

A Bayesian Analysis of Post-Entry Outcomes: An Application to Shopping Mall Sales

Doug J. Chung
Kyoungwon Seo
Reo Song

Working Paper 17-071



A Bayesian Analysis of Post-Entry Outcomes: An Application to Shopping Mall Sales

Doug J. Chung
Harvard Business School

Kyoungwon Seo
Seoul National University

Reo Song
California State University, Long Beach

Working Paper 17-071

Copyright © 2017, 2018 by Doug J. Chung, Kyoungwon Seo, and Reo Song

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

A Bayesian Analysis of Post-Entry Outcomes: An Application to Shopping Mall Sales*

Doug J. Chung, Harvard University

Kyoungwon Seo, Seoul National University

Reo Song, California State University, Long Beach

March 2018

Abstract

We propose a Bayesian model of post-entry outcomes. Endogenous firm entries are modeled as a discrete game of complete information. The model encompasses many realistic elements, such as common aggregate shocks and within-/cross-category spillovers. Multiple equilibria are addressed by estimating a selection function from the observed data. The Bayesian approach that combines decision making with parameter inference, as well as multiple equilibria in complete information games, require demanding computational power for model estimation. We overcome computational burden by utilizing an improved simulator for likelihood evaluation and the general-purpose computing on graphics processing units (GPGPU) technology that takes advantage of multiple processing cores in a GPU to increase computational speed. We apply our model to shopping mall configuration and sales. We find that competition effects dominate within retail store categories, but that agglomeration effects exist across store categories. We find positive causal brand effects for midscale and upscale stores, above and beyond market effects, but find negative causal brand effects for discount stores on mall sales. We find that a developer is better off creating a mall in an affluent versus a populated market.

Key words: firm entry, discrete game, complete information, multiple equilibria, endogeneity, GPGPU technology, Bayesian estimation, shopping mall, spillover.

* Comments are welcome to Kyoungwon Seo, Seoul National University, Seoul, South Korea, seo8240@snu.ac.kr.

The authors thank Jose Alvarez, Jinhee Jo, the seminar participants at Carnegie Mellon University, CEIBS, Harvard University, HKUST, and the conference participants at the 2016 Marketing Science Conference for their comments and suggestions. The authors also thank Harvard Business School, and the Institute of Finance and Banking, and the Institute of Management Research at Seoul National University for providing financial support.

1. Introduction

Analyzing or predicting post-entry outcomes in a market is a crucial question for both firms and policy makers. Specifically, firms decide whether to open a business in a new market, and policy makers decide on regulations for governance of a market. For example, if an airline opens up a new route and, thus, enters the LA-Boston market, how much would prices drop due to tightened competition? Would the expected profit justify an airline’s entry? How would the entry affect revenues of the existing airlines in the market? Should the entrant airline be allowed to merge with an existing airline? Or should it not be allowed in order to prevent future predatory behavior? To answer these questions, past studies have proposed structural models that take into account not only market characteristics, but also endogeneity stemming from interactions among competing firms and sample selection biases (Reiss & Spiller, 1989; Mazzeo, 2002a; Zhu, Singh & Manuszak, 2009; and Ellickson & Misra, 2012).

In this paper, we propose a novel model of post-entry outcomes. The model encompasses many important realistic elements not included in previous models, mainly due to computational burden. Our model has the following features. First, we adopt the Bayesian approach, unlike the aforementioned papers. Second, we model firms’ entry decisions and total market sales to be affected by a common demand shock that captures endogeneity. If an unobserved characteristic of a market causes a firm to enter, higher sales are not purely due to the entry of the firm. Our model can separate out the pure causal effect of a firm’s entry on total market sales. Third, entries of several heterogeneous firms are modeled as a simultaneous-move discrete game of complete information. The game reflects interactions among firms and can identify their spillover effects. Fourth, we address the multiplicity of equilibria by specifying an equilibrium selection function whose parameter is estimated from the observed data. An equilibrium selection function provides useful information in predicting total market sales. Fifth, we use a new simulator that computes the model likelihood significantly faster than the usual frequency simulator—in particular, when there are as many as nine potential entrants in our application. Finally, we further reduce the computational burden of evaluating the empirical likelihood by utilizing a state-of-the-art

technology: general-purpose computing on graphics processing units (GPGPU), which uses multiple cores in a GPU to increase computational speed.

An important distinguishing feature of this paper is that we adopt a Bayesian approach to analyze post-entry outcomes with an equilibrium model of entry that addresses the endogeneity problem. There are two main reasons for our choice. First, the Bayesian approach is a unique tool that has the advantage of unifying parameter inference and decision making. After obtaining the posterior distribution of unknown parameters, the Bayesian approach properly incorporates the parameter posterior into the distribution of the post-entry outcomes. Then, the latter can be used for policy and business decision-making analyses (e.g., whether to introduce a regulation or whether to open a business in a market).¹ Second, unlike other approaches, such as the maximum likelihood estimation (MLE), Bayesian inference can be precise when the likelihood is computed via simulations. The simulated likelihood is not precise because of simulation errors—for example, in MLE, simulated errors are fixed (after the initial generation) throughout the optimization procedure. As a result, the difference between the true likelihood and the realized value of the simulated likelihood affects the MLE. In contrast, Bayesian estimation can overcome this problem. Andrieu and Roberts (2009) propose a pseudo-marginal approach that provides exact parameter posteriors even when the likelihood is simulated. A key difference from the simulated MLE is that the pseudo-marginal approach generates new simulated errors each time the likelihood is evaluated.²

However, technical difficulties have prohibited researchers from adopting the Bayesian approach to post-entry outcome analysis when equilibrium market structures are modeled to overcome endogeneity. In particular, a closed-form likelihood for a simultaneous-move entry game of complete information is unknown, and the evaluation of the likelihood requires time-consuming simulations. Even worse, the possibility of multiple equilibria forces us to find all equilibria for

¹ Refer to Savage (1972) and Rossi, Allenby & McCulloch (2005) for details.

² Refer to Flury and Shephard (2011) who illustrate the method.

each simulation. As a result, Bayesian analysis of a complete-information discrete game that addresses multiple equilibria requires immense computational power.

Despite its advantages, to the best of our knowledge, the Bayesian approach has never been used to estimate discrete games of complete information because of this computational complexity. Furthermore, even in a traditional (frequentist) setting, most studies have relied on the two-step (or four-step) estimation to overcome the computational burden (Mazzeo, 2002a; Zhu, Singh & Manuszak, 2009; Ellickson & Misra, 2012). This estimation procedure can be viewed as an extension of Heckman (1979). In the first step, entry game coefficients are estimated, and then these estimates are used to correct selection biases in the later steps. Multi-step estimations are computationally less demanding than Bayesian estimation, which requires repetitive evaluations of the joint likelihood of discrete games and post-entry equations. Despite its computational advantage, for parameter inference, multi-step estimation is known to perform more poorly than maximum likelihood estimation (MLE) because it does not fully exploit the full information contained in the joint likelihood (Nawata, 1994). Similar to MLE, Bayesian estimation fully exploits information from the data. As mentioned earlier, Bayesian analysis is the only methodology that unifies parameter inference and decision making.

In a simultaneous-move discrete game of complete information, the number of players in the game exponentially increases the number of strategy profiles and, thus, the computational complexity. To confront this issue, past studies have made arbitrary assumptions and limited the number of equilibrium strategy profiles. For example, Reiss and Spiller (1989) scaled back the problem by allowing, at most, one airline to enter a market. Mazzeo (2002a) considered six players and 15 entry configurations with limited heterogeneity.³ Zhu, Singh and Manuszak (2009) examined Wal-Mart, Kmart and Target, allowing for seven entry configurations. Our empirical application involves nine retail stores equaling $2^9=512$ entry configurations, which need to be checked for Nash equilibrium in each simulation draw. For example, if there are 1,000 markets,

³ There are two product types, each allowing up to three players, which makes 15 configurations.

the total number of strategy profiles evaluated via simulation (with 1,000 simulation draws) to compute a likelihood once becomes 512,000,000.

The entry game is assumed to be a simultaneous-move discrete game of complete information in which a firm’s profit (and, thus, its entry decision) is affected not only by market characteristics, but also by spillover effects from other firms’ entry decisions. We refer to these spillovers *as strategic effects* because they are caused by the endogenous entry decisions of other potential collocating firms.

This framework has several advantages. As is the case with most discrete games, this approach does not require revenue or price data because the observed actions of entry—the equilibrium outcome—can be mapped onto firms’ profits. Furthermore, by allowing flexible strategic effects, we are able to capture both the negative and positive effects of collocation. We also allow the strategic effects to be heterogeneous across firms. Finally, because our data are cross-sectional, the observed equilibrium outcome is the result of a steady-state, long-term equilibrium, in which firms have made adjustments regarding their choices (of entry). The complete-information structure of the game fits this setting. We emphasize that we do not shy away from the complete-information framework, despite the challenges of both multiple equilibria and heavy computational burden. The remainder of this section and Sections 2 and 3 describe the methods that we used to address these challenges.

In addition to the discrete game, we jointly estimate a sales model that corrects the endogeneity problem with regard to firm entry, allowing us to evaluate the pure causal effect of a firm’s presence on market sales. We supplement the market structure outcome with a sales model to jointly predict both the types of firms that will constitute a market and the total market sales.

There are a couple of challenges involved in the joint modeling (and estimation) of market structure and market sales. First, as with most discrete games, we face the problem of multiple equilibria, which makes it difficult to either define the likelihood for estimation or conduct accurate counterfactual policy simulations. As a result, past research has scaled back the problem (Bresnahan & Reiss, 1990; Berry, 1992); specified the sequence of moves (Berry, 1992; Mazzeo, 2002b); made arbitrary assumptions related to equilibrium selection (Hartmann, 2010); or adopted

a partial identification approach—i.e., estimated a range of parameters instead of point estimates (Ciliberto & Tamer, 2009). In this research, we address multiple equilibria by implementing the selection function method of Bajari, Hong, and Ryan (2010) to empirically estimate the equilibrium selection rule from the observed data. After obtaining the posterior of unknown parameters, the equilibrium selection rule pins down the distribution of future outcomes uniquely and, thus, is useful in predicting post-entry outcomes.

Second, to identify the pure causal effect of firm entry on total market sales, we must correct for the endogeneity issue—that is, do high expected market sales cause a firm to enter, or does firm entry cause high market sales? To address this problem, we include aggregate shocks in both our entry game and the sales model to isolate the pure causal effect of a specific firm’s entry on market sales.

As in the estimation of discrete games of complete information with multiple players, computational processing power is crucial, especially in our setting, where we not only take into account equilibrium selection, but also jointly estimate a model of market sales. To overcome the computational burden of our analysis, we rely on two computational techniques.

First, we use an improved simulator for the likelihood evaluation. A simulator randomly generates many instances of error terms and counts the number of instances that are consistent with a given observation. But if there are nine firms, as in our application, there are 512 strategy profiles. Then, to compute a likelihood for the 512 outcomes with reasonable precision, we need to generate a large number of simulated instances of error terms. Our simulator is designed to not generate error terms that are obviously inconsistent with a given observation (out of equilibrium). Thus, it accelerates the simulation process by significantly reducing the number of simulations. Previous studies, such as Bajari, Hong and Ryan (2010), often use importance sampling to reduce computation time. Importance sampling works well for many simulation-based frequentist estimators but does not avoid the imprecision of the simulated likelihood because importance sampling also relies on fixed simulations. Our simulator can be used to compute the exact posterior. As discussed earlier, the pseudo-marginal approach by Andrieu and Roberts (2009) performs exact Bayesian inference. The pseudo-marginal approach draws new simulated errors

each time the likelihood is evaluated, making importance sampling inapplicable. Acceleration of our simulator makes it possible to adopt the pseudo-marginal approach.

Second, we utilize the GPGPU technology, using multiple processing cores in a GPU of a graphics card to increase the computation speed of our model estimation. The parallel computing of a GPU has been used in various areas, including artificial intelligence research. In particular, evaluating a simulated likelihood for a complete-information game consists of solving many unrelated games for equilibrium and, thus, is a so-called “embarrassingly parallel problem.” Fast evaluation of simulated likelihoods is a key to our Bayesian estimation of the joint model of entry and post-entry outcomes. We observe that our estimation method utilizing GPGPU in a single graphics card runs 500 times faster than a method using traditional Matlab codes. While our estimation procedure takes approximately one day, a traditional method without parallel multi-core processing would produce the same result in 500 days. We believe that GPGPU will be widely used in estimating games of complete information in future studies.

We apply our model to shopping malls for our empirical application. Mall sales are affected by the presence of retail stores at the mall, but retail stores decide to enter the mall endogenously. In terms of empirical context, our paper shares similarities with the work of Vitorino (2012). However, there are several key differences. First, we incorporate a complete information game structure. Our data is cross-sectional, and thus the observed equilibrium outcome is the result of a steady-state long-term equilibrium. We believe the complete information game structure adequately captures this setting. Second, we use the Bayesian approach to estimate the equilibrium selection function and consider all equilibria in both parameter estimation and counterfactual simulations. Finally, we jointly examine both configuration and sales. Hence, our model can help a market developer predict not only market structure (retail configuration), but also total market (shopping mall) sales from a given location. This information can eventually be used to decide where to create a market—that is, in our empirical setting, to choose the location for a shopping mall.

The results indicate that population and income are the key forces that drive retail stores’ profits: upscale retailers prefer to locate in affluent and populated areas, while midscale stores

prefer to locate in lower-income and less-populated areas. We find that the negative effect of competition is prevalent within store categories, especially for discount and midscale stores, but positive agglomeration effects exist across store categories. We also find substantial competitive effects not only within, but also across, store categories. Consistent with the results of a mall’s store configuration, population and income are also the main drivers of total mall sales. Furthermore, our results suggest that, above and beyond the effects of market characteristics, midscale and upscale retailers have a positive causal brand effect, whereas discount stores have a negative causal brand effect, on mall sales. Our counterfactual simulations reveal that certain stores would decide not to enter due to the competitive entry of other stores, even when market conditions are favorable. We find that a developer is moderately better off choosing a mall location that has high average income versus large population. In terms of equilibrium selection, we find no evidence that the highest joint payoff equilibrium is more frequently selected—an assumption commonly used in the literature in both estimation and counterfactual policy simulations of discrete games.

The remainder of the paper is organized as follows. In Sections 2 and 3, respectively, we present the model and the estimation. Section 4 discusses the data and industry details for our empirical application. Section 5 discusses the estimation results and counterfactual analyses. Section 6 concludes.

2. Model

We build a joint model of firm entry and sales to examine the formation of market structure and its effect on total market sales. First, we model a firm’s decision to enter a particular market, with the firm’s payoff depending not only on firm and market characteristics, but also on other firms’ entry decisions. Second, along with the model of firms’ entry decisions, we utilize a model of total market sales that corrects the endogeneity issue associated with firm entry—that is, it address the question: do high expected market sales cause firm entry, or does firm entry cause high market sales? We combine the market structure model with a sales model to jointly predict the types of firms that would enter a specific market and the total market sales, clearly important

issues for businesses and policy makers. We jointly utilize these two models to address endogeneity in predicting total market sales.

2.1. Discrete Game of Firm Entry

We model the entry decisions of firms in a specific market as a simultaneous-move discrete game of complete information. Suppose that there is a sequence of markets, indexed by $m=1, \dots, M$. In each market, there are J potential entrants, and the profit of firm j when entering market m is defined as

$$\pi_{mj}(a_m) = \pi_{mj}(a_{mj} = 1, a_{m(-j)}) = \beta_j x_{mj} + \sum_{j' \neq j}^J \delta_j^j a_{mj'} + \varphi_j \xi_m + \varepsilon_{mj}, \quad (1)$$

where $a_m = (a_{m1}, \dots, a_{mJ}) \in \{0, 1\}^J$ is the vector of action profiles of all firms in market m , with $a_{mj} = 1$ if firm j enters and $a_{mj} = 0$ otherwise. Similarly, $a_{m(-j)}$ is the vector of action profiles for all firms in market m other than firm j . The vector x_{mj} represents the characteristics of firm j in market m , which may include market-level characteristics such as population, average income, and household size. The vector of parameters $\beta_j = (\beta_{j1}, \dots, \beta_{jK})$ represents the marginal effect of firm- and market-level characteristics on firm profit, and $\delta^j = (\delta_{j'}^j)_{j'=1, j' \neq j}^J$ represents the vector of strategic effects of other firms' entries on firm j 's profit. Notice that we do not restrict these strategic effects to be negative, but also allow them to be positive—that is, we allow both negative competitive effects and positive agglomeration effects (Ciliberto & Tamer, 2009; Vitorino, 2012).

The last two terms in Equation (1) are the unobserved components. The first term ξ_m is the market-level shock common to all firms in a specific market, and the second term ε_{jm} is the firm-market specific shock assumed to be independent across markets and firms. We assume that both terms follow a standard normal distribution and are observed by potential entrants but unobserved by the econometrician. It is important to recognize that the market-level shock common to all potential entrants in a market is the source of correlation among entrants' payoffs. The market-level shock can be thought of as any factors that influence all firms in a market. We denote $\varphi = (\varphi_1, \dots, \varphi_J)$ as the vector of marginal effects corresponding to market-level shocks—that

is, the degree to which market-level shocks affect firm j 's payoff. We normalize the profit of a firm not entering the market to be zero.

2.2. Multiple Equilibria

We assume that a pure strategy Nash equilibrium is observed in each market. Firm j enters market m if and only if $\pi_{mj}(a_{mj} = 1, a_{m(-j)}) > 0$. It is generally the case that a pure strategy Nash equilibrium is not unique in discrete games. Figures 1a and 1b offer an illustrative example similar to that shown in Bresnahan and Reiss (1990). Suppose that two firms, A and B, are playing a simultaneous-move entry game of complete information. For simplicity, assume that each player's payoff is dependent only on the strategic effect (entry choice) of the other player and an idiosyncratic shock. Thus, the profit function in Equation (1) has only the second and the last components. In such a case, the profits for firm A in market m would be $\pi_{mA} = -\delta_B^A a_{mB} + \varepsilon_{mA}$, and the profits for firm B would be $\pi_{mB} = -\delta_A^B a_{mA} + \varepsilon_{mB}$. If we assume that entry is competitive ($\delta > 0$), given a set of parameters, there would be a region in the $(\varepsilon_{mA}, \varepsilon_{mB})$ plane where one would observe more than one equilibrium outcome (shaded region in Figure 1a). Similarly, if the entry is complementary—that is, the profit function for firm A is $\pi_{mA} = \delta_B^A a_{mB} + \varepsilon_{mA}$, and that of firm B is $\pi_{mB} = \delta_A^B a_{mA} + \varepsilon_{mB}$ ($\delta > 0$)—there would be a region in which an equilibrium outcome could either be all firms entering or none entering (shaded area in Figure 1b).

<Figure 1>

Two potential problems are associated with the multiplicity of equilibria. First, because there are specific regions (shaded regions in Figures 1a and 1b) where no unique outcome exists, the model is termed incomplete (Tamer, 2003)—that is, we are not able to define the likelihood of certain types of outcomes. Second, with counterfactual policy simulations, we cannot accurately predict an outcome because we cannot determine which equilibrium is selected. To overcome this problem, researchers typically simplify the scale of the problem—e.g., use the number of firms rather than their identities (Bresnahan & Reiss, 1990); postulate a structure on the sequence of moves (Berry, 1992; Mazzeo, 2002b); set arbitrary assumptions with regard to equilibrium

selection (Hartmann, 2010); or rely on a partial-identification framework that estimates a range of parameters instead of point estimates (Ciliberto & Tamer, 2009). We mitigate the problem of multiple equilibria by empirically estimating—from the observed data—the selection rule proposed by Bjorn and Vuong (1984) and formalized by Bajari, Hong, and Ryan (2010).

The detailed process is as follows. Let Γ be the set of pure strategy Nash equilibria given the discrete game payoffs. Thus, the probability that profile a_m is played is

$$\rho(a_m; \Gamma) = \begin{cases} \frac{\exp(\kappa z(a_m))}{\sum_{a'_m \in \Gamma} \exp(\kappa z(a'_m))} & \text{if } a_m \in \Gamma \\ 0 & \text{otherwise} \end{cases}, \text{ where } z(a_m) = \begin{cases} 1 & \text{if } \Psi_m(a_m) = \max_{a'_m \in \Gamma} \Psi_m(a'_m) \\ 0 & \text{otherwise} \end{cases}.$$

The joint payoff of firms taking action profile a_m is represented as $\Psi_m(a_m) = \sum_j \pi_{mj}(a_m)$, and the parameter κ captures how often the highest joint payoff equilibrium is selected.⁴ For example, if $\kappa > 0$, the highest joint payoff equilibrium is more likely to be selected instead of other equilibria.

A discussion about the identification of κ will be presented in Section 3.2.

It is worth noting that the existing literature typically does not explicitly address equilibrium selection and makes arbitrary assumptions—e.g., that the highest total payoff equilibrium is always selected. As specified earlier, this research has two main objectives. First, we seek to provide projections about market structure—firm entry in the presence of other potential entrants—under specific market characteristics. Second, we utilize market structure information to predict market-level sales, informing market developers about where to create a market. Performing such tasks require counterfactual analyses—that is, simulations of future discrete game outcomes where an equilibrium selection rule should be specified to choose equilibrium and predict the market structure.

⁴ We have estimated other selection rules in Bajari, Hong, and Ryan (2010). The estimation results were essentially identical.

2.3. Sales Model

Along with a model to predict the market structure (type of firms in a specific market), to take into account the market developer’s objectives, we consider the following model of total sales for a particular market as

$$y_m = \lambda w_m + \sigma_\xi \xi_m + \sigma_\nu \nu_m, \quad (2)$$

where w_m is the vector of characteristics of market m , such as population and average income. We can also include dummy variables for firms (indicating their presence) in w_m . Market-level unobservable effects are captured by ξ_m and ν_m , both of which are assumed to come from the standard normal distribution. Furthermore, the two terms are assumed to be independent of each other, as well as independent of market-firm specific shocks ε_{mj} in firm j ’s profit function in Equation (1).

As ξ_m affects the entrant’s payoff, our model allows correlation between total sales of a market and its entrants’ profits, given all the observable variables. More specifically, the covariance between the error terms of an entrant firm j and the total sales of market m is

$$\text{cov}(\varphi_j \xi_m + \varepsilon_{mj}, \sigma_\xi \xi_m + \sigma_\nu \nu_m) = \varphi_j \sigma_\xi.$$

Next, we consider the implication of the correlation between total market sales and entrants’ profits. Suppose that there is a large market-level shock that affects both market sales and firms’ decisions to enter that market. Such shocks can be any unobservable (to the econometrician) elements not included in x_{mj} and w_m in equations (1) and (2), respectively. Thus, firm entry may be correlated with high sales because of the market-level shock, but not because an entrant draws more demand and boosts the total sales of the market. If the sales model in Equation (2) is run alone, the estimated effect of a firm’s entry will be biased because of the correlation between firm dummy variables and the regression error ($\sigma_\xi \xi_m + \sigma_\nu \nu_m$). Hence, we include the market-level shock ξ_m in both the entrant payoff in Equation (1) and the total market sales in Equation (2), and we correct for the endogeneity bias by jointly estimating the parameters of the discrete game and the sales model.

3. Estimation

We adopt a Bayesian approach because it provides a unified methodology for inference and decision. One of our goals is to propose a model that predicts post-entry outcomes that will help both policy makers and businesses. Through the Bayesian approach, we can properly reflect the parameter uncertainty including the equilibrium selection when evaluating the desirability of a choice (a location choice in our empirical setting). We have more than 100 parameters in our empirical exercise, and Bayesian inference on this large number of parameters turns out to be feasible.

In this section, we discuss how we compute the posterior distribution for parameter inference. We use the Metropolis algorithm, one of the Markov Chain Monte Carlo (MCMC) methods common in Bayesian analysis when direct sampling from the posterior distribution is not feasible.

3.1. Likelihood Evaluation for Parameter Inference

For parameter inference, we must compute the likelihood—the probability of observing the data given parameter values. Because the likelihood in our model does not have a closed-form solution, we use a frequency simulator similar to that of Ciliberto and Tamer (2009). We improve upon their simple simulator via methods we discuss in detail below. For notational simplicity, we will omit subscripts whenever the meanings are clear. Note that some components of w are whether firms enter a market (i.e., firm dummies). Let w' denote all the variables except for the firm dummies, such that $w = (w', a)$. Variables in w' include population, average income, etc. The joint probability of observing firms' entry decisions and market sales can be represented as

$$\begin{aligned} p(a, y | \theta, x, w') &= \int p(a | \theta, x, w', \xi) \cdot p(y | a, \theta, x, w', \xi) \phi(\xi) d\xi \\ &= \int p(a | \theta, x, \xi) \cdot p(y | a, \theta, w', \xi) \phi(\xi) d\xi, \end{aligned} \tag{3}$$

where $\phi(\cdot)$ is the density function of the standard normal distribution. The second equality holds because a is independent of w' conditional on (θ, x) , and y is independent of x given (a, θ, w') .

We examine the two elements in Equation (3) separately. First, the marginal probability of observing firms' entry decisions $a \in \{0,1\}^J$ in market m with a common shock ξ is

$$p(a | \theta, x, \xi) = \int \rho(a; \Gamma(\varepsilon)) \cdot \phi(\varepsilon) d\varepsilon, \quad (4)$$

where $\Gamma(\varepsilon)$ is the set of equilibria with explicit dependence on ε . Because there is no closed-form solution for the set of equilibria $\Gamma(\varepsilon)$, to evaluate this, we may rely on a simple frequency simulator such as

$$p(a | \theta, x, \xi) \approx \frac{1}{R} \sum_{r=1}^R \rho(a; \Gamma(\varepsilon^r)), \quad (5)$$

where $\varepsilon^1, \dots, \varepsilon^R$ are J -dimensional independent standard normal draws. We refer to the simulator in Equation (5) as the *simple simulator*. While the simple simulator can approximate the integral in Equation (4), it is not computationally efficient because a majority of the draws may not provide any information to help determine the likelihood.

Note that if one of the followings holds for at least one $j = 1, \dots, J$:

- i) $\varepsilon_j < -(\beta_j x_j + \sum_{j' \neq j} \delta_{j'}^j a_{j'} + \varphi_j \xi)$ and $a_j = 1$; or
- ii) $\varepsilon_j > -(\beta_j x_j + \sum_{j' \neq j} \delta_{j'}^j a_{j'} + \varphi_j \xi)$ and $a_j = 0$,

then $\rho(a; \Gamma(\varepsilon)) = 0$ because neither i) nor ii) can be in equilibrium and, thus, is not consistent with observation a . Let E be the event in which either condition i) or ii) holds for some $j = 1, \dots, J$.

Then, a more efficient simulator can be written as

$$\begin{aligned} & p(a | \theta, x, \xi) \\ &= \Pr(E^c) \int \rho(a; \Gamma(\varepsilon)) \cdot \phi(\varepsilon | E^c) d\varepsilon + \Pr(E) \int \rho(a; \Gamma(\varepsilon)) \cdot \phi(\varepsilon | E) d\varepsilon \\ &= \Pr(E^c) \int \rho(a; \Gamma(\varepsilon)) \cdot \phi(\varepsilon | E^c) d\varepsilon \\ &\approx \Pr(E^c) \frac{1}{R} \sum_{r=1}^R \rho(a; \Gamma(\varepsilon^r)), \end{aligned}$$

where E^c is the complement of E and $\varepsilon^1, \dots, \varepsilon^R$ are drawn from the density $\phi(\varepsilon | E^c)$. Note that $\phi(\varepsilon | E^c)$ is the density of a truncated standard normal distribution. To draw $\varepsilon = (\varepsilon^1, \dots, \varepsilon^R)$, we

draw each ε_j independently from a truncated standard normal, where the truncation level is set at

$$-\left(\beta_j x_{mj} + \sum_{j' \neq j} \delta_{j'}^j a_{mj} + \varphi_j \xi_m\right)$$

and the truncation direction (left or right) is determined by a_{mj} . The term $\Pr(E^c)$ is computed by the following product:

$$\Pr(E^c) = \prod_{j=1}^J \left[\Phi \left(-\beta_j x_j - \sum_{j' \neq j} \delta_{j'}^j a_{j'} - \varphi_j \xi_m \right) \right]^{1-a_j} \left[1 - \Phi \left(-\beta_j x_j - \sum_{j' \neq j} \delta_{j'}^j a_{j'} - \varphi_j \xi_m \right) \right]^{a_j},$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

By avoiding simulation draws that give an obvious $\rho(a; \Gamma(\varepsilon)) = 0$, our improved simulator performs more efficiently than the simple simulator in Equation (5). For example, assume that $\beta_j x_j + \sum_{j' \neq j} \delta_{j'}^j a_{j'} + \varphi_j \xi = 0$ for all $j=1, \dots, J$ with $J=9$ —that is, the entry decisions of nine firms depend only on the draws of each ε . In addition, assume that $a=(1, \dots, 1)$ is observed—that is, all firms enter. Thus, the event E^c is mapped onto the positive orthant of the J -dimensional Euclidean space of ε . Via the simple simulator in Equation (5), a draw falls in E with probability $1 - \frac{1}{2^9} \approx 0.998$. Hence, only two out of 1,000 draws will, on average, lie in E^c and determine the value of the likelihood, whereas in our improved simulator, all 1,000 draws will lie in E^c .

Next, we discuss the computation of $\Gamma(\varepsilon^r)$, the set of all pure strategy Nash equilibria. To find all equilibria, we check whether each strategy profile is in equilibrium. For each market and each draw of ε^r , there are 2^J strategy profiles to check for equilibrium. For example, if there are $J=9$ potential entrants in 1,000 markets, and if we use 1,000 random draws of ε^r for simulation, we would need to check $2^9 \times 1,000 \times 1,000 = 512$ million cases for equilibria in each MCMC step, which would not be feasible using conventional computational methods. Because all of these cases in a given MCMC step can be checked in parallel, we capitalize on the parallel-processing power of the

GPGPU,⁵ a state-of-the-art technology that uses a graphics processing unit and its many cores to implement computation.

The second part of the likelihood in Equation (3) is given by

$$p(y | a, \theta, w', \xi) = \frac{1}{\sigma_\nu} \phi\left(\frac{y - \lambda w - \sigma_\xi \xi}{\sigma_\nu}\right).$$

The details of our estimation procedure are presented in the appendix.

3.2. Identification

We briefly discuss the intuition behind the variation in the data that helps one make inferences on the parameters. For detailed arguments, see Bajari, Hong, and Ryan (2010). The idea is based on identification at infinity. We omit the market index m here. For each a_j , we can find large values of x such that playing a_j is a dominant strategy of all firms $-j$ with probability close to 1. For these x values, a small variation in x_j identifies β_j . Then, find x and x' such that $\beta_j x = \beta_j x'$ and a_j differ only on a_j . The observed distribution of a_j on x and x' identifies δ . Given β and δ , variation in φ induces variation in correlation among firms' payoffs and, hence, their entry decisions. Thus, when the equilibrium of the discrete game is unique, β , δ and φ are identified.

If equilibrium is not unique, the equilibrium selection plays a critical role in parameter estimation. Given values of the parameters (β, δ, φ) , observations of markets with multiple equilibria will identify the selection probability $\rho(a_m; \Gamma)$, in particular, κ . Finally, the correlation between entry decisions $a_m = (a_{m1}, \dots, a_{mJ})$ and y_m helps identify σ_ξ . The parameters λ and σ_ν are identified by variations in y_m and w_m .⁶

4. Empirical Application: Industry Details and Data

⁵ Some papers use graphics processing units to improve computation time. For example, Aldrich et al. (2011) solve macroeconomic models, and Durham and Geweke (2014) propose a posterior simulator for Bayesian estimation. To the best of our knowledge, we are the first to implement the GPGPU technology in the estimation of discrete games.

⁶ We have verified parameter identification with simulated datasets randomly generated according to our model.

We apply our model to shopping malls for our empirical application. This is a natural setting in which to illustrate the benefits of our model, as mall sales are affected by the presence of retail stores at the mall, but retail stores decide to enter the mall endogenously. We first discuss industry details and then explain the data we use for empirical analyses.

The retail sector constitutes one of the largest segments of the U.S. economy, generating sizeable annual sales that considerably bolster total GDP. In 2015, retail sales totaled \$4.7 trillion in the U.S. alone (approximately 26 percent of GDP) and nearly \$21 trillion globally,⁷ and the market continues to grow. Despite the recent e-tail surge, traditional brick-and-mortar stores remain the core of the retail industry. In 2015, U.S. retail e-commerce sales accounted for only 7.3 percent of total U.S. retail sales. A considerable proportion of brick-and-mortar retail in both developed countries, such as the United States, and newly industrialized countries, such as China, involve a market structure typically referred to as a shopping mall or a shopping center. According to government statistics, China's retail sales have more than quadrupled in a ten-year span, from 6.7 trillion yuan in 2005 to 30 trillion yuan in 2015,⁸ creating one of the world's largest consumer markets. In an effort to keep up with this explosive increase in consumption, global shopping-center development is increasing rapidly. According to a CBRE report, in 2015, 39 million square meters of new mall space was under construction worldwide, with nine out of the ten most active locations in China. Emerging markets such as Istanbul, Bangkok, Moscow, Abu Dhabi, and Kuala Lumpur also were highly active.⁹

Despite this massive commercial activity, there has been very limited research to guide developers in deciding where to build a shopping mall—i.e., create a market. We apply our methodology to the development of a shopping mall to provide insight into how a developer can choose a mall site. For example, should a developer construct a mall in a highly populated area or

⁷ Worldwide Retail Ecommerce Sales: The eMarketer Forecast for 2016, [http://totalaccess.emarketer.com/Reports/Viewer.aspx?R=2001849&dsNav=Ntk:basic%7cglobal+retail+sales%7c1%7cRo:4,Nr:NOT\(Type%3aComparative+Estimate\)](http://totalaccess.emarketer.com/Reports/Viewer.aspx?R=2001849&dsNav=Ntk:basic%7cglobal+retail+sales%7c1%7cRo:4,Nr:NOT(Type%3aComparative+Estimate)).

⁸ National Bureau of Statistics of China, "Total Retail Sales of Consumer Goods in December 2015", January 20, 2016, http://www.stats.gov.cn/english/PressRelease/201601/t20160120_1307123.html.

⁹ CBRE, "Global Shopping Centre Development", April 2015, <http://www.cbre.com/research-and-reports/global-shopping-centre-development>.

in an affluent area? This seemingly simple question turns out to be quite complicated, as it involves the endogenous decisions of various retail stores, the key constituents of a shopping mall. To predict the financial outcome of a mall in a particular location, we must anticipate what types of retail stores would join if the mall were developed. In addition to mall configuration (market structure), we have to ascertain the causal effect of a specific retail store on total mall sales to evaluate the expected payoff to the developer.¹⁰

A typical shopping mall consists of a large cluster of retail stores located in physical proximity, sharing amenities such as restrooms, food courts, and customer parking. Naturally, physical proximity of collocation has both benefits and costs. The benefits include an economy of scale achieved by sharing amenities, as well as increased overall demand from reduced consumers' transportation costs from one-stop shopping. The obvious cost comes from competition from other retail stores located in the vicinity.

We focus our attention on anchor stores because they are, by definition, the key tenants in a mall, occupying most of the mall's gross leasable area (GLA) and generating much of the foot traffic (see Figure 2 for an example of a mall layout in terms of GLA). Anchor stores are competing national department stores or retail chains such as Nordstrom, Macy's, and Sears.

<Figure 2>

Our mall configuration data come from the Directory of Major Malls, a data provider that supplies information about shopping centers and their tenants operating in the U.S. General-purpose shopping centers are organized by size and trade area into the following categories: strip/convenience center; neighborhood center; community center; regional mall; and super-regional mall (see Table 1 for the definition of different types of shopping hubs by The International Council of Shopping Centers—ICSC, 2016).

<Table 1>

¹⁰ We define a shopping mall as a market and, thus, in generalizable terms, market structure and market sales denote, in our empirical context, mall configuration and mall sales, respectively. Hereafter, we use these terms interchangeably.

According to the definition in Table 1, shopping hubs, generally referred to as shopping malls, are either regional or super-regional. We define these malls as separate markets in which individual retail stores locate, a reasonable assumption because both regional and super-regional malls are independent shopping hubs with the space and capacity to provide a wide range of goods and services, operating in a large and separate geographical trade area. Therefore, we focus our analysis on regional and super-regional shopping malls and their tenant anchor stores.

As previously mentioned, anchor stores are large national chain stores that drive the majority of customer traffic to a shopping mall. These retailers are typically classified into three broad categories—discount, midscale, or upscale—based on target customer segments (J.D. Power and Associates, 2007; Levy & Weitz, 2012). Focusing on price rather than service, discount department stores, such as Sears and Target, sell a variety of merchandise at lower prices than typical retail stores. Midscale department stores, including Macy’s and Dillard’s, offer a wide selection of both brand-name and non-brand-name merchandise, seeking to offer good value to their customers. Upscale department stores, such as Nordstrom and Bloomingdale’s, sell goods at above-average prices; their customers typically prefer exclusive designer brands and value customer service over low price (see Table 2 for a detailed categorization of each department store).

<Table 2>

In addition to mall configuration, we obtain demographic data from the Scan/US demographic database, which includes population, size of household, and average household income within five miles of each shopping mall. Furthermore, taking the location of each anchor store’s headquarters, we compute the distance to every shopping mall in which it has a store. After data cleanup by excluding unusable observations that are missing information, we utilize 1,188 regional and super-regional shopping malls with a total of 6,500 anchor stores for our empirical analysis. The summary statistics of the key variables are presented in Table 3.

<Table 3>

5. Results

We discuss, first, the results of each firm’s profit function in Equation (1) and, then, the results of the sales model in Equation (2). We then conduct several counterfactual policy simulations.

5.1. Firm Profits

Table 4 shows the parameter estimates with regard to market- and firm-specific effects, and Table 5 shows estimates of the strategic spillover effects of the discrete game of firm entry; combined, they represent the retail store’s profit function in Equation (1). There is clearly a substantial degree of heterogeneity via stores. We discuss in detail some noticeable patterns according to store categories (discount, midscale, and upscale).

<Table 4> and <Table 5>

The constants are negative except for Macy’s and other midscale stores. This indicates that, given market characteristics, midscale stores are more likely to enter a specific market when not accounting for strategic effects of competitors’ entry.¹¹

The effect of population is positive for discount and upscale stores, indicating that these store types would prefer to locate in populated areas. In contrast, midscale stores would rather locate in less-populated areas. The offerings of midscale stores typically consist of products with quality similar to that at upscale stores, but with less service and at lower prices. Hence, midscale stores would find densely populated areas too costly (due to high labor and rental costs) to effectively operate in those areas.

For upscale stores, such as Nordstrom and Bloomingdale’s, the inclination to enter populated areas is not surprising. Upscale stores, by definition, service upper-income segments of the population, which exist in critical mass only in highly populated areas. Correspondingly, parameters associated with average income (a proxy for purchasing power) indicate that Nordstrom and Bloomingdale’s prefer to locate in affluent neighborhoods. On the other hand, the

¹¹ We will refrain from interpreting insignificant parameters.

negative effect on average income implies that most discount and midscale stores prefer to locate in less-affluent areas because they cater to low- to mid-income customers.

The effect of household size is positive and significant for midscale stores, suggesting that large households with many family members appreciate the good value-per-price of midscale department stores. The effect of site size is positive for all midscale and upscale firms. Customers who shop in midscale and upscale firms value not only merchandise shopping but also other amenities, such as restaurants, cafes, movie theaters, valet parking, etc. Larger shopping malls have the space to provide more such amenities.

The variable distance-to-HQ is insignificant for most firms, suggesting that there are no economic benefits from locating closer to a firm's headquarters. All of our sample stores—anchor stores—are large chain department stores, which typically have achieved economies of scale in various dimensions and operate many distribution centers across the U.S. Hence, for most firms, distance to a firm's headquarters does not seem to have a significant effect on the strategic choice of whether to operate in a particular market. Other market-level factors, such as population and purchasing power, seem to be more important when firms make strategic decisions to enter a specific market.

The market-level shocks represent any unobserved (by the econometrician) factors that influence firms in a particular market. Examples of such shocks include convenient transportation (highway access) and the presence (or proximity) of other attractions, such as amusement parks. These common market-level shocks seem to have a strong influence on midscale department stores but a lesser one on upscale department stores.

The coefficient with regard to equilibrium selection is negative and statistically insignificant, providing no evidence that the highest joint payoff equilibrium is more frequently selected. This finding contradicts the previous literature, which commonly assumes that the highest joint payoff equilibrium is chosen in both estimation and counterfactual policy simulations. In fact, our results actually show suggestive evidence (a negative but insignificant parameter estimate with regard to equilibrium selection) that the highest joint payoff equilibrium is less commonly selected compared to other equilibria, implying stores do not seem to coordinate to achieve the highest joint payoff.

The parameter estimates for the strategic effects of firm entry, shown in Table 5, are consistent with our earlier inference with regard to market- and firm-specific effects—for example, midscale stores’ reluctance to enter highly populated areas due to high operating costs. One can see that competition is the dominant effect within store categories, especially for midscale firms. Furthermore, midscale firms suffer from the entry of upscale firms, indicating competitive effects not only within store categories but also across store categories. We would expect this result because product offerings of midscale firms are comparable to those of upscale firms, but with less service at lower prices. In contrast, there is a positive agglomeration effect between discount and midscale firms. The product offerings of discount stores are typically different from those of midscale stores, and customers take advantage of one-stop shopping when these types of firms collocate. We will discuss these strategic effects in more detail when we perform our counterfactual policy simulations in Section 5.3.

5.2. Sales Model

Table 6 shows the results of the sales model in Equation (2). Consistent with the findings from the discrete game of retail store entry, we find that population and income are the key drivers of total mall sales. The main objective of Equation (2) is to infer the causal effect of each anchor store’s presence on total mall revenue. In order to precisely extrapolate the causal effect, we need to address the endogeneity issue associated with store entry. That is, we have to ask: do high expected mall sales cause a retail store to enter, or does a store’s entry cause high mall sales? This question is especially critical for a developer who seeks to evaluate the causal effect of a specific department store’s presence on total mall sales.

<Table 6>

As previously explained, we address the endogeneity issue through the common unobserved term ξ_m in equations (1) and (2). Table 6 demonstrates that midscale and upscale stores have a positive causal effect on total mall sales. The positive causal effect can be due to two possible factors. First, the direct sales of a particular anchor store above and beyond the demand

characteristics of the mall—that is, any effect on sales after controlling for market characteristics such as population and income. Second, the positive (or negative) spillover effect of an anchor store’s presence on total mall sales. Note that the spillover effect referred to here is different from the strategic spillover effect in Equation (1). The former explains spillovers between anchor stores (competition and agglomeration), whereas the latter describes spillovers from a single anchor store to other smaller stores in a mall. Anchor stores attract mall customers through their brand and, as a result, attract smaller stores, which increases total mall sales. We refer to the combination of these two effects as *causal brand effects*.

Note that there is a negative causal effect for all discount stores. Because of harsh price competition, the presence of discount anchor stores discourages smaller stores from entering and, as a result, can have a diminishing effect on mall revenue.

To illustrate the importance of our joint model (of firm entry and total market sales) that controls for the endogeneity of store entry, Table 7 shows the result of a sales model without common market-level shocks that affect store entry in a particular market—that is, without the presence of common shock ξ_m in equations (1) and (2). This result would be obtained by simply running an OLS (ordinary least square) regression of Equation (2). We can see that there is a considerable difference between the results in Table 6 and Table 7. Although the directions of the parameter estimates are similar, most estimates become smaller in magnitude and statistically insignificant when the common shock is not included. For instance, consider the only significant firm-dummy variable, that of Nordstrom. The positive effect of Nordstrom’s entry is smaller than that in Table 6. Because any unobserved market-level shock that jointly affects firm entry and market sales—such as attractions or transportation convenience—can positively influence Nordstrom’s likelihood of entry, an OLS regression will underestimate Nordstrom’s true causal effect on total mall sales.

<Table 7>

5.3. Counterfactual Analyses

Using the structural parameters estimated in the previous section, we examine the marginal effect of market characteristics on market structure and total market sales to guide the developer. For this analysis, we first predict market structure—that is, compute the probability of each firm entering a market, given market characteristics—and then compute the expected sales conditional on the configuration of market structure.

For example, a developer needs to select a site for a new shopping mall. The choice of a site m is summarized as c_m , where c_m contains site characteristics (x and w') such as population and average income. The developer will choose the site that generates the highest total expected sales.¹² Formally, let C be the set of choices. Then, the developer maximizes $E[y | c]$ by choosing $c \in C$ such that

$$E[y | c] = \iint E[y | c, \theta, \varepsilon, \xi] d\Phi(\varepsilon, \xi) dp(\theta | data),$$

where $\Phi(\cdot)$ is the cumulative distribution function of an independent multivariate standard normal random vector, and $p(\theta | data)$ is the posterior distribution of the model parameters where $data$ represents all the data used in our model estimation. To compute the inner integral, given θ , we simulate (ε^r, ξ^r) for $r=1, \dots, R$. For each (ε^r, ξ^r) , we compute the set of all pure strategy Nash equilibria $\Gamma(\varepsilon^r)$. If there exist multiple equilibria, potential entrants' entry decisions, a^r , are drawn according to the equilibrium selection rule $\rho(a; \Gamma(\varepsilon^r; \theta))$. This step is necessary to determine $w = (c, a)$ because w includes firm dummy variables a . We use notation $w(c, a)$ to denote w 's dependence on a as well as on c . This procedure gives us the value of $E[y | c, \theta, \varepsilon, \xi] = \lambda w(c, a^r) + \sigma_\xi \xi$, where ν in Equation (2) is ignored because its expectation is zero. Hence, we can compute the inner integral as

¹² We assume that marginal cost is the same for all locations and that the developer's profits are proportional to total mall sales. Then, higher sales imply higher profits. With additional information on cost structures, it is easy to extend our analysis to more-general cases.

$$\int E[y | c, \theta, \varepsilon, \xi] d\Phi(\varepsilon, \xi) \approx \frac{1}{R} \sum_{r=1}^R \lambda w(c, a^r) + \sigma_\xi \xi^r. \quad (6)$$

Finally, recall that we have samples of θ , $\{\theta_1, \dots, \theta_S\}$, drawn from the posterior distribution, $p(\theta | data)$ through our MCMC draws. Because each sample of θ gives us one value of Equation (6), we can compute expected sales, given market characteristics c , as

$$E[y | c] \approx \frac{1}{S} \sum_{s=1}^S \int E[y | c, \theta, \varepsilon, \xi] d\Phi(\varepsilon, \xi) \approx \frac{1}{SR} \sum_{s=1}^S \sum_{r=1}^R \lambda w(c, a^r) + \sigma_\xi \xi^r.$$

Figure 3 shows the probability of selected firms entering, conditional on different population levels. Consistent with the results of firms' profit function in Table 4, we see that while the probability of entry for midscale stores, specifically Dillard's and other midscale stores, decreases with population, the entry probability of upscale stores, specifically Nordstrom and Bloomingdale's, increases. However, contrary to Table 4, Macy's entry probability actually increases with population. What explains this difference?

We should note that a firm's entry probability is a function of three components: the main effect from firm- and market-specific characteristics (Table 4); the within-category spillover effects (diagonal elements in Table 5); and the cross-category spillover effects (off-diagonal elements in Table 5). As a result, even though Macy's main effect with regard to population is negative, other midscale stores' effect due to population is also negative and, thus, discourages Macy's competitors from entering the market. Less competition increases Macy's likelihood of entry. Furthermore, discount stores are more likely to enter as population increases. Macy's benefits from positive cross-category spillover effects, further increasing its probability of entry.

Similarly, Nordstrom's and Bloomingdale's entry decisions depend not only on the main effect of population, but also on positive within-category spillover effects. We can see the amplifying effect reflected in the convex relation between population and the entry probability of Nordstrom and Bloomingdale's shown in Figure 3.

<Figure 3>

To gain deeper insights into the spillover effects, consider the results of another analysis in which we examine the correlation between site size and firms' probability of entry (Figure 4).

Consistent with the main effects in Table 4, the entry probability of Target and other discount stores decreases as mall size increases. However, inconsistent with the (negative) main effects, Sears' probability of entry seems to increase. Again, recall that our counterfactuals take into account not only the direct effect, but also the equilibrium behavior of other firms and, thus, the spillover effect.

<Figure 4>

We now examine the spillover effect more closely. Figure 5 separates the effect of site size on Sears' profits into the direct (main) effect and the indirect (spillover) effect from other firms. The thick downward sloping dotted line represents the main effect of site size on Sears' profits, and the thin lines represent the eight spillover effects from other firms—that is, the change in Sears' profits due to the entry of other firms. The thick, upward-sloping dashed line shows the aggregate of all spillover effects. Finally, the thick solid line represents the sum of the main and overall spillover effects. Even though the main effect is negative, the overall effect of site size on Sears' profits (and, therefore, the probability of entry) is positive because the positive spillover effects are greater than the negative main effect.

<Figure 5>

Having explained all the working parts of our model, we now ask the question of where to create a market—that is, where to develop a mall. Table 8 shows realistic scenarios facing a developer choosing a mall location. Site M possesses the average values of population and income, as well as the average values of other variables in our data. The population of site A is ten percent higher than that of site M (with the same average income as site M). Similarly, the average income of site B is ten percent higher than that of site M (with the same population as site M). Other market characteristics in sites A and B are set to the mean values in our data. We compare the effects of these changes in population and income in sites A and B to determine which variable has a higher impact on mall sales.

<Table 8>

Once again, because mall configuration outcome is a result of complex effects—main effect, within-category effect and cross-category effect—the final outcome may be different from the

direction of the main effects in Table 4. For example, although the main effect of population is positive and highly significant for Target, the probability of Target’s entry actually falls as population increases (compare Target’s entry probability for site M and site A). This is due to the fact that in highly populated areas, the likelihood of entry increases for other discount stores, Target’s main competitors. Because of the negative competitive effect indicated by the diagonal elements in Table 5, Target’s profits (and, thus, the probability of entry) decrease with an increase in population.

The results of the mall configuration simulation suggest that more discount stores enter site A and more midscale and upscale stores enter site B, resulting in a higher number of anchor stores entering site B. Because of the greater number of stores, along with the negative casual brand effect of discount stores (discouraging smaller stores from entering), site B (with higher income) outperforms site A (with larger population) by \$14 million in total mall revenue.

6. Conclusion

We propose a model of post-entry outcomes that includes many working components resembling real-world settings. Key elements of the model are as follows. First, we jointly estimate a discrete game of complete information with a sales model to identify the pure causal effect of a firm’s entry on market sales. As a result, within- and across-firm category spillover effects, as well as market spillover effects, are modeled. Second, we adopt the Bayesian approach, which combines decision making with parameter inference for estimation and counterfactual analyses. Finally, we utilize an improved simulator for likelihood evaluation and a state-of-the-art technology, GPGPU, using multiple processing cores of a graphics processing unit to significantly increase computational speed to consider and solve all equilibria in our Bayesian estimation—an effort that would not have been feasible with conventional computational methods.

We apply our model to a shopping mall setting in which mall sales are affected by the presence of retail stores at the mall, but retail stores decide to enter the mall endogenously. In spite of the recent surge in e-commerce, brick-and-mortar retail, specifically in the form of large-scale shopping malls, is still the dominant venue for consumer purchases in the developed world.

Furthermore, coinciding with recent growth in real estate development in newly industrialized countries such as China, the construction of mass-scale shopping malls has experienced tremendous growth. Yet there is little research about the dynamics governing the formation of the retail cluster—that is, the types of stores that will join a shopping mall—and, most importantly, the overall profitability of the mall. Our analysis helps us assess (1) the types of stores that will join a shopping hub; and (2) the expected financial performance of the shopping mall given its location and store configuration.

We find that population and income are the key drivers of retail stores' profit: upscale stores locate in highly populated, affluent areas, whereas midscale stores locate in less-populated, lower-income areas. Furthermore, both midscale and upscale stores prefer to locate in large shopping malls. Our analysis of strategic effects suggests that the negative effect of competition is the dominant force *within* store categories but that positive agglomeration effects exist *across* store categories. We also find substantial competitive effects both within and across store categories.

We find that population and income are highly correlated with mall sales and that midscale and upscale stores have a positive causal brand effect above and beyond the effects of market characteristics. Conversely, discount stores have a negative causal brand effect on mall sales. Our counterfactual simulations suggest that some stores would decide not to join a mall, despite favorable market conditions, due to the expected competitive entry of other stores. Finally, we find that a developer is moderately better off choosing a mall location that has high average income versus a large population.

Regarding equilibrium selection, we find no direct evidence that the highest joint payoff equilibrium is more frequently selected. In contrast, we find suggestive evidence that the highest joint payoff equilibrium is less commonly selected than other equilibria, challenging the common assumption made in the previous literature.

In summary, this research provides a rigorous yet practical framework to understand and evaluate firm entry in a market and its consequences for total market sales. Although our empirical application is in the retail shopping mall domain, our model can be extended and applied to a number of settings in which a decision maker must choose among alternative sites to

construct a market—for example, for transportation hubs such as airports or train stations. In addition, our modeling framework can be applied to assessing the impact of regulatory factors on firms' entry decisions and overall sales of site developers to gauge their implications for consumer welfare. We believe that these substantive areas will be exciting venues for future research.

Appendix: Estimation Details

This appendix provides the details of our Bayesian estimation procedure. To obtain the posterior distribution, we draw samples of θ by the Metropolis algorithm. At each iteration t , the Metropolis algorithm draws θ' from a proposal distribution and determines if θ' is accepted as a posterior sample θ_t with probability $\max\{1, p(\theta')/p(\theta_{t-1})\}$, where $p(\theta)$ is the posterior distribution of parameters. If θ' is not accepted, we set $\theta_t = \theta_{t-1}$.

Recall that $p(\theta) = (\text{prior density at } \theta) \cdot (\text{likelihood at } \theta)$. Our choice of prior distribution for all parameters is set as the independent joint normal distribution with mean 0 and standard deviation 10. Because the likelihood does not admit a closed form, we rely on the simulated likelihood explained in Section 3.

Our choice of the proposal distribution is a random walk. The proposal distribution at iteration t is normal with mean vector θ_t and variance matrix s^2V , where s is a positive number and V is a positive definite matrix. We tune the variance matrix to improve the performance of the Metropolis algorithm, a variation of Roberts & Rosenthal (2009) and Haario, Saksman, and Tamminen (2001). For s , we set our target acceptance probability between 0.1 and 0.3. If the number of accepted proposals in the last 100 iterations is below 10 or above 30, we adjust s accordingly. We set

$$V = \frac{(2.4)^2}{(\text{dimension of } \theta)} (\text{covariance of } t-1 \text{ samples}) + e(\text{identity matrix})$$

at iteration t . Here, e is a very small positive number such that the second term guarantees positive definiteness of V . During the initial tuning stage, we run 100,000 Metropolis iterations. To evaluate the simulated likelihood, $R=128$ errors are generated once, and the same errors are used throughout all of the Metropolis iterations.

During the main MCMC stage, s and V are fixed. We set the matrix V equal to the one obtained from the initial tuning stage, and s is set to 0.6, which we determined by trial and error. For the main MCMC stage, $R=1024$ errors are simulated and fixed throughout the Metropolis

iterations.¹³ One million Metropolis iterations are run. Every 1000 samples are recorded and the first half are discarded as burn-in. Thus, we have 500 samples for parameter inference and counterfactual analyses. We have experimented with various numbers of iterations and sets of simulated errors, but the results were not qualitatively different.

We run the estimation procedure with Matlab R2016b on 64 bit Windows 7, a desktop computer with Intel Core i7-6700K @4GHz, RAM 64GB, and a graphics card, GeForce GTX TITAN Black 6GB. The graphics card has 2880 CUDA (Compute Unified Device Architecture) cores that may be viewed as parallel processors. The graphics card executes our OpenCL kernel code to compute the simulated likelihood on these cores in parallel. The OpenCL kernel code execution is programmed to be controlled by our C++ code, which is called by Matlab when evaluating the likelihood.

As our dataset contains more than 1,000 shopping malls and $R=1024$ errors are simulated, an evaluation of the simulated likelihood requires solving for Nash equilibria of more than one million complete information entry games. All of these games are solved in parallel by our graphics card. The main MCMC stage that evaluates the likelihood one million times took 20 hours.

¹³ Preliminary analyses show that the results from the pseudo-marginal approach are similar to the fixed simulated error approach. We report the results from the latter.

References

- Aldrich, E.M., Fernández-Villaverde, J., Gallant, A.R. & Rubio-Ramírez, J.F. (2011). Tapping the supercomputer under your desk: Solving dynamic equilibrium models with graphics processors. *Journal of Economic Dynamics and Control*, 35(3), 386-393.
- Andrieu, C., & Roberts G. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. *The Annals of Statistics*, 37(2), 697-725.
- Bajari, P., Hong, H., & Ryan, S. P. (2010). Identification and estimation of a discrete game of complete information. *Econometrica*, 78(5), 1529-1568.
- Berry, S.T. (1992). Estimation of a model of entry in the airline industry. *Econometrica*, 60(4), 889-917.
- Bjorn, P.A., & Vuong, Q.H. (1984). Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation. Working paper 537, California Institute of Technology.
- Bresnahan, T.F., & Reiss, P.C. (1990). Entry in monopoly markets. *The Review of Economic Studies*, 57(4), 531-553.
- Ciliberto, F., & Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6), 1791-1828.
- Durham, G., & Geweke, J. (2014). Adaptive sequential posterior simulators for massively parallel computing environments. *Bayesian Model Comparison (Advances in Econometrics Vol. 34)*; Jeliazkov, I., Poirier, D.J., Eds, 1-44.
- Ellickson, P. B., & Misra, S. (2012). Enriching interactions: Incorporating outcome data into static discrete games. *Quantitative Marketing and Economics*, 10(1), 1-26.
- Flury, T., & Shephard, N. (2011). Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models. *Econometric Theory*, 27(5), 933-956.
- Haario, H., Saksman, E. & Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7(2), 223-242.
- Hartmann, W.R. (2010). Demand estimation with social interactions and the implications for targeted marketing. *Marketing Science*, 29(4), 585-601.
- Heckman, J. J. (1979). Sample Selection Bias as a Specification Error, *Econometrica*, vol. 47 (1). 153-161.

- ICSC (The International Council of Shopping Centers) (2016). U.S. shopping-center classification and definitions. Accessed October 10, 2016, http://www.icsc.org/uploads/research/general/US_CENTER_CLASSIFICATION.pdf
- J.D. Power and Associates Reports (2007). Department Store Experience Study. Available at businesscenter.jdpower.com/news/pressrelease.aspx?ID=2007239.
- Levy, M., & Weitz, B.A. (2012). *Retailing Management*, 8th ed., McGraw-Hill Education.
- Mazzeo, M. J. (2002a). Competitive outcomes in product-differentiated oligopoly. *Review of Economics and Statistics*, 84(4), 716-728.
- Mazzeo, M.J. (2002b). Product choice and oligopoly market structure. *The RAND Journal of Economics*, 33(2), 221-242.
- Nawata, K. (1994). Estimation of sample selection bias models by the maximum likelihood estimator and Heckman's two-step estimator. *Economics Letters*, 45(1), 33-40.
- Reiss, P., & Spiller, P. (1989). Competition and entry in small airline markets. *Journal of Law and Economics*, 32(2), 179-202.
- Roberts, G.O. & Rosenthal, J.S. (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2), 349-367.
- Rossi, P., Allenby, G. & McCulloch, R. (2005). *Bayesian Statistics and Marketing*, Wiley.
- Savage, L.J. (1972). *The foundations of statistics*, Courier Corporation. Su, C., & Judd, K. L. (2012). Constrained Optimization Approaches to Estimation of Structural Models. *Econometrica*, 80(5), 2213-2230.
- Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies*, 70(1), 147-165.
- Vitorino, M.A. (2012). Empirical entry games with complementarities: An application to the shopping center industry. *Journal of Marketing Research*, 49(2), 175-191.
- Zhu, T., Singh, V., & Manuszak, M. D. (2009). Market structure and competition in the retail discount industry. *Journal of Marketing Research*, 46(4), 453-466.

Table 1: U.S. Shopping Center Classification and Characteristics

Type of Shopping Center	Concept	Average Size (Sq. Ft.)	Typical GLA* Range (Sq. Ft.)	Typical Number of Tenants	Trade Area Size
Super-Regional Mall	Similar in concept to regional malls, but offering more variety and assortment.	1,255,382	800,000+	NA	5-25 miles
Regional Mall	General merchandise or fashion-oriented offerings. Typically, enclosed with inward-facing stores connected by a common walkway. Parking surrounds the outside perimeter.	589,659	400,000-800,000	40-80 stores	5-15 miles
Community Center (“Large Neighborhood Center”)	General merchandise or convenience-oriented offerings. Wider range of apparel and other soft goods offerings than neighborhood centers. The center is usually configured in a straight line as a strip, or may be laid out in an L or U shape, depending on the site and design.	197,509	125,000-400,000	15-40 stores	3-6 miles
Neighborhood Center	Convenience oriented.	71,827	30,000-125,000	5-20 stores	3 miles
Strip/ Convenience	Attached row of stores or service outlets managed as a coherent retail entity, with on-site parking usually located in front of the stores. Open canopies may connect the store fronts, but a strip center does not have enclosed walkways linking the stores. A strip center may be configured in a straight line, or have an “L” or “U” shape. A convenience center is among the smallest of the centers, whose tenants provide a narrow mix of goods and personal services to a very limited trade area.	13,218	<30,000	NA	<1 mile

Source: The International Council of Shopping Centers, January 2017

* GLA: Gross leasable area

Table 2: Department Store Categorization

Type	Definition	Stores
Discount	With a focus on price rather than on service, discount department stores sell a variety of merchandise at a lower price than typical retail stores. Many discount stores can be categorized as big-box stores, which offer a wide selection of products and grocery.	Kmart, Sears, Target, and Walmart
Midscale	Midscale department stores offer a wide selection of both brand-name and non-brand-name merchandise, seeking to offer good value to their customers.	Dillard's, JCPenney, Kohl's, and Macy's
Upscale	Upscale department stores sell goods at above-average prices; their customers are more interested in exclusive designer brands and value customer service over low price.	Bloomingdale's, Neiman Marcus, Nordstrom, and Saks Fifth Avenue

Source: J.D. Power and Associates (2007); Levy & Weitz (2011)

Table 3: Variable Summary Statistics

Variables	Mean	S.D.	
Sears	0.493	0.500	
Target	0.194	0.396	
Other Discount	0.359	0.480	
Dillard's	0.221	0.415	
Macy's	0.471	0.499	
Other Midscale	0.608	0.488	
Nordstrom	0.117	0.322	
Bloomingdale's	0.025	0.157	
Other Upscale	0.205	0.404	
Mall Sales (\$)	200,000,000	184,000,000	
Population	584,203	682,545	
Age	40.946	2.798	
Household Size	1.329	1.041	
Household Income (\$)	74,630	19,590	
Site Size (square feet)	915,498	364,859	
Open 2 (Malls opened during 1973 –1980)	0.232	0.422	
Open 3 (Malls opened after 1980)	0.476	0.500	
Distance to Headquarter (km)	Sears	9,978	5,585
	Target	10,423	5,112
	Dillard's	9,623	3,515
	Macy's	9,659	5,934
	Nordstrom	10,872	3,931
	Bloomingdale's	9,059	3,575

Table 4: Parameter Estimates—Profit Function (Market- and Firm-Specific Effects)

	Sears	Target	Other Disc	Dillard's	Macy's	Other Mid	Nordstrom	Bloomingtondale's	Other Up
Constant	-2.493 (.593)	-4.311 (.823)	-3.699 (.729)	-5.634 (1.479)	6.721 (1.82)	7.163 (1.934)	-1.191 (.286)	-2.834 (.441)	-0.152 (.299)
Population	-0.004 (.125)	0.407 (.18)	0.631 (.199)	-4.397 (.923)	-1.246 (.757)	-5.603 (1.067)	0.509 (.096)	0.982 (.23)	-0.340 (.081)
Age	0.150 (.085)	0.010 (.127)	-0.013 (.136)	-1.018 (.627)	1.669 (.644)	0.136 (.581)	-0.004 (.075)	0.341 (.162)	0.176 (.048)
Size HH	0.036 (.096)	-0.189 (.135)	-0.251 (.143)	2.094 (.701)	1.794 (.707)	3.359 (.816)	-0.132 (.098)	-0.449 (.231)	0.183 (.057)
Income	-0.058 (.099)	0.148 (.129)	-0.233 (.15)	-2.306 (.761)	1.396 (.685)	-0.880 (.657)	0.377 (.07)	0.364 (.113)	-0.015 (.054)
Site Size	-0.751 (.284)	-1.496 (.362)	-1.997 (.428)	11.825 (1.84)	13.669 (2.148)	11.443 (1.819)	0.712 (.204)	0.661 (.218)	0.655 (.195)
Open 2	-0.601 (.307)	-1.367 (.41)	-1.198 (.44)	8.427 (1.88)	5.224 (1.723)	7.774 (1.937)	0.240 (.21)	0.061 (.321)	0.194 (.172)
Open 3	-0.230 (.205)	1.004 (.354)	1.114 (.351)	0.856 (1.348)	-6.183 (1.574)	-1.689 (1.384)	-0.046 (.151)	-0.080 (.25)	-0.160 (.129)
Distance to HQ	0.009 (.059)	0.129 (.08)		0.581 (.259)	-0.124 (.332)		-0.077 (.061)	0.130 (.108)	
Mkt Level Shock	-2.440 (.649)	-3.637 (.807)	-4.047 (.861)	19.168 (2.831)	19.832 (3.026)	18.466 (2.839)	0.670 (.34)	0.489 (.368)	0.470 (.34)
Equilibrium Selection					-0.393 (.277)				

Note: Standard errors are reported in parentheses. Significance is in bold.

Table 5: Parameter Estimates—Profit Function (Spillover Effects)

	To discount	To midscale	To upscale
From discount	-1.915 (.526)	8.355 (1.397)	-0.027 (.202)
From midscale	2.854 (.656)	-15.531 (2.423)	-0.591 (.344)
From upscale	0.200 (.288)	-4.500 (2.04)	0.022 (.118)

Note: Standard errors are reported in parentheses. Significance is in bold.

Table 6: Parameter Estimates—Sales Model (controlling for endogeneity)

Constant	18.315 (.307)	Sears	-0.567 (.224)
Population	0.169 (.086)	Target	-0.667 (.265)
Age	0.036 (.059)	Other discount	-0.631 (.251)
Size HH	0.042 (.064)	Dillard's	0.857 (.385)
Income	0.220 (.065)	Macy's	0.825 (.364)
Site size	-0.113 (.199)	Other midscale	0.704 (.362)
Open 2	-0.294 (.178)	Nordstrom	0.518 (.215)
Open 3	0.226 (.141)	Bloomingdale's	0.381 (.337)
		Other upscale	0.117 (.157)

Note: Standard errors are reported in parentheses. Significance is in bold.

Table 7: Parameter Estimates—Sales Model (without controlling for endogeneity)

Constant	18.857 (.165)	Sears	-0.135 (.117)
Population	0.080 (.069)	Target	-0.159 (.142)
Age	0.039 (.05)	Other discount	-0.115 (.117)
Size HH	0.107 (.052)	Dillard's	0.050 (.123)
Income	0.210 (.06)	Macy's	0.055 (.121)
Site size	0.284 (.074)	Other midscale	-0.061 (.113)
Open 2	-0.014 (.117)	Nordstrom	0.360 (.171)
Open 3	0.146 (.121)	Bloomingdale's	0.270 (.306)
		Other upscale	-0.036 (.115)

Note. Standard errors are reported in parentheses. Significance is in bold.

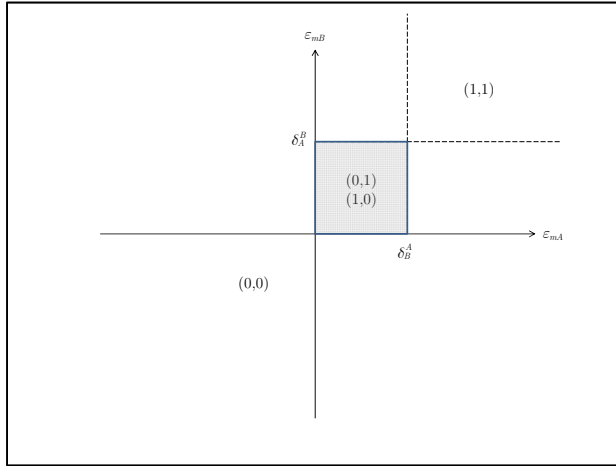
Table 8: Counterfactual Simulation—Choice of Market

		Site M	Site A	Site B
Population		380,914	419,005	380,914
Average Income		72,413	72,413	79,654
Entry probability	Sears	37.49%	36.29%	36.57%
	Target	26.12%	26.07%	27.21%
	Other discount	43.56%	43.99%	41.47%
	Dillard's	23.45%	22.68%	20.14%
	Macy's	42.12%	42.27%	44.33%
	Other midscale	54.32%	52.95%	53.84%
	Nordstrom	4.08%	4.60%	5.47%
	Bloomingdale's	0.08%	0.10%	0.12%
	Other upscale	17.16%	16.63%	17.24%
Sales (\$Billion)		1.89	1.93	2.07

Note: All variables are set to the mean values for site M. The population of site A is 10% higher than that of site M, and the income of site B is 10% higher than that of site M. Other variables are set to their mean values for sites A and B.

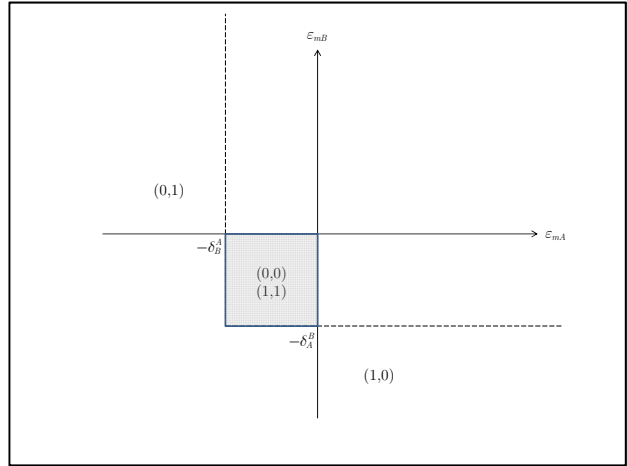
Figure 1: Multiple Equilibria—Illustrative Example

a) Competitive Entry (Negative Spillover)



The profit of firm A in market m is $\pi_{m,A} = -\delta_B^A a_{m,B} + \varepsilon_{m,A}$ and that of firm B is $\pi_{m,B} = -\delta_A^B a_{m,A} + \varepsilon_{m,B}$ ($\delta > 0$).

b) Complimentary Entry (Positive Spillover)



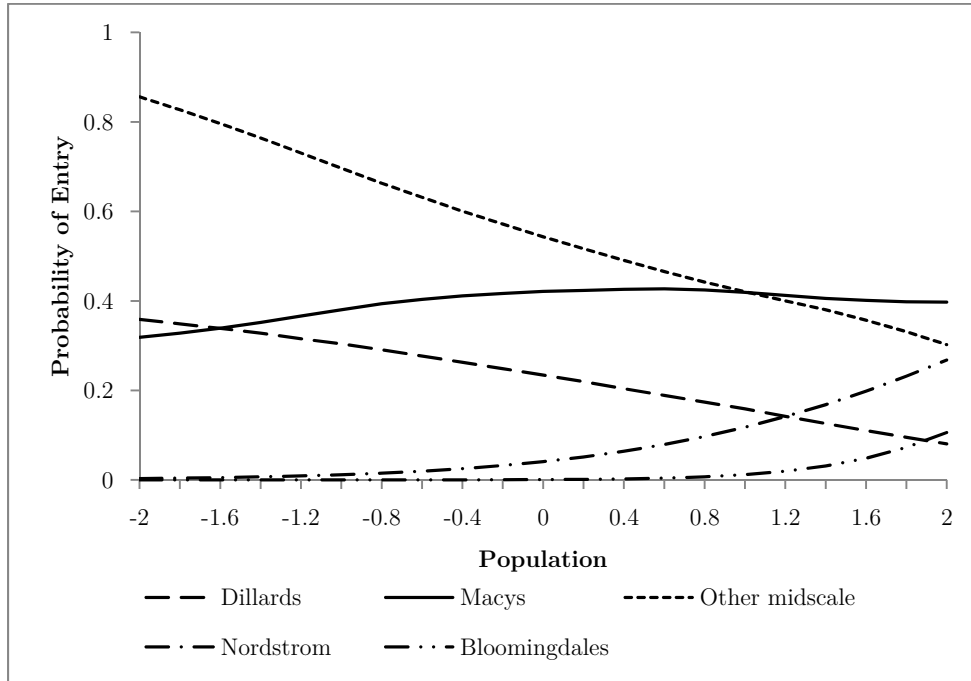
The profit of firm A in market m is $\pi_{m,A} = \delta_B^A a_{m,B} + \varepsilon_{m,A}$ and that of firm B is $\pi_{m,B} = \delta_A^B a_{m,A} + \varepsilon_{m,B}$ ($\delta > 0$).

Figure 2: Mall Floor Plan



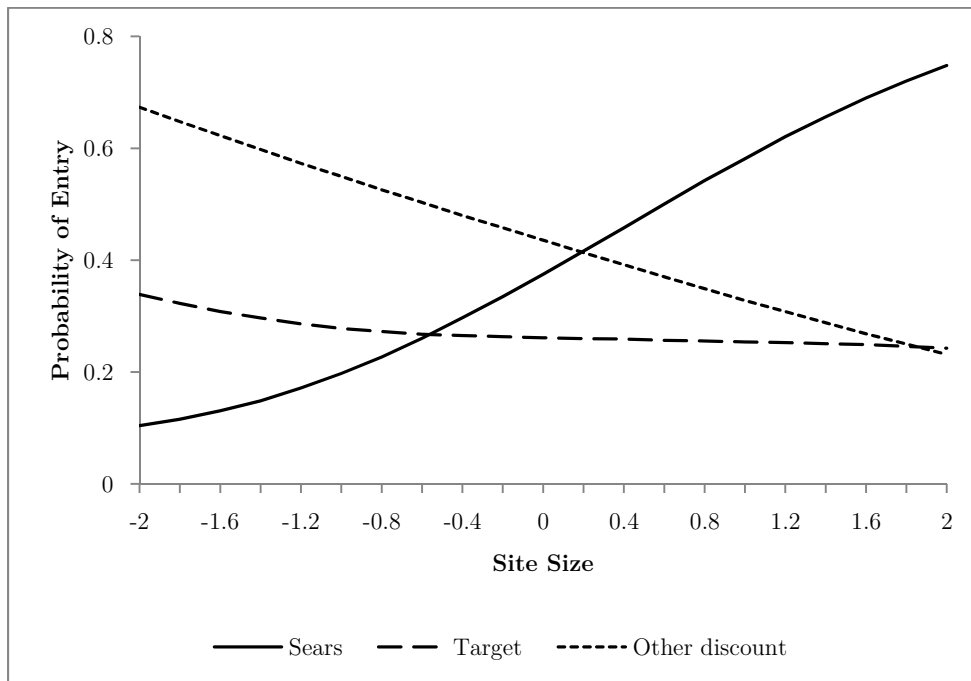
South Coast Plaza, Costa Mesa, CA 92626, <http://www.southcoastplaza.com/store-directory/>

Figure 3: Relation between Firm Entry and Population



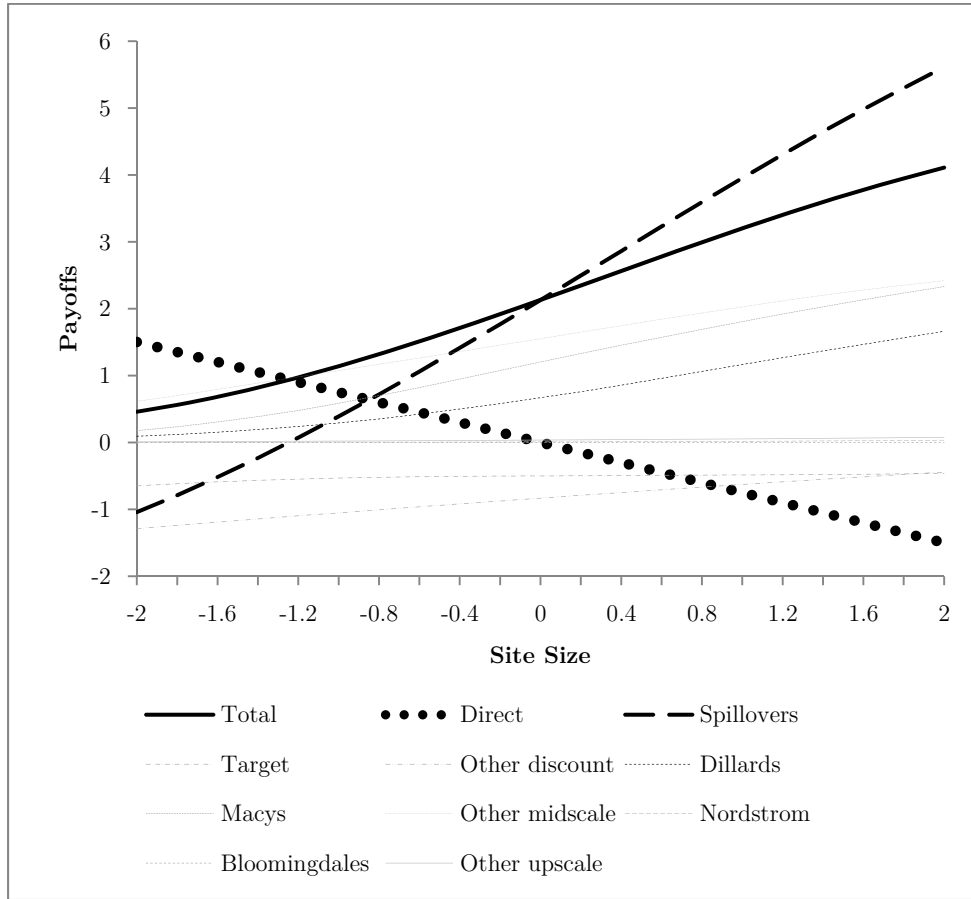
Note: Population is standardized value of log (population). That is, the x-axis represents standard deviations from the mean.

Figure 4: Relation between Firm Entry and Site Size



Note: Site size is standardized value of log (site size). That is, the x-axis represents standard deviations from the mean.

Figure 5: Main Effect vs. Spillover Effect of Site Size (Sears)



Note: The thick, downward-sloping dotted line represents the main effect, and the thick, upward-sloping dashed line represents the aggregate spillover effect. The thick, upward-sloping solid line is the sum of the main and the aggregate spillover effects. The eight thin lines are spillover effects from other firms. Site size is standardized value of log (site size). That is, the x-axis represents standard deviations from the mean.