center size is a result of the interaction between several 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). The date when the mall opened is included 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 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, because older malls often have lower sales than more recent malls. A categorical variable with the identity of the mall developer is included because different developers develop long-term relationships with different types of anchors.

Other variables important for the success of a shopping center involve its accessibility and visual exposure. Because these variables are seen as essential requirements when planning larger shopping centers, it can be assumed that such characteristics will be captured by the constant terms, because they are, in general, shared by all the centers in the sample.

Table 4 presents several firm-specific variables that are also considered—namely, the identity of each anchor store, the size (in square feet) of the store,² and the distance of each store to its headquarters and nearest distribution center. Several comments are in order regarding these variables.

Because a store's size is the result of a complex interaction between many different factors, it is not possible to

²Note that there is no information on the size of a given anchor store for markets in which the store does not enter. For such markets, the store size variable was imputed using the observed store sizes for the markets in which the store entered.

Table 4
DESCRIPTIVE STATISTICS: STORE-SPECIFIC VARIABLES

Variable	<i>Sears</i>	<i>Target</i>	<i>Other Discount</i>	<i>JCPenney</i>	<i>Mervyn's</i>	<i>Other Midscale</i>	<i>Dillard's</i>	<i>Macy's</i>	<i>Other Upscale</i>
Number of Stores	370	60	32	334	54	229	155	88	230
<i>Store Sqft</i>									
M	120,589.31	122,165.55	—	105,043.64	81,013.02	—	132,433.31	164,127.78	—
SD	41,200.33	13,334.25	—	37,769.37	4703.83	—	33,138.73	32,665.55	—
Min	49,360	75,883	—	34,340	58,275	—	40,000	50,000	—
Max	385,000	191,719	—	263,255	152,728	—	280,000	385,000	—
<i>Dist HQ</i>									
M	1271.59	1486.38	—	1560.76	2877.98	—	1357.96	1214.65	—
SD	833.34	662.30	—	702.02	1191.80	—	757.44	974.34	—
Min	16.95	9.48	—	19.97	4.40	—	4.83	16.83	—
Max	6771.46	6359.21	—	6078.82	4472.37	—	6481.02	7114.37	—
<i>Dist DC</i>									
M	672.63	379.43	—	257.81	2615.23	—	766.12	293.03	—
SD	814.04	320.24	—	310.08	1,161.20	—	497.57	290.94	—
Min	6.92	2.63	—	1.89	13.44	—	12.58	5.52	—
Max	6097.28	3929.13	—	3935.29	4341.93	—	4685.70	2323.30	—

Notes: Store Sqft is the square footage of the store in a given mall. Dist HQ and Dist DC are the distances from each store to its headquarters and nearest distribution center (in kilometers), respectively.

model explicitly how a specific store size for a particular mall is determined. Therefore, in the model, store size is taken as given by stores when making their entry decisions. This implicitly assumes that store size decisions (similar to mall size and mall location decisions) are made before the stores' entry decisions. That is, when a store considers entering a specific mall, its decision is already conditional on a specific store size, which may have been the result of a previous stage in the game. Furthermore, I assume that the size of a given store is excluded from the profit equations of the other stores. As such, store size is a firm-specific variable that shifts the individual profit functions without affecting the other stores' profit functions. Because these are large anchor stores, the effect of a competitor's store size on firms' profits is second order relative to the competitive effects (which are first order).

The use of the last two firm-specific variables (distance of each store to its headquarters and nearest distribution center) is motivated by the work of Zhu and Singh (2009), who use these variables to generate exclusion restrictions in their study of store location decisions of retail discount firms. In the current application to the shopping center industry, these two distance variables lead to valid exclusion restrictions because they can be interpreted as being entry shocks specific to each store that only enter competitors' profit functions indirectly through that store's entry probabilities. In other applications, however, a variable such as the distance to nearest distribution center may not lead to valid exclusion restrictions. This happens in situations in which, for example, distance from nearest distribution center affects prices (and thus demand) at the local level. For an extensive discussion of why such variables do not lead to valid exclusion restrictions in their supermarket application, see Basker and Noel (2009). With regard to department stores, logistics/operations managers at well-known department stores confirmed that the distance to distribution centers does not affect local prices. Therefore, these distance variables can serve as proxies to capture logistical issues in department stores' entry decisions into each specific mall. For a more detailed discussion of the importance of using firm-specific variables as a way to generate exclusion restrictions, see the "Model Identification and Estimation Method" section.

Market Unobservables

Given that there may be some exogenous determinants of profits that the econometrician might not observe, this may raise concerns regarding the endogeneity of players' decisions. Intuitively, if there is a market characteristic that is not included in the model but shifts the profits of two players in the same direction, it might be inferred that there are strategic effects between these firms when this is actually not what is driving the correlation between firms' entry decisions.

To control for unobserved market characteristics that may lead to spurious within-market correlation in the actions of the players, a random-coefficients model specification is used. Including random coefficients plays a dual role in this setup. Apart from helping address market endogeneity issues, as described previously, it helps uncover different "market types," bringing additional insight to the estimation results.

If a discrete mixing distribution is used for the store intercepts and each of the stores' intercepts is allowed to take $r =$

1, ..., R distinct values (depending on the unobservable market type) with a given probability, the entry probabilities in Equation 6 become

$$(7) \quad p_i^*(a_i = 1 | x_m, \theta, T, r) = \sum_{r=1}^R s_r \times \frac{\exp \left(\alpha_{ir} + \beta'_i x_m + \delta_{TT} \sum_{j \neq i} p_j^*(a_j = 1 | x_m, \theta, T) + \sum_j \delta_{\tilde{T}T} \times p_j^*(a_j = 1 | x_m, \theta, \tilde{T} \neq T) \right)}{1 + \exp \left(\alpha_{ir} + \beta'_i x_m + \delta_{TT} \sum_{j \neq i} p_j^*(a_j = 1 | x_m, \theta, T) + \sum_j \delta_{\tilde{T}T} \times p_j^*(a_j = 1 | x_m, \theta, \tilde{T} \neq T) \right)}$$

for $i = 1, \dots, n,$

where R is the number of unobserved market types and the mixing weights $s = (s_1, s_2, \dots, s_r)$ are interpreted as the unobserved market-type sizes.

Potential Entrants

The number of potential entrants in each market is assumed to be nine for all markets: three upscale, three mid-scale, and three discount department stores. The variables "distance to headquarters" and "distance to nearest distribution center" help control for the fact that some stores are more likely to enter in some geographical areas than others. So, for example, a store such as Mervyn's, which tends to locate more on the West Coast (where its headquarters and most of its distribution centers are located), will be more likely to enter in a mall on the West Coast, and this will be captured by the distance variables' coefficients for this store. Robustness checks were conducted to test that the results obtained in the empirical estimates section hold over a reasonable range of the number of potential entrants (see the Web Appendix at www.marketingpower.com/jmr_webappendix).

MODEL IDENTIFICATION AND ESTIMATION METHOD

One of the most important dimensions in which entry models differ is 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 the current empirical application, stores are assumed to play an incomplete information game (i.e., like the econometrician, potential entrants have incomplete information about one another's profits). Consequently, this section focuses on the identification and estimation of this type of game. For a brief review on the identification issues encountered in complete information games, which can be used as background for the following discussion, refer to the Web Appendix (www.marketingpower.com/jmr_webappendix).

Conditions for Identification of the Model's Structural Parameters

The main purpose of this section is to show how the assumptions made in the model setup section are required to achieve identification. I choose to study the nonparamet-

ric identification of the model. Although the estimation strategy used in this research relies on parametric methods, it is well known that a (parametric) estimation approach is more robust if it does not rely exclusively on functional form assumptions. I proceed by first establishing sufficient conditions for nonparametric identification of the deterministic part of the expected payoff functions $E[\Pi_i(a_i, a_{-i}, x_m, \varepsilon_i, T; \theta)]$ and then by investigating the identification of the deterministic part of profits, $\Pi_i(a_i, a_{-i}, x_m, \varepsilon_i, T; \theta)$, taking the expected profits as given.³ The insights for this identification strategy are borrowed from the work of Hotz and Miller (1993) and Magnac and Thesmar (2002) in the context of dynamic games and from Bajari et al. (2010) in the static game setting.

Rust (1994) and Magnac and Thesmar (2002), among others, demonstrate that, in the absence of a distribution for the private information (ε), the model is clearly not identified. Furthermore, Bajari et al. (2010) demonstrate that when the deterministic part of profits is allowed to be nonparametric, an independence assumption is required for identification. Given this discussion, Assumption 1 is imposed, under which the error terms ε are distributed i.i.d. Type I extreme value across actions and players. On the basis of this assumption, the equilibrium in the model must satisfy the following:

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

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

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

Equation 9 maps $\{E[\Pi_i(a_i = 1, \cdot)] - E[\Pi_i(a_i = 0, \cdot)]\}$ into $p_i^*(a_i = 1 | x_m, \theta, T)$. These equations can be rewritten as follows:

$$(10) \quad p_i^*(a_i = 1 | x_m, \theta, T) = \varphi \{E[\Pi_i(a_i = 1, a_{-i}, x_m, T; \theta)] - E[\Pi_i(a_i = 0, a_{-i}, x_m, T; \theta)]\},$$

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

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

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

Although the particular functional form of p_i in Equation 9 depends on the probability distribution chosen for ε , the preceding mapping holds for error distributions other than the Type I extreme value distribution as long as the chosen distribution has full support.

From Equation 11, it is evident that one cannot infer the absolute values $\Pi_i(a_i = 1)$ and $\Pi_i(a_i = 0)$ but only their difference. This means that a payoff normalization is required for identification because payoffs are only determined relative to the payoff under action $a_i = 0$ (no entry decision). So, to allow for identification, Assumption 2 is imposed. This assumption—that the deterministic part of profits associated with the no-entry decision is zero—is a standard identifying assumption in any multinomial choice model.

Now that the assumptions needed to identify the expected profit functions $E[\Pi_i]$ are laid out, I turn to the identification of the profit functions Π_i given choice probabilities and expected profits. The relationship 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 follows:

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

where $Pr_{-i}(a_{-i} | x_m, \theta) = \prod_{j \neq i} p_j(a_j | x_m, \theta)$.

Heuristically, the problem of nonparametric identification of the profit functions Π_i is equivalent to the problem of finding a solution to the system of linear equations in Equation 12. Even after having established conditions for the identification of the values of $E[\Pi_i]$ and p_{-i} in Equation 12, it is clear that, without further restrictions, the system is not identified for Π_i . Holding x_m fixed and making use of Assumption 2, this system has n equations and $n \times 2^{n-1}$ unknowns (assuming here that there is one player of each type). This identification problem is related to the so-called reflection problem, which appears in models with social interactions (Manski 1993).

Note that the degree of underidentification in Equation 12 increases with the number of agents. So, for example, if there is a game with two players and one additional player is added, the number of equations in the system increases by only 1, while the number of unknowns increases by 4.

To identify the payoffs Π_i in Equation 12 nonparametrically, cross-equation restrictions (i.e., exclusion restrictions) must be imposed.⁴ I 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. The Appendix provides an example with two players, which formally demonstrates how adding exclusion restrictions may help with identification.

The intuitive basis for the need of exclusion restrictions in order to achieve identification is quite strong. Suppose, for example, that two firms ($i = 1, 2$) frequently enter together in the same markets (i.e., p_1 is highly correlated

³Nonparametric identification of the model is then studied by observing the space of expected payoffs. An alternative and equivalent approach relies on the examination of the system of equations with the choice probabilities in Equation 4 (see Schmidt-Dengler and Pesendorfer 2008).

⁴Related 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 Schmidt-Dengler and Pesendorfer 2008). Bajari et al. (2010) and Bajari, Hong, and Ryan (2010) also demonstrate the need for such exclusion restrictions in the context of static games.

with p_2). This can be strong evidence that the firms are strategic complements and/or that both firms have incentives to enter in markets with the same characteristics. To separate these stories, I use firm-specific variables that should have a strong effect on the profits of firm 1 but should have no direct effect on the profits of firm 2. If, indeed, the strategic effects are large and positive, shifting up the profits of firm 1 through the firm-specific variable should also increase the probability of firm 2 entering the market (p_2). If there are no spillovers and the observed pattern is a result of firms 1 and 2 having the same preferences with respect to some market characteristic, p_2 should remain unchanged. This reasoning applies to situations in which the market characteristic can be either observable or unobservable.

This section demonstrates that to identify the profit function parameters, it is necessary to rely on the mapping from the choice probabilities to the expected profit functions in Equation 11. Because these choice probabilities represent the equilibria of an incomplete information game, it is necessary to make sure that an equilibrium exists and make an assumption regarding the (possible) multiplicity of equilibria.

Existence and Multiplicity of Equilibria

Ensuring the existence of equilibria in this setting corresponds to making sure that the system of equations given by Equation 4 has at least one solution.⁵ 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 Brouwer's fixed-point theorem (see McKelvey and Palfrey 1995).

Because, 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 Equation 4), 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 standard 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).

Estimation Method

Su and Judd (2008) propose a new optimization approach to estimate a general class of structural models using maximum likelihood. The proposed approach fits in a class of problems known as MPECs (mathematical programs with equilibrium constraints). Although this research constitutes the first attempt to apply this approach empirically, MPEC can be used to estimate a variety of economic models other than entry models.

⁵This research only examines pure strategy equilibria because, provided the distribution of $\varepsilon_i(a_i)$ is atomless, the best reply function in Equation 4 is unique, and thus there is no need to consider mixed strategies.

Following Su and Judd's (2008) insight, the structural parameters and endogenous economic variables are chosen to maximize the likelihood function subject to the constraint that the endogenous economic variables are consistent with the equilibrium defined by the structural parameters.⁶ 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:

$$(13) \quad L(\theta) = \prod_{m=1}^M \prod_{i=1}^n \left[p_i^*(a_i = 1 | x_m, T; \theta) \right]^{1\{a_i=1, T\}} \times \left[1 - p_i^*(a_i = 1 | x_m, T; \theta) \right]^{1\{a_i=0, T\}}$$

subject to

$$p_i^*(a_i = 1 | x_m, T; \theta)$$

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

Then, using one of the state-of-the-art constrained optimization programs,⁷ the estimates of the parameters of the model are obtained.

Previous studies have proposed other methods of estimating discrete games of incomplete information, the most common being nested fixed-point algorithms and two-step estimators (see, e.g. Aguirregabiria and Mira 2002, 2007; Bajari et al. 2010). For a brief comparison of the MPEC method with these alternative estimation methods, refer to the Web Appendix (www.marketingpower.com/jmr_webappendix).

THE ENTRY MODEL IN PRACTICE: EMPIRICAL ESTIMATES

Tables 5, 6, and 7 report the estimates of the structural parameters of the profit functions in Equation 5. Table 5 reports the estimates of the parameters associated with the market- and store-specific continuous variables for each store, Table 6 lists the parameter estimates associated with the market-specific dummy variables, and Table 7 displays the parameter estimates for the strategic effects. This section discusses the results for the model with unobserved market heterogeneity with two market types (i.e., for $R = 2$).⁸

⁶For an illustration of how MPEC helps with multiple equilibria issues, refer to the Web Appendix (www.marketingpower.com/jmr_webappendix).

⁷To obtain the parameter estimates shown in the "The Entry Model in Practice: Empirical Estimates" section, 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.

⁸Going from $R = 2$ to $R = 3$ did not substantially change the results and implied that one of the market types had a very small probability mass. So, although the model can be generalized to an R larger than 2, I chose to report the results for when $R = 2$. The choice of a discrete mixing distribution for the market types as opposed to a continuous mixing distribution is due to computational reasons.

Table 5
PROFIT FUNCTIONS' PARAMETER ESTIMATES (MARKET- AND STORE-SPECIFIC CONTINUOUS VARIABLES)

Variable	Discount Department Stores			Midscale Department Stores			Upscale Department Stores		
	Sears	Target	Other Discount	JCPenney	Mervyn's	Other Midscale	Dillard's	Macy's	Other Upscale
Intercept 1	.36 (.31)	-.99* (.31)	-1.40* (.44)	-1.01* (.25)	-4.47* (.33)	.34 (.21)	-4.26* (.41)	-3.10* (.31)	-1.12* (.23)
Intercept 2	-6.34* (0.74)	-2.04* (.36)	-.38 (.38)	-2.56* (.28)	-2.06* (.25)	-.76* (.17)	-4.30* (.40)	-2.42* (.25)	-.26 (.15)
Pop	-2.06* (.22)	.04 (.12)	-.20 (.11)	-1.12* (.10)	.43* (.08)	-.43* (.06)	-.20 (.11)	.65* (.09)	.53* (.07)
Age	7.65* (1.05)	-1.10 (1.84)	10.40* (1.66)	5.96* (.74)	-6.81* (1.02)	-2.97* (.89)	-1.37 (.90)	5.58* (1.26)	-2.03* (.73)
Size HH	5.32* (1.60)	-2.22 (1.63)	4.61* (1.87)	7.88* (1.14)	3.78* (1.42)	-3.55* (.92)	-1.00 (1.22)	-1.59 (1.13)	-7.69* (.93)
Sing Fam House	-1.62* (.36)	3.18* (.70)	-.07 (.45)	-1.62* (.25)	5.18* (.83)	.80* (.40)	-1.50* (.53)	-1.69* (.29)	.24 (.22)
House Med Val	.26 (.20)	2.85* (.90)	-3.02* (.56)	-.88* (.16)	1.45* (.25)	-.98* (.17)	-2.00* (.42)	2.75* (.23)	.51* (.14)
Parking	1.52* (.40)	.98* (.47)	.38 (.34)	1.88* (.22)	2.18* (.38)	.20 (.19)	2.44* (.27)	1.63* (.30)	1.67* (.24)
Store Sqft	2.66* (.35)	-2.97* (1.10)	—	2.09* (.28)	-4.31* (1.59)	—	-.24 (.33)	-.57 (.37)	—
Dist HQ	.23 (.21)	-.91* (.22)	—	-.10 (.06)	-.77* (.20)	—	-1.28* (.16)	1.03* (.29)	—
Dist DC	-.46* (.06)	.03 (.12)	—	-.17* (.06)	-.18 (.16)	—	-1.38* (.15)	-.26* (.07)	—

Notes: Standard errors are in parentheses. Values statistically significant at the 5% level are marked with an asterisk. Pop is the total population within a 20-mile radius of the mall. 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 Fam House is the percentage of houses that are single-family houses in the same three-digit zip code of the mall, and House Med Val is the median value of the houses located also in the same three-digit zip code of the mall. Parking is the number of parking spaces of the mall. Store Sqft is the square footage of the store in a given mall. Dist HQ and Dist DC are the distances from each store to its headquarters and nearest distribution center, respectively. Two intercepts are reported for each store because the latent model specification allows for two different unobserved market types. All continuous X variables are transformed as follows: $X^{\text{transformed}} = \ln[(X^{\text{original}}/\text{mean}(X^{\text{original}}))]$ so that each variable has a value of zero at its mean (across malls) value.

To obtain the standard errors presented in the tables, a bootstrap resampling approach is used to generate several (500) synthetic data samples of the same size as the original problem (i.e., each with 561 markets). Then, the estimation problem is resolved many times (i.e., once for each of the synthetic data sets). The standard errors reported are the standard errors generated by the bootstrap.

The Role of Market- and Store-Specific Characteristics

The intercepts in Table 5 are negative for all stores. Thus, at the median level of the continuous variables and at the omitted categories of the dummy variables, none of the anchor stores has incentives to enter on its own.

The coefficient estimates associated with the population size variable seem to indicate that discount and midscale department stores prefer to locate in less populated areas, while upscale stores prefer more populated markets.

The purchasing power of the area surrounding each shopping center (measured here by the proxy variable median house value) has, in general, a positive effect for upscale stores and a negative effect for discounters and midscale stores. This reflects the notion that, all else being equal, discount and midscale department stores usually cater to middle- to lower-income buyers who are more cost conscious and buy lower-quality merchandise at a discount because they may not be able to afford better brands on a regular basis.

Notably, the estimated parameter of the purchasing power variable is positive for Target, confirming the common perception that Target, compared with other discounters, is not solely patronized by lower-income customers.

The coefficient estimates associated with the variables median age, average household size, and percentage of single-family houses in the area surrounding the mall do not exhibit a clear pattern across or within store types, with the exception of a few particular cases. For example, discount department stores seem to prefer to enter in markets with an older population, 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 as a function of the area available for construction of the mall. All stores, with the exception of Target and other discount department stores (e.g., Wal-Mart, Kmart, Big Lots), seem to prefer to anchor larger centers; it is possible that, because these stores are mostly one-stop destination stores (i.e., stores to which consumers make a special trip

Table 6
PROFIT FUNCTIONS' PARAMETER ESTIMATES (MARKET-SPECIFIC DUMMY VARIABLES)

Variable	Discount Department Stores			Midscale Department Stores			Upscale Department Stores		
	Sears	Target	Other Discount	JCPenney	Mervyn's	Other Midscale	Dillard's	Macy's	Other Upscale
Date 2	.33 (.18)	-.02 (.18)	.40 (.26)	.92* (.18)	.31 (.17)	.05 (.18)	.86* (.19)	-.35* (.17)	.29* (.12)
Date 3	.75* (.21)	.50* (.26)	-.03 (.39)	1.38* (.17)	.99* (.21)	.37* (.17)	1.67* (.20)	.09 (.17)	.59* (.14)
Developer GGP	1.04* (.28)	-.06 (.17)	-.05 (.40)	1.39* (.20)	.57* (.25)	-.46* (.20)	.57* (.21)	1.54* (.22)	.55* (.16)
Developer SPG	1.24* (.29)	-.60 (.37)	-2.02 (2.76)	1.22* (.21)	.11 (.26)	.48 (.18)	1.06* (.24)	2.07* (.24)	1.08* (.18)
Other Dev Medium	1.26* (.28)	-.71* (.26)	-.72* (.37)	1.61* (.22)	.39 (.21)	.75* (.18)	.65* (.18)	.56* (.23)	.93* (.11)

Notes: Standard errors are in parentheses. Values statistically significant at the 5% level are marked with an asterisk. Date 2 and Date 3 are dummy variables reflecting the opening date of the mall. The omitted reference date group is Date 1 (not reported), which is a variable that takes the value of 1 if the mall was opened between 1952 and 1972. Date 2 is a dummy variable that takes the value of 1 if the mall was opened between 1973 and 1980, and Date 3 is a dummy variable that takes the value of 1 if the mall was opened between 1981 and 2006. Developer GGP, Developer SPG, and Other Dev Medium are dummy variables reflecting the developer of a given mall. The omitted reference developer group is All Other Developers (not reported), which takes the value of 1 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.

Table 7
PROFIT FUNCTIONS' PARAMETER ESTIMATES
(STRATEGIC EFFECTS)

	Discount	Midscale	Upscale
Discount	-.24* (.10)	-.88* (.40)	.83* (.41)
Midscale	.09 (.28)	.60* (.26)	.48* (.19)
Upscale	-.73* (.30)	-1.15* (.17)	-.67* (.18)

Notes: Each value in the table corresponds to the impact of row store type *i* on the column store type *j*. Standard errors are in parentheses. Values statistically significant at the 5% level are marked with an asterisk.

with the intent of shopping), they do not benefit as much as “impulse” or “backup”⁹ stores from being located in larger malls.

The last three variables in Table 5 are store-specific variables: store’s square footage and store’s distance to headquarters and to the nearest distribution center (Store Sqft, Dist HQ, and Dist DC, respectively). Regarding the estimated coefficients associated with the variable Store Sqft, Sears, JCPenney, and Dillards are more likely to enter in malls that provide them with stores larger than their median size stores. The opposite seems to be true for others.

As expected, the Dist HQ and Dist DC variables are useful predictors of most stores’ entering decisions. The coeffi-

⁹Conversations with mall managers indicate that consumers are likely to view a department store as a “backup” store, meaning 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 (2003) provide an excellent discussion of these effects.

icients associated with these variables are negative for almost all stores suggesting that, *ceteris paribus*, stores tend to prefer markets that are closer to their respective headquarters and existing distribution centers (because, by doing so, stores enjoy reduced costs—e.g., logistics and distribution-related costs).

Table 6 shows the estimated coefficients for the dummy variables associated with the opening date of the mall and the identity of the mall developer. As expected, stores in centers constructed more recently seem to have, all else being equal, larger profits than those located in older centers. Table 6 also indicates 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 (e.g., Target). 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 such as 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 malls built and managed by bigger developers because this seems to ameliorate the risks of the developments and guarantee higher expected profits.

Strategic Effects

Estimating a structural model of strategic entry can determine whether stores are strategic complements or strategic substitutes. Table 7 reports the parameter estimates associated with the stores’ strategic effects (i.e., δ_{TT} and $\delta_{\bar{T}T}$). 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., Table 7’s main diagonal), we find that midscale 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 seem to benefit considerably from the presence of discount and midscale department stores but that the reverse is not true. This finding is consistent with practitioner beliefs that many shoppers in discount and midscale department stores would be willing to trade up to a higher-end store but that the reverse is seldom true. This finding is reminiscent of the Blattberg and Wisniewski (1989) results for store brands, in which they find that cross-price elasticities between store brands and national brands are 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.

There are several mechanisms that can be used to explain why the gains from agglomeration are larger than the competitive effects for midscale department stores but not for other types. One possible theory is that consumers who patronize midscale department stores are likely to engage in comparison shopping. Another explanation is that a typical midscale department store consumer may need to engage in a wider search to find a product that suits his or her needs. This is plausible because, typically, there is a higher degree of differentiation in the brands offered by different midscale department stores than in those offered by their discount and upscale counterparts. Furthermore, when compared with higher-income shoppers, midscale department store patrons are more price sensitive, so they are more prone to scout out several different stores looking for deals. Upscale department stores (also called "specialty" department stores) usually are not as broad in scope as midscale department stores. In addition, consumers who typically shop at these stores tend to be more loyal because they develop special relationships with their sales personnel (many times making use of the personal shopping services) and are more time constrained (thus not exhibiting a tendency to shop around but rather engaging in more targeted shopping).

Unobserved Market Heterogeneity Parameters

The intercept estimates reported in the first rows of Table 5 measure the overall preference of each store for locating in each of the two latent market types (hereinafter called type 1 and type 2 markets). Table 5 indicates that different stores exhibit distinct preferences for each of the market types, as reflected in the different magnitudes of the esti-

mated intercepts. The estimated probability masses for each of the underlying market types are .78 for market type 1 and .22 for market type 2.

A more detailed investigation of the estimated coefficients in Table 5 shows that Sears and Target exhibit a strong preference for markets of type 1. So, given that the own strategic effects for discount stores are negative, type 1 markets (holding everything else constant) tend to have only one discount store, and this store is more likely to be a Sears or a Target than any other discount store. As for the midscale department stores, the estimates reveal that JCPenney and Mervyn's prefer different market types (i.e., the market unobservables do not seem to explain the joint entry decisions of these stores). For the upscale department stores, Dillard's does not seem to have a distinct preference for one of the market types, and Macy's tends to prefer type 2 markets. Moreover, a notable pattern is also revealed across different store types: The results suggest that Sears, Target, and JCPenney tend to prefer type 1 markets, while Mervyn's and Macy's prefer type 2 markets.

Profit Decomposition: Relative Importance of the Estimated Effects

The parameter estimates also indicate the importance of each source of profit for the overall profits of each store. The profits of each store are decomposed into the following sources: (1) market- and store-specific characteristics (which include demographics), (2) spillovers from stores of the same type, and (3) spillovers from stores of other types. Table 8 presents the results of this decomposition. The market- and store-specific 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 only store type that seems to benefit somewhat from the presence of stores of the same type is the midscale department store. In addition, the negative competitive effects from stores of different types seem to be an important determinant of long-term profits for midscale department stores. In contrast, the positive effects from stores of different types are a particularly important determinant of long-term profits for upscale department stores.

Table 8
DECOMPOSITION OF AVERAGE PREDICTED PROFITS OF EACH STORE BY SOURCE OF PROFITS

	<i>Average Total Profits</i>	<i>Market Specific</i>	<i>Store Specific</i>	<i>Strategic Same Type</i>	<i>Strategic Other Types</i>
Sears	.91	2.51	-1.05	-.04	-.52
Target	-2.29	-.51	-1.09	-.17	-.52
Other discount	-1.88	-.01	-1.18	-.18	-.52
JCPenney	-.33	2.45	-1.39	.30	-1.69
Mervyn's	-4.78	.03	-3.71	.60	-1.69
Other midscale	-.43	.75	.10	.41	-1.69
Dillard's	-1.63	1.26	-3.72	-.38	1.21
Macy's	-2.53	-.17	-3.11	-.46	1.21
Other upscale	.33	.31	-.93	-.29	1.24

Notes: Average Total Profits is the average of the in-sample predicted profits. Market Specific and Store Specific refer to the average profits that can be attributed to characteristics of the market and to store characteristics, respectively. Strategic Same Type is the average of profits that can be attributed to the entry of stores of the same type, and Strategic Other Type is the average of profits that can be attributed to the entry of stores of other types.

APPLICATIONS AND EXTENSIONS

The empirical methodology used here relies on a model of equilibrium outcomes, which allows not only for an economic interpretation of the estimated structural parameters but also for conducting some policy experiments. These experiments will yield results that have managerial implications for both mall developers and retailers and might also be useful from a public policy perspective.

It is in the financial interest of the mall developers to know, according to not only their own experience but also that of other mall developers, what type and size of externalities each anchor generates. This information may help mall developers more easily address questions such as how much to subsidize anchor stores (e.g., How much subsidy would a mall developer be willing to give to avoid the exit of a given anchor [in a given market setting]?) or local communities (e.g., How much would a mall developer be willing to subsidize a local community to get approval to include a given type of store in the center?). In general, the model’s results and insights, supplemented with developers’ internal data, will enable developers to better decide how to allocate resources.

These results could also potentially be used by department stores to inform decisions about entry. For example, consider an increase in population for a given market. On the basis of the stores already in the center and the other market characteristics, a department store could compute whether it is profitable to enter this shopping center. Even if it is not profitable, and depending on the spillovers for the other stores, the mall developer may be willing to subsidize the entry of this store.

The model can also be used for predictions about the effect of changes in other exogenous demographic variables. The demographics and character of the market surrounding a shopping center can change dramatically, requiring a center to reposition itself to accommodate the new market. Using new demographics, several hypotheses for expansion or reconfiguration of a given center might be explored.

The results may also be useful from a public policy perspective. Several programs, such as the federal EZ program,¹⁰ have been implemented to help overcome the underprovision of commercial services in low-income communities. The government may use a model like the one in this article to calculate which lump-sum grant (or other type of benefits) to give to a shopping center to make the entry of the stores (and thus of the mall) profitable.

Unfortunately, as is common in this type of strategic interaction models, the possibility of multiple equilibria makes it impractical to run more complicated counterfactual scenarios that require solving for the equilibrium (equilibria) in a particular market. Developing methods that allow for performing such counterfactuals is an important area for further research.¹¹

¹⁰The concept of the EZ (“Enterprise Zone”) program was invented by Peter Hall, a U.K. geographer, in the late 1970s. Inspired by the success of Hong Kong and the free trade zones in Singapore, he proposed to reduce government regulations and taxes in distressed urban communities to stimulate economic development.

¹¹For example, Aguirregabiria’s (2009) recent study is an attempt to make headway in this direction.

Last, the proposed framework and method could be used outside the realm of the shopping center industry to understand not only the entry and market structure determinants of other industries but also the relative positioning of different brands within a product category.

CONCLUSION

This article studies the joint entry decisions of stores in a retail cluster. A strategic model of entry capable of quantifying the magnitude of interstore spillovers is proposed, and the research shows how these effects can help explain the composition of a given market. The model is applied to a novel data set containing information about the store configurations of all U.S. regional shopping centers and is estimated using a new econometric approach that is robust to multiple equilibria.

While the main contribution of this research is methodological, the analysis conducted provides greater understanding of the entry behavior in the shopping center industry, and its results have managerial implications for mall developers and anchor stores. The results show that several demographic variables are important in stores’ entry decisions. Also, while most stores seem to generate negative spillovers for the other stores, the results suggest that mid-scale department stores are strategic complements and that upscale department stores benefit considerably from the presence of discount and midscale department stores. Therefore, allowing for positive spillovers seems to be crucial when explaining the store configuration and variety existent in regional shopping centers. Mall developers and anchor stores can use these results to make more informed decisions (e.g., regarding the amounts for entry subsidies).

Future studies could use the proposed framework and method to answer questions related to the entry and market structure determinants of this and other industries. Furthermore, the methodology presented here could also be used in many other settings.

APPENDIX: IDENTIFICATION: TWO-PLAYER EXAMPLE

Let N be the number of firms, $\#A$ be the number of actions each firm can take, and K be the number of points in the support of a variable X that is market specific. In a market characterized by $X = x$ (which implies that $K = 1$), with $N = 2$ potential entrants ($I = 1, 2$) and $\#A = 2$ (“enter” and “not enter”), let $p_i = p_i(a_i = 1 | x)$. The system in Equation 12 can then be rewritten as follows:

$$(A1) \left\{ \begin{aligned} E[\Pi_1(a_1 = 1 | x, \theta)] &= \Pi_1(a_1 = 1, a_2 = 1 | x, \theta) \times p_2 \\ &\quad + \Pi_1(a_1 = 1, a_2 = 0 | x, \theta) \times (1 - p_2) \\ E[\Pi_1(a_1 = 0 | x, \theta)] &= \Pi_1(a_1 = 0, a_2 = 1 | x, \theta) \times p_2 \\ &\quad + \Pi_1(a_1 = 0, a_2 = 0 | x, \theta) \times (1 - p_2) \\ E[\Pi_2(a_2 = 1 | x, \theta)] &= \Pi_2(a_1 = 1, a_2 = 1 | x, \theta) \times p_1 \\ &\quad + \Pi_2(a_1 = 0, a_2 = 1 | x, \theta) \times (1 - p_1) \\ E[\Pi_2(a_2 = 0 | x, \theta)] &= \Pi_2(a_1 = 1, a_2 = 0 | x, \theta) \times p_1 \\ &\quad + \Pi_2(a_1 = 0, a_2 = 0 | x, \theta) \times (1 - p_1) \end{aligned} \right.$$

In line with Assumption 2, the second and last equations in Equation A1 can be omitted, so the result is a system with two equations and four unknowns. It is clear that, without any further restrictions, the $\Pi(\cdot)$ terms cannot be nonparametrically identified using exclusively the values for $E[\Pi_i]$ and p_i .

Suppose X is a variable that may take different values for different firms. For example, assume that the variable X can take one of two values $\{H, L\}$ (high or low). This implies that $K = 2$. In this case, and making use of Assumption 2, the system in Equation 12 can be written as follows:

$$(A2) \left\{ \begin{array}{l} E[\Pi_1(a_1 = 1 | X_1 = H, X_2 = H; \theta)] \\ \quad = \Pi_1(a_1 = 1, a_2 = 1 | X_1 = H, X_2 = H; \theta) \times p_2 \\ \quad + \Pi_1(a_1 = 1, a_2 = 0 | X_1 = H, X_2 = H; \theta) \times (1 - p_2) \\ E[\Pi_1(a_1 = 1 | X_1 = H, X_2 = L; \theta)] \\ \quad = \Pi_1(a_1 = 1, a_2 = 1 | X_1 = H, X_2 = L; \theta) \times p_2 \\ \quad + \Pi_1(a_1 = 1, a_2 = 0 | X_1 = H, X_2 = L; \theta) \times (1 - p_2) \\ \dots \\ E[\Pi_2(a_2 = 1 | X_1 = H, X_2 = H; \theta)] \\ \quad = \Pi_2(a_1 = 1, a_2 = 1 | X_1 = H, X_2 = H; \theta) \times p_2 \\ \quad + \Pi_2(a_1 = 0, a_2 = 1 | X_1 = H, X_2 = H; \theta) \times (1 - p_1) \\ \dots \end{array} \right.$$

The system in Equation A2 has (1) $N \times K^N = 2 \times 2^2 = 8$ equations and (2) $N \times K^N \times \#A^{N-1} = 2 \times 2^2 \times 2^1 = 16$ unknowns, so it is not identified. By allowing X to be not only market specific but also firm specific, both the number of equations and unknowns increased.

However, if there is (*ex ante*) reason to believe that each firm's profits are not affected by the value that X takes for the other firm, exclusion restrictions can be imposed that will make the system identified. This is because, while the number of equations will remain the same as in Equation A2, the number of unknowns will decrease considerably, because now $\Pi_i(a_i, a_j | X_i = x_i, X_j = x_j; \theta) = \Pi_i(a_i, a_j | X_i = x_i, X_j = x_j'; \theta) \forall x_j \neq x_j' \forall i \neq j$. With eight equations and $N \times K \times \#A^{N-1} = 2 \times 2 \times 2^1 = 8$ unknowns in the system, nonparametric identification of the deterministic part of profits can now be achieved.

Note that the necessary condition for identification using exclusion restrictions relies not on the *number* of firm-specific variables that are used but rather on the *variability* of such variables. According to Theorem 1 in Bajari et al. (2010) and to Theorem 3 in Bajari, Hong, and Ryan (2010), as long as there is sufficient variability in the firm-specific shifters, the deterministic part of profits $\Pi(\cdot)$ is nonparametrically identified. In the context of this two-firm example, one *single* discrete X variable with two-point support led to eight exclusion restrictions. Extending the same logic to the three-firm case results in $N \times K^N = 24$ equations and $N \times K^N \times \#A^{N-1} = 96$ mean profit parameters. Imposing the cross-equation restrictions results in only $N \times K \times \#A^{N-1} = 24$ parameters, which become identified. Again, even with one single discrete X .

Essentially, from items 1 and 2, it is evident that the number of unknown parameters in the system in Equation A2 depends linearly on K but the number of equations grows exponentially with K . By using variables that are firm specific and that have sufficient variability (i.e., that have sufficient high values for K), the model is identified.

REFERENCES

- Aguirregabiria, Victor (2009), "A Method for Implementing Counterfactual Experiments in Models with Multiple Equilibria," working paper, University of Toronto.
- and Pedro Mira (2002), "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70 (4), 1519–43.
- and — (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75 (1), 1–53.
- Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov (2010), "Estimating Static Models of Strategic Interaction," *Journal of Business and Economic Statistics*, 28 (4), 469–82.
- , —, and Stephen P. Ryan (2010), "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 78 (5), 1529–68.
- Basker, Emek and Michael D. Noel (2009), "The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry," *Journal of Economics and Management Strategy*, 18 (4), 977–1009.
- Berry, Steven T. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60 (4), 889–917.
- Beyard, Michael D. and W. Paul O'Mara (1999), *Shopping Center Development Handbook*. Washington, DC: Urban Land Institute.
- Blattberg, Robert C. and Kenneth J. Wisniewski (1989), "Price-Induced Patterns of Competition," *Marketing Science*, 8 (4), 291–309.
- Bresnahan, Timothy F. and Peter C. Reiss (1990), "Entry in Monopoly Markets," *Review of Economic Studies*, 57 (4), 531–53.
- Dubé, Jean-Pierre H., Jeremy T. Fox, and Che-Lin Su (2009), "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," NBER Technical Working Paper 14991, National Bureau of Economic Research.
- Dudey, Marc (1990), "Competition by Choice: The Effect of Consumer Search on Firm Location Decisions," *The American Economic Review*, 80 (5), 1092–1104.
- Ellickson, Paul B. and Sanjog Misra (2008), "Supermarket Pricing Strategies," *Marketing Science*, 27 (5), 811–28.
- Fujita, Masahisa and Jacques-François Thisse (1996), "Economics of Agglomeration," *Journal of the Japanese and International Economies*, 10 (4), 339–78.
- Gould, Eric D., B. Peter Pashigian, and Canice J. Prendergast (2005), "Contracts, Externalities, and Incentives in Shopping Malls," *Review of Economics and Statistics*, 87 (3), 411–22.
- Hotz, V. Joseph and Robert A. Miller (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60 (3), 497–529.
- Konishi, Hideo and Michael T. Sandfort (2003), "Anchor Stores," *Journal of Urban Economics*, 53 (3), 413–35.
- Magnac, Thierry and David Thesmar (2002), "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70 (2), 801–816.
- Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60 (3), 531–42.
- Mazzeo, Michael J. (2002), "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33 (2), 221–42.
- McKelvey, Richard D. and Thomas R. Palfrey (1995), "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10 (1), 6–38.

- Musalem, Andrés and Shin Woochoel (2009), "A Structural Model of Product Line and Pricing Decisions," working paper, Fuqua School of Business, Duke University.
- Orhun, A. Yeşim (2005), "Spatial Differentiation in the Supermarket Industry," working paper, Chicago Booth School of Business, University of Chicago.
- Rust, John (1994), "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, R. Engle and D. McFadden, eds. North Holland: Elsevier, 3082–3146.
- Schelling, Thomas C. (1960), *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Schmidt-Dengler, Philipp and Martin Pesendorfer (2008), "Asymptotic Least Squares Estimators for Dynamic Games," *The Review of Economic Studies*, 75 (3), 901–928.
- Seim, Katja (2006), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37 (3), 619–40.
- Stahl, Konrad (1982), "Location and Spatial Pricing Theory with Nonconvex Transportation Cost Schedules," *Bell Journal of Economics*, 13 (2), 575–82.
- Su, Che-Lin and Kenneth L. Judd (2008), "Constrained Optimization Approaches to Estimation of Structural Models," Discussion Papers 1460, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- Zhu, Ting and Vishal Singh (2009), "Spatial Competition with Endogenous Location Choices: An Application to Discount Retailing," *Quantitative Marketing and Economics*, 7 (1), 1–35.
- , ———, and Mark D. Manuszak (2009), "Market Structure and Competition in the Retail Discount Industry," *Journal of Marketing Research*, 46 (August), 453–66.

Copyright of Journal of Marketing Research (JMR) is the property of American Marketing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.