

Strategic Clustering and Competition by Alcohol Retailers: An Empirical Analysis of Entry and Location Decisions

Yi Deng*

Gabriel Picone†

September 2012

Abstract

We develop and estimate a spatial game-theoretic model of entry and location choices to examine firms' strategic clustering decisions. The model identifies two contradictory effects that determine firms' geographical location choices: a competition effect and a clustering effect. We also separate firms' strategic clustering incentives from the observed clustering behavior due to exogenous factors such as population and topographic desirability or constraints. In particular, we examine two closely related industries that share similar location limitations but have different strategic incentives to cluster, jointly estimate the Bayesian Nash equilibrium of a two-industry entry and location game, and quantify the strategic clustering incentives.

JEL Classifications: L13, L81, R12, R30

*Department of Economics, University of South Florida, Tampa, FL,USA. Email: ydeng@usf.edu.

†Department of Economics, University of South Florida, Tampa, FL,USA. Email: gpicone@usf.edu.

1. Introduction

Since the pioneering work of Hotelling (1929), firm clustering has been examined by many researchers, mostly in a theoretical framework. While firms have an incentive to disperse (either in physical distance or in the space of other product attributes) because firms in close proximity engage in more intense price competition and have less market power (D'Aspremont et al. 1979, Eaton and Lipsey 1989), they also have an incentive to cluster to capture market share (Hotelling 1929). In general, theoretical literature on product differentiation has agreed that, in a location game with multiple product characteristics, firms will choose to maximize differentiation in some dominant product characteristics but minimize differentiation in others (Irmen and Thisse 1998). Therefore, on the dimension of main product attributes, the market power effect will dominate the market share effect and we should observe substantial spatial differentiation among firms.

The early empirical literature on clustering versus spatial differentiation often finds evidence contradicting the above theoretical prediction. For instance, Pinkse and Slade (1998) present evidence suggesting that gasoline stations have a tendency for clustering, and Stavins (1995) finds evidence of clustering (in product space) for computer models. A few studies also find evidence of clustering by manufacturers (Ellison and Glaeser 1997, Duranton and Overman 2008). However, these studies usually only examine a single product market in a single city, and do not take into account factors such as population density, zoning restrictions, road layouts, waterways or other topographic constraints, i.e., exogenous factors that determine the availability and desirability of location choices. Later studies in this area, including Netz and Taylor (2002) and Seim (2006), find that retailers that sell homogeneous products will spatially differentiate to avoid price competition.

Whether firms cluster or spatially differentiate has important policy implications. For example, restrictions on the maximum number of alcohol licenses for an area of the city will encourage potential alcohol retailers to spread across all areas if the competition effect dominates or will encourage the growth of another cluster in a different part of the city. Laws intended to encourage good businesses to locate in certain neighborhoods will also produce different predictions depending on whether firms cluster or spatially differentiate.

This paper studies entry decisions and location choices of on-site and off-site alcohol retailers. These two related industries have very different possibilities for product differentiation to avoid price competition. Off-site retailers such as liquor stores sell homogenous products and the main

product characteristic with which they can differentiate from their competitors is their spatial location. On the other hand, on-site retailers such as bars sell products with high product differentiation, since they can also differentiate on menu, drink mixture, and ambiance to avoid price competitions. As shown in Fisher and Harrington (1996), in such industries consumers' search costs are high, which tend to increase the firms' incentives to cluster. However, as more on-site retailers choose to locate in the same location, the opportunities for product differentiation and the marginal benefits of clustering may decrease.

We develop and estimate a spatial game-theoretic model of entry and location choices to examine firms' strategic clustering decisions, in an extended Seim (2006) framework. In particular, each potential entrant chooses a specific store location from among a set of discrete geographical locations, by comparing the expected profit among all possible locations based on the entrant's belief of the profitability at each location and other competitors' location choices. We estimate the Bayesian Nash equilibrium of this location game, and quantify the two contradictory effects of clustering and competition in affecting firms' geographical location choices. The current paper is the first in the literature in pursuing a game-theoretic approach to empirically separate the firms' competing and clustering behaviors.

Our methodology separates strategic clustering from the observed clustering due to factors such as population and topographic desirability or constraints. Note that in conducting empirical research, detailed information such as zoning restrictions, road layouts, or waterway distribution is usually not available, so researchers lack a direct way to control for such variables.¹ To solve this problem, we follow Picone, Ridley, and Zandbergen (2009)'s idea and examine two closely related industries that share similar location limitations but have different strategic incentives to cluster. We jointly estimate the location choices in both industries in a single spatial game-theoretic model, and are thus able to quantify the strategic clustering incentives within the industry. The advantage of our structural econometric model compared with Picone, Ridley, and Zandbergen (2009), is that instead of relying on a geographical index which only presents

¹As Picone, Ridley, and Zandbergen (2009) argue, almost any spatial pattern of firms or human establishments will reveal some degree of clustering, since this reflects the set of variables that create human settlements. They presented as an example the geographical distribution of public elementary schools. Both the map and the statistical indices indicate a significant clustering of schools, yet city planners should have little incentive to cluster the public schools. Rather, they should have tried to strategically disperse the public schools in populated areas. This underscores the importance of controlling for such variables when examining firms' clustering behavior.

a descriptive statistics of firms' clustering behavior as in their study, our model sheds light on the firms' decision-making process in a more structural, game-theoretical framework.

Estimation results reveal a strong negative competition effect for both on-site and off-site alcohol retailers. However, on-site alcohol retailers also benefit from a substantially positive clustering effect, which on average outweighs the negative competition effect. This explains why on-site alcohol retailers are more likely to cluster than off-site retailers. On the other hand, the benefits of clustering decrease rapidly as the number of existing on-site alcohol retailers increases in a given location, and estimation results suggest that on average, when there are two or three on-site alcohol retailers in a census tract, the marginal clustering effect is equal to marginal competition effect, consistent with the observed clustering pattern of bars in data. We also find that a larger population and a higher ratio of youths tend to increase the demand for alcohol, and that retailers also face demand and competition from other locations, which decay as geographic distance between locations increases.

The rest of the paper is organized as follows. Section 2 develops a spatial game-theoretic model of entry and location choices, and Section 3 describes the data and our estimation methodology. In Section 4 we present estimation results and discuss their implications. Section 5 concludes.

2. A Spatial Game-Theoretic Model of Entry and Location Choices

2.1. The Equilibrium of Single-Industry Game

We examine retailers' entry and location choices in a spatial game-theoretic model with incomplete information. We begin with a location game in a single industry. In particular, a group of F^m potential entrants simultaneously choose whether or not to enter a market m , as well as where to locate, among possible store locations $j = 0, 1, 2, \dots, J^m$, where $j = 0$ denotes the decision of no entry. A location is defined as a geographical area with a clear boundary within a market, such as a few blocks, or a census tract. Each potential entrant makes his entry and location choice by comparing the expected profit among all possible locations, based on the entrant's belief of the profitability at each location and other competitors' location strategies. In order to compute an equilibrium with many possible locations, we choose to examine a static, single-period game. Computing and estimating a dynamic game with many locations turns out to be computationally infeasible.

A representative retailer i 's profit in location j depends on a set of demand and cost characteristics in this location as well as the geographical distribution of other stores. In particular, the store profit is determined by

$$\pi_{ij}^m = X_{j,-j}^m \beta + \xi^m + f(N^m, j) + \varepsilon_{ij}^m \quad (2.1)$$

where $X_{j,-j}^m$ is a vector of demand and cost characteristics specific to location j and other locations in this market, such as population density, household income, etc. ξ^m is a market-level fixed effect, normally distributed, capturing the exogenous demand and cost characteristics that are not included in X_j^m , such as consumer preferences in this market. For instance, residents in one city may be more prone to go to bars instead of liquor stores than in other cities, or some cities are more tourist-oriented, etc. ε_{ij}^m represents the idiosyncratic component of retailer i 's profit if operating in location j , which is private information known only to retailer i himself. Other players on the market only know the distribution of ε_{ij}^m .

The term $f(N^m, j)$ describes the influences of other stores on store i 's profits in location j , due to both competition and clustering effects, where N^m is a vector containing the number of stores in each of the J^m locations in market m . To separately estimate the competition and clustering effects on store profits, structural assumptions on the functional forms of $f(N^m, j)$ must be made. In particular, here we specify $f(N^m, j)$ as the logarithm of a double-exponential function (Jaffe and Trajtenberg 1996) which consists of two terms, one measuring the competition effect and the other capturing the clustering effect:

$$\begin{aligned} f(N^m, j) &= \log\{\left[\exp(N_j + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'})\right]^{-\gamma_1} (1 - e^{-\theta N_j})^{\gamma_2}\} \\ &= -\gamma_1 [N_j + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'}] + \gamma_2 \log(1 - e^{-\theta N_j}) \end{aligned} \quad (2.2)$$

N_j and $N_{j'}$ are the number of competing stores in the same location j and some other location j' , respectively. $\kappa_{j,j'}^R$ are weights used to weigh the number of competing stores in different locations j' (to be explained in details below), and γ_1, γ_2 , and θ are model parameters to be estimated.

The above double-exponential functional form has been widely used in the literature to capture two contradictory effects. For instance, when estimating patent citation dynamics over time, statisticians often have to separate a diffusion process, which tends to increase patent citations as the time lag increases, and an obsolescence process, which tends to decrease citations

as the time lag increases, and the double-exponential functional form has been found to be very successful in separating these two contradictory effects due to its great flexibility (Caballero and Jaffe 1993, Jaffe and Trajtenberg 1996, Hall, Jaffe, and Trajtenberg 2005, Deng 2008) . As our needs here are very similar, this functional form becomes our natural choice.

We take the logarithm of the double-exponential function so that the first term in equation (2.2) is linear in the number of competing stores, the same specification that Seim (2006) and many other researchers have used to quantify the competition effect alone. We also interpret this linear term as the competition effect. The second term in the equation captures the clustering effect and increases store profit as the number of competing stores *in the same location* increases. For retailers this clustering effect arises from consumers' preferences for variety and firms' opportunities to differentiate their products. However, as more and more stores enter the same location, this clustering effect may increase at a diminishing pace (when $\theta > 0$), as it may become more and more difficult for a store to differentiate its product, so the marginal benefit may decline.

Stores in other locations imply less intense competition than stores in the same location. Thus in specifying the competition effect on store profit, we discount the numbers of competing stores in other locations by a set of weights $\kappa_{j,j'}^R$, which decreases as geographical distance between the stores increases. Previous studies such as Seim (2006) often first draw several distance bands around the location (for instance, with radii of 1 mile, 1 to 3 miles, and more than 3 miles) and then assign all the stores that fall in the same band equal weights, no matter which locations they belong to. Here we use a more accurate method to calculate weights, as a continuous function of the distance between different locations. In particular,

$$\kappa_{j,j'}^R = \exp(-d_r D_{j,j'}) \tag{2.3}$$

i.e., the weight used to discount the number of stores in location j' when specifying the competition effect for store in location j is an exponential function of the distance between these two locations $D_{j,j'}$. $d_r > 0$ is a decay parameter to be estimated, so the weight $\kappa_{j,j'}^R$ lies between zero and one. As the distance increases, competition from other locations declines.

Next we model other profit determinants that are specific to location j . Instead of assuming that store profits in location j depend only on market characteristics in this location X_j^m , we choose to include characteristics from all other locations in the same market as well, weighed by $\kappa_{j,j'}^P$, which lies between zero and one depending on the distance between location j and

any other location j' . We make this comprehensive assumption because in our study different locations in the same market are *not* completely segmented. Rather, we allow for the possibility that residents in one location can go to retailers in locations other than their home location, as long as they are willing to tolerate the associated inconvenience and pay transportation cost. Therefore, the first term of store i 's profit in location j , as shown in the right hand side of equation (2.1), is expanded into

$$X_{j,-j}^m \beta = X_j^m \beta + \sum_{j' \neq j} \kappa_{j,j'}^P X_{j'}^m \beta = (X_j^m + \sum_{j' \neq j} \kappa_{j,j'}^P X_{j'}^m) \beta \quad (2.4)$$

where X_j and $X_{j'}$ are a list of demand and cost characteristics specific to locations j and j' , including population, household income, age structure, ethnic background, etc, respectively. The weight $\kappa_{j,j'}^P$ is defined in a similar way as in equation (2.3)

$$\kappa_{j,j'}^P = \exp(-d_p D_{j,j'}) \quad (2.5)$$

Thus, the profit function is given by

$$\pi_{ij}^m = \xi^m + (X_j^m + \sum_{j' \neq j} \kappa_{j,j'}^P X_{j'}^m) \beta - \gamma_1 [N_j + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'}] + \gamma_2 \log(1 - e^{-\theta N_j}) + \varepsilon_{ij}^m \quad (2.6)$$

Next we compute the Bayesian Nash equilibrium of this single-industry location game. Due to imperfect information about his opponents' profitability, retailer i can form only an expectation of their optimal location strategies, based on which he will then choose his store location. Thus, his expected profit in location j is

$$\begin{aligned} E\pi_{ij}^m &= \xi^m + (X_j^m + \sum_{j' \neq j} \kappa_{j,j'}^P X_{j'}^m) \beta - \gamma_1 [EN_j + \sum_{j' \neq j} \kappa_{j,j'}^R EN_{j'}] + \gamma_2 \vartheta_j \log(1 - e^{-\theta EN_j}) + \varepsilon_{ij}^m \\ &= E\bar{\pi}_{ij}^m + \varepsilon_{ij}^m \end{aligned} \quad (2.7)$$

where EN_j and $EN_{j'}$ represent his expectation of the store numbers in location j and j' , respectively. $E\bar{\pi}_{ij}^m$ contains the part of retailer i 's profit at j that is observable to all other players, and ε_{ij}^m as noted above is the private information known only to retailer i himself. ϑ_j is a convexity-adjustment term which depends on retailer i 's belief of the probabilistic distribution of store locations. ²

²Note that when we take the expectation of the profit function (2.6), $E \log(1 - e^{-\theta N_j})$ does not equal to

Due to the symmetry of his rivals' types, retailer i 's perception of the location choice of any other retailer r would be the same for all competitors. In particular, to retailer i , the probability that any retailer r chooses location j , p_{rj} , is given by

$$p_{rj} = pr(E\bar{\pi}_{rj}^m + \varepsilon_{rj}^m \geq E\bar{\pi}_{rj'}^m + \varepsilon_{rj'}^m), \quad \forall j' \neq j \quad (2.8)$$

The number of competitors that i expects to face in location j is thus $(\Xi^m - 1)p_{rj}$, or $(\Xi^m - 1)p_j$ because of the symmetry of player types. Ξ^m is the total number of entrants to market m and is to be solved below.

We assume that the idiosyncratic component of store profit, ε 's, are *i.i.d.* draws from a type I extreme-value distribution. This leads to multinomial logit probabilities of stores' location distribution with a closed-form expression, thus simplifying the model solution and estimation enormously. In particular, to retailer i , the probability of store r entering location j is given by

$$p_{rj} = \frac{\exp(E\bar{\pi}_{rj}^m)}{\sum_{k=1}^{J^m} \exp(E\bar{\pi}_{rk}^m)} \quad (2.9)$$

The equilibrium is a symmetric Bayesian Nash equilibrium under which each player's conjecture about other competitors' choice strategies coincide with the optimal responses that other players make, given their conjectures of all players' strategies. The symmetry of player types implies that every player has the same equilibrium conjecture about his rivals' location choice, namely $p_i = p_r = p^*$, where p^* is a vector of equilibrium conjectures of other players' strategies over all locations:

$$p_j^* = \frac{\exp(E\bar{\pi}_j^m)}{\sum_{k=1}^{J^m} \exp(E\bar{\pi}_k^m)} \quad (2.10)$$

where

$\log(1 - e^{-\theta EN_j})$ because of convexity of the function. Therefore, we take the following approximations: $\log(1 - e^{-\theta N_j}) \approx -e^{-\theta N_j}$ since $e^{-\theta N_j}$ is very small. Thus we have $E \log(1 - e^{-\theta N_j}) \approx -E(e^{-\theta N_j}) = -e^{-\theta EN_j + \frac{\theta^2}{2} V(N_j)}$ where $V(N_j)$ is the variance of retailer i 's belief of the probabilistic distribution of the store number at location j . Therefore, we adjust the convexity by defining $\vartheta_j = e^{\frac{\theta^2}{2} V(N_j)} = \exp[\frac{\theta^2}{2} N p_j (1 - p_j)]$ and re-write $E \log(1 - e^{-\theta N_j}) \approx \vartheta_j \log(1 - e^{-\theta EN_j})$ as in (2.7). We thank an anonymous referee for pointing out the need to make such a convexity adjustment to us.

$$\begin{aligned}
E\bar{\pi}_j^m &= \exp(\xi^m + X_{j,-j}^m \beta - \gamma_1[EN_j + \sum_{j' \neq j} \kappa_{j,j'}^R EN_{j'}] + \gamma_2 \vartheta_j \log(1 - e^{-\theta EN_j})) \\
&= \exp(\xi^m + X_{j,-j}^m \beta - \gamma_1[(\Xi^m - 1)p_j^* + \sum_{j' \neq j} \kappa_{j,j'}^R (\Xi^m - 1)p_{j'}^*] + \gamma_2 \vartheta_j \log(1 - e^{-\theta(\Xi^m - 1)p_j^*}))
\end{aligned}$$

and is obtained by substituting equation (2.7) into equation (2.9) and realizing that the expected number of competitors in any location j is simply $(\Xi^m - 1)p_j^*$.

The system of equations (2.10) defines the equilibrium location conjecture p^* as a fixed point of mapping from a retailer's conjecture of his rivals' strategies into his rivals' conjecture of the retailer's own strategy. It contains J^m unknowns and J^m equations, one for each of the J^m locations in market m , and can be numerically solved by a non-linear equation solver. Seim (2006) proves the existence and uniqueness properties of such an equilibrium.

Finally, we conclude this entry and location choice game by solving for the total number of entrants Ξ^m . In equilibrium, a potential entrant will enter the market if and only if he expects a non-negative profit. The probability of entry depends on a comparison of expected profits across all locations with the zero payoff of not entering the market. Given the assumption of *i.i.d.* type-I extreme-value distribution of ε 's, the probability of entry is thus given by

$$\begin{aligned}
\Pr(\text{entry}) &= \frac{\sum_{k=1}^{J^m} \exp(E\bar{\pi}_j^m)}{1 + \sum_{k=1}^{J^m} \exp(E\bar{\pi}_k^m)} \\
&= \frac{\{\exp(\xi^m) \sum_{k=1}^{J^m} \exp(X_{j,-j}^m \beta - \gamma_1[(\Xi^m - 1)p_j^* + \sum_{j' \neq j} \kappa_{j,j'}^R (\Xi^m - 1)p_{j'}^*] \\
&\quad + \gamma_2 \vartheta_j \log(1 - e^{-\theta(\Xi^m - 1)p_j^*}))\}}{\{1 + \exp(\xi^m) \sum_{k=1}^{J^m} \exp(X_{k,-k}^m \beta \\
&\quad - \gamma_1[(\Xi^m - 1)p_k^* + \sum_{k' \neq k} \kappa_{k,k'}^R (\Xi^m - 1)p_{k'}^*] + \gamma_2 \vartheta_k \log(1 - e^{-\theta(\Xi^m - 1)p_k^*}))\}}
\end{aligned} \tag{2.11}$$

Note that the market-specific fixed effect ξ^m is cancelled in the numerator and denominator of equation (2.10), as it does not affect the relative attractiveness of different locations once the retailer has decided to enter this market. However, it appears in entry probability equation (2.11) as it affects the overall attractiveness of the market. Therefore, the expected number of entrants is simply

$$\Xi^m = F^m * \Pr(\text{entry}) \tag{2.12}$$

This concludes our description of the equilibrium of the entry/location choice game when all players are from the same industry.

2.2. The Equilibrium of A Two-Industry Game

For our purpose we need to expand the above single-industry model and include two types of retailers, on-site retailers and off-site retailers. As shown in Irmen and Thisse (1998), with multiple product characteristics to decide on, firms tend to maximize differentiation in one dominant characteristic but minimize differentiation in others. Our hypothesis here is that these two kinds of retailers have different clustering incentives: on-site retailers are more likely to differ on menus and ambience or atmosphere and are less likely to differ on geographical distance, whereas off-site retailers are more likely to spatially differentiate to reduce price competition since they sell a more homogeneous product. As noted by Picone, Ridley, and Zandbergen (2009), there may be substantial rewards for on-site retailers to cluster, such as greater demand from decreasing consumer travel costs to a location containing more than one retailer, or more varieties for consumers to choose from, but much less so for off-site retailers.

We modify the above entry and location choice game by allowing the location distribution of one type of retailers to affect the other type's profit and location strategy. We begin by re-specifying the profit function of an on-site retailer i in a location j as (the superscript m is ignored)

$$\begin{aligned} \pi_{ij}^{on} = & \xi^{on} + X_{j,-j}\beta^{on} - [\gamma_1^{on-on}(N_j^{on} + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'}^{on}) + \gamma_1^{on-off}(N_j^{off} + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'}^{off})] \\ & + \gamma_2 \vartheta_j \log(1 - e^{-\theta N_j^{on}}) + \varepsilon_{ij}^{on} \end{aligned} \quad (2.13)$$

and the profit function of an off-site retailer i in location j as

$$\pi_{ij}^{off} = \xi^{off} + X_{j,-j}\beta^{off} - [\gamma_1^{off-on}(N_j^{on} + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'}^{on}) + \gamma_1^{off-off}(N_j^{off} + \sum_{j' \neq j} \kappa_{j,j'}^R N_{j'}^{off})] + \varepsilon_{ij}^{off} \quad (2.14)$$

with N_j^{on} and N_j^{off} representing the number of on-site and off-site retailers in location j , respectively, and the parameters to be estimates are

$$\{\beta^{on}, \beta^{off}, d_p, d_r, \gamma_1^{on-on}, \gamma_1^{on-off}, \gamma_1^{off-on}, \gamma_1^{off-off}, \gamma_2, \theta, \mu^{on}, \mu^{off}, \sigma^{on}, \sigma^{off}\}$$

(μ^{on}, σ^{on}) and $(\mu^{off}, \sigma^{off})$ are the mean and variance of the normally distributed on-site and off-site fixed effects ξ^{on} and ξ^{off} , respectively.

Equation (2.13) implies that the profit of an on-site retailer includes a competition effect and a clustering effect, similar to equation (2.6), except that now an on-site retailer faces competition from not only the other on-site retailers but also off-site retailers. We introduce a new parameter γ_1^{on-off} here, because we would like to allow for the possibility that an additional on-site retailer may have a quantitatively different competition effect on the store profit than an additional off-site retailer.

The profit of an off-site retailer, on the other hand, turns out to be *not* affected by the clustering effect and only includes a competition effect from both on-site and off-site retailers.³ Thus equation (2.14) is very similar to Seim (2006)'s specification of store profit in her study of video-rental stores, as in both cases stores sell homogeneous goods (video rental and bottled alcohol) and thus would engage in spatial differentiation.

We preserve several assumptions from the single-industry game, including that 1) retailers possess incomplete information regarding other retailers' idiosyncratic shocks ε 's; 2) within each industry (on-site or off-site), player types are symmetric; 3) ε 's follow an *i.i.d.* type-I extreme-value distribution which is common knowledge to all players. Thus the Bayesian Nash equilibrium is defined as a pair of strategies (p^*, q^*) , where p^* represents any on-site retailer's equilibrium conjectures about other on-site retailers' location choices, and q^* denotes any off-site retailers' equilibrium conjectures about other off-site retailers' location choices. In particular, (p^*, q^*) is the solution of the following system of $2J^m$ equations

³When estimating the model, rather than arbitrarily assuming bars or liquor stores to be on-site or off-site, we take an agnostic approach and let the data tell. For details, please refer to Section 4 and Table 3.

$$\begin{aligned}
p_j^* &= \{ \exp(X_{j,-j} \beta^{on} - [\gamma_1^{on-on} ((\Xi^{on} - 1)p_j^* + \sum_{j' \neq j} \kappa_{j,j'}^R (\Xi^{on} - 1)p_{j'}^*) \\
&\quad + \gamma_1^{on-off} (\Xi^{off} q_j^* + \sum_{j' \neq j} \kappa_{j,j'}^R \Xi^{off} q_{j'}^*)] + \gamma_2 \vartheta_j \log(1 - e^{-\theta(\Xi^m - 1)p_j^*}) \} / \\
&\quad \{ \sum_{k=1}^{J^m} \exp(X_{k,-k} \beta^{on} - [\gamma_1^{on-on} ((\Xi^{on} - 1)p_k^* + \sum_{k' \neq k} \kappa_{k,k'}^R (\Xi^{on} - 1)p_{k'}^*) \\
&\quad + \gamma_1^{on-off} (\Xi^{off} q_k^* + \sum_{k' \neq k} \kappa_{k,k'}^R \Xi^{off} q_{k'}^*)] + \gamma_2 \vartheta_j \log(1 - e^{-\theta(\Xi^m - 1)p_k^*}) \} \quad (2.15)
\end{aligned}$$

$$\begin{aligned}
q_j^* &= \{ \exp(X_{j,-j} \beta^{off} - [\gamma_1^{off-on} (\Xi^{on} p_j^* + \sum_{j' \neq j} \kappa_{j,j'}^R \Xi^{on} p_{j'}^*) \\
&\quad + \gamma_1^{off-off} ((\Xi^{off} - 1)q_j^* + \sum_{j' \neq j} \kappa_{j,j'}^R (\Xi^{off} - 1)q_{j'}^*)] \} / \\
&\quad \{ \sum_{k=1}^{J^m} \exp(X_{k,-k} \beta^{on} - [\gamma_1^{off-on} (\Xi^{on} p_k^* + \sum_{j' \neq j} \kappa_{k,k'}^R \Xi^{on} p_{k'}^*) \\
&\quad + \gamma_1^{off-off} ((\Xi^{off} - 1)q_k^* + \sum_{k' \neq k} \kappa_{k,k'}^R (\Xi^{off} - 1)q_{k'}^*)] \} \quad (2.16)
\end{aligned}$$

And finally, the number of entrants (Ξ^{on}, Ξ^{off}) are solved in a similar fashion as in equations (2.11) and (2.12).

3. Data Description and Estimation Approach

3.1. Sample Characteristics

We extract our data set from the National Establishment Time-Series (NETS) database, and focus on liquor retailers in five metropolitan areas in the U.S.: Birmingham, AL; Oakland, CA; Tampa, FL; Chicago, IL; and Minneapolis-St. Paul, MN. These five areas create a nationally representative sample of large U.S. cities and at the same time are also very different in terms of population clustering, demographic composition, religious beliefs, and other characteristics that might influence zoning, shopping, and preferences for alcohol.

The NETS data base contains street addresses of all the alcohol retailers, their Standard Industrial Classification (SIC) code, the start and the end of their operating years, and sales,

as well as other characteristics of the stores. For our purpose, we need to locate the stores on the map, preferably to a very detailed level, and examine their spatial pattern. In the following analysis, we define a “market” as a census-designated place, which in most cases simply corresponds to an incorporated town, and a “location” as a census tract, as the census tract is the most disaggregated level at which detailed data on population density and household characteristics can be found at the U.S. Census Bureau. Therefore, we geocode the street addresses of each and every liquor store using ArcGIS 10, and map them into the census tracts to which they belong. The ArcGIS 10 successfully mapped 84% of the liquor stores into census tracts to which they belong. For the remaining 16%, we calculate the spherical distance between each store and the neighboring census tracts based on their longitude and latitude coordinates, and map them into the closest census tract. This procedure turns out to be very successful: among those 16% of stores which the ArcGIS 10 cannot map, the median of the distance between stores and the center of the closest census tracts is less than 0.5 miles for Oakland, Chicago, and Minneapolis-St. Paul, 0.6 miles for Tampa, and 0.9 miles for Birmingham, much smaller than the normal radius of a census tract. Thus we conclude that in most cases, the closest tract should be the actual census tract to which they belong.

Table 1 shows the number of designated places in each of the five metropolitan areas, along with other characteristics of the places and census tracts in which we are interested. Populations of places vary greatly in these areas, from a little over 1000 to almost 1 million. However the populations of census tracts are much more similar, in most cases ranging from 2000 to 5000, with mean and median both between 4,000 and 5,000 in all five metropolitan areas. The ratio of non-hispanic whites varies from 47 percent in Oakland, CA to 85 percent in Minneapolis-St. Paul, MN. Median household income also varies substantially, from the lowest of \$39,189 in Tampa, FL, to the highest of \$61,507 in Oakland, CA. The ratio of youths (defined as those aged between 25 and 34), on the other hand, is not very different across the metropolitan areas, ranging from 13 percent in Birmingham, AL to 15 percent in Oakland, CA.

We focus on estimating the location game of two very closely related industries, the on-site and the off-site alcohol retailing industries. In particular, we define the on-site alcohol retailers as all “drinking places” with a 4-digit SIC code 5813 (“bars” thereafter), and off-site alcohol retailers as all “liquor stores” with a 4-digit SIC code 5921, based on the main type of businesses the stores reported in year 2005. Both industries serve alcoholic drinks and are subject to the

same zoning restrictions. However, they show very different patterns of spatial distribution. As Table 2 reveals, of the 1,149 census tracts in our sample that have at least one bar within the tract, more than 43% of the tracts have two or more bars, compared at merely less than 16% for liquor stores (i.e., more than 84% of the liquor stores are the only liquor store in the census tracts they are located). This seems to suggest that bars are more likely to cluster than liquor stores, and liquor stores tend to spatially separate from each other. The model estimation below tries to quantify their incentives.

3.2. Estimation Method

The system of equations (2.11) and (2.12) is highly nonlinear and thus numerically very difficult to solve. To simplify model estimation, we assume that the expected number of entrants predicted by the model in equation (2.12) is exactly the same as the number of entrants observed in the data. This approach of using an unobservable effect to equal the actual and predicted numbers of entrants has also been adopted by Berry (1994), Berry, Levinsohn, and Pakes (1995), and Seim (2006). In particular, from equations (2.11) and (2.12) we can solve the market-specific fixed effect ξ^m as a function of location characteristics and equilibrium location conjecture

$$\xi^m = \ln(\Xi^m) - \ln(F^m - \Xi^m) - \ln\left(\sum_{j=1}^{J^m} \exp(E\pi_j^m)\right) \quad (3.1)$$

The model is then estimated using a Maximum Likelihood Estimator (MLE). Each market is treated as an independent entry and location game, and the likelihood function is given by

$$L(\omega) = \prod_{m=1}^M pdf(d^m | \xi^m, M^m, \Xi^m) * pdf(\xi^m | M^m, \Xi^m, F^m) \quad (3.2)$$

where ω represents a vector of model parameters. $pdf(d^m | \xi^m, M^m, \Xi^m)$ is the likelihood of entrants' location choices in market m conditional on the market characteristics M^m , the fixed effect ξ^m , and number of entrants Ξ^m , and $pdf(\xi^m | M^m, \Xi^m, F^m)$ denotes the probability of observing the particular realization of ξ^m . We assume ξ^m is normally distributed, with a mean and standard deviation to be estimated. We assume that the number of potential entrants in each market is 150% of the actual number of stores in the market. A similar assumption is made in Seim (2006) and other studies.

Therefore, for a given set of parameter value ω and observed market characteristics, location choices and actual and potential entry in each market, the system of equations (2.10) is

numerically solved by a Matlab non-linear equation solver, market by market, for an equilibrium location-choice probability p^* , or (p^*, q^*) in a two-industry game. The solution is then substituted into equation (3.1) to yield an equilibrium realization of ξ^m . Finally, the fixed-point equilibrium is nested into a MLE routine to estimate the model parameter θ . We use a combination of grid search and hill-climbing algorithms to estimate the model parameters.

4. Estimation and Simulation Results

We start model estimation by first separately estimating a single-industry location game for bars and liquor stores, and then a two-industry joint game among both bars and liquor stores. In the first stage single-industry game estimation, rather than arbitrarily assuming bars or liquor stores to be on-site or off-site, we take an agnostic approach and let the data tell. In particular, since the off-site game is nested in the on-site game (by setting γ_2 in the profit function to zero), we start by estimating an on-site location game for bars and liquor stores, separately, and then decide whether the clustering effect as determined by γ_2 and θ is significant or not.

Table 3 reports the parameter estimates. The first and second columns display estimation results of a single-industry on-site game for bars and liquor stores, respectively. Estimates of parameter γ_2 in these two columns clearly suggest that bar industry can be characterized as an on-site industry, as the estimate of γ_2 in the first column of the table is statistically significant. The liquor stores, on the other hand, can be characterized as an off-site industry, as the estimates of both γ_2 and θ are statistically insignificant. Thus in the third column of Table 3 we present estimation results of an off-site location game for liquor stores, by restricting γ_2 and θ to zero. The associated $\chi^2(2)$ statistic is 0.89, suggesting that such restrictions do not hurt model fit. Therefore, in the following estimation and simulations, we keep the on-site specification for bar industry and off-site specification for liquor stores.

Next we turn to discussion of model estimates. Most of the parameter estimates are of the anticipated sign. For instance, population has a positive and statistically significant effect on the profits of bars and liquor stores, when the model is separately estimated for each industry as well as when they are jointly estimated. A higher ratio of youths implies higher profits for both bars and liquor stores, although the latter is not statistically significant. The ratio of non-hispanic whites has a positive but insignificant effect on the liquor store profits, and its effect on bar profits is negligible. Median household income has almost no effect on the profits of bars

and liquor stores.

To quantify the effects of exogenous changes in the demographic variables on location-choice probabilities, we simulate the model and numerically differentiate the location probabilities with respect to each variable. In particular, for each game as well as for each demographic variable, we separately compute the response to an exogenous change of the variable, in each market and location, and report the average response of the location probabilities across all locations and markets.

Table 4 reports the simulated effect on location probabilities. For instance, in the single-industry location game among on-site retailers (column 1), a 10 percent increase in the population of a location (census tract) implies an approximate 0.6 percentage point increase in the likelihood that a potential bar will open in the same location, and a 10 percentage point increase in the ratio of youths in a location will increase the likelihood of a potential bar being located in the same location by more than 0.4 percentage point. In the single-industry location game for off-site retailers, as shown in column 2, such effects on location probabilities are 0.7 percent and 0.1 percent, respectively. A 10 percentage point increase in the ratio of non-hispanic whites in the population of a location will increase the location likelihood of a liquor store by 0.3 percent, yet its effect on the location choice of bars is negligible. The marginal effects of changes in median household income are negligible for both bars and liquor stores.

Parameter estimates of γ_1 indicates a negative competition effect on store profit and thus a lower probability when there are more existing retailers at the same location. To quantify its effect in terms of location probabilities, a simulation exercise similar to the one above is also conducted, by adding one more retailer in each location in the sample, one at a time, and calculating the difference in the model-implied location probabilities. Table 5 displays results of this simulation. For instance, having one additional liquor store in a location will decrease the likelihood for a new liquor store to open in the same location by approximately 0.6 percentage point. Similarly, an increase in the number of existing bars in a location implies a lower location probability for a new bar to open there, by approximately 0.8 percentage point (assuming the clustering effect is zero).

On the other hand, positive estimates of γ_2 and θ for the on-site retailers reveal a strong clustering effect within bars, and thus having more bars may imply stronger incentives for new bars to open in the same location. Again, a simulation is conducted to quantify the average

magnitude of this effect. As reported in column 1 of Table 5, the clustering effect of having one more bar in a location is to increase the location choice probability of a new bar in the same location by 4.9 percentage point. Thus the net effect of having one more bar, when aggregating the competition and clustering effects, is to increase the location probability for a new bar to open there by 4.1 percent, as reported in the last row of Table 5, rather than a decrease in the location likelihood.

Retailers face demand as well as competition not only from their own location, but also from other locations in the same market. Because of the associated inconvenience and traveling cost, demand and competition from another location decays as distance between the two locations increases. For instance, parameter estimates of d_p for bars (column 1 of Table 3) implies that a population increase of 1,000 in a location that is 1.5 miles away will be equivalent to an increase of 730 in the local population (note that the pair-wise distance between census tracts from the same place is usually between 0.5 and 2 miles). On the other hand, competition from another location decays much more quickly: the parameter estimate of d_r for bars implies that the competition from one additional bar 1.5 miles away is only 11 percent of the competition from a local bar.

The decay speeds of product demand and competition for liquor stores are different from those of bars. In particular, estimates of d_p and d_r (column 2 of Table 3) imply that product demand for liquor store decays at a faster speed than the competition as geographical distance increases. In particular, a population increase of 1,000 in a neighboring location 1.5 miles away will generate an additional demand equivalent to that implied by 126 local persons, i.e., the demand shrinks by 87 percent at this distance, yet the competition effect brought by an additional liquor store 1.5 miles away would be equivalent to what is brought by 0.50 local liquor store, or shrinks by 50 percent.

Now we turn to the estimation of the two-industry joint location game. The fourth column of Table 3 reports the parameter estimates. Again, the estimates of β 's indicate that a larger population and a higher youths ratio have a positive and significant effect on profits of bars and thus their location probabilities, but the ratio of non-hispanic whites and median household income have negligible effect. For liquor stores, only population increases have positive and significant effects on the store profits and location probabilities, and the effects of ratios of youths, non-hispanic whites, and household income are all insignificant.

Estimates of γ'_1 s continue to show a significant competition effect on a retailer's profit from others retailers in the same industry. Moreover, small but significant estimates of γ_1^{on-off} and γ_1^{off-on} (both less than 0.1) indicate a mildly negative cross-industry competition effect. Overall, an increase of one more existing liquor store in a location tends to decrease the probability for a new liquor stores to open at the same location, by 0.6 percentage point. Such a competition effect is slightly larger for bars, about 0.7 percentage point. Moreover, more bars in the same location also induce a stronger clustering effect and increase the attractiveness of the location, as the estimates of γ_2 and θ are again positive and significant. The associated clustering effect of having one more existing bar is to increase the location likelihood of a new bar to open there by 6.6 percentage point on average. Thus the net effect of having one more existing bar in a location is to increase the choice probability by 5.9 percentage point, rather than to decrease this probability (Table 5). Estimates of d_p and d_r for bars and liquor stores suggest that the serving radius of bars is much larger than that of liquor stores, as the decaying speed of demand for bars is much slower than that for liquor stores, yet the competition from other locations for liquor stores is stronger than for bars, as the competition effect for liquor stores shrinks at a slower speed than for bars.

The positive estimate of θ also indicates a decreasing marginal clustering effect as the number of bars increases in a given location. As Chart 1 reveals, the marginal clustering effect of having one more bar in a location is the highest when there is only one existing bar, and the effect decreases rapidly as the number of bars increases, to almost negligible when the number of existing bars reaches 6. Moreover, when comparing the marginal competition effect with the marginal clustering effect (note that competition effect on store profit is constant), we find that the two marginal effects equal each other when the number of existing bars is slightly more than 2. This is consistent with the observation that the average number of operating bars in a census tract is 2.19 (Table 2), suggesting that on average, two or three bars in a census tract is probably the equilibrium number that equates the competition effect and clustering effect.

Finally, estimates of the μ 's and σ 's indicate different distributions of the market-specific fixed effects for these two industries. Note that from equation (2.11) the probability of not entering the market is

$$1 - \Pr(\text{entry}) = \frac{1}{1 + \exp(\xi^m) \sum_{k=1}^{J^m} \exp(E\bar{\pi}_k^m)}$$

thus the larger the ξ^m , the more likely that a potential entrant will be able to enter the market. Therefore, a much lower mean of the fixed effect for off-site retailers (μ^{off} at -7.52) than on-site retailers (μ^{on} at 2.74) indicates a much lower probability for a potential liquor store owner to enter the market than a potential bar owner, other things being equal. This may suggest a higher entry cost for opening new liquor stores. The estimates of the standard deviations of ξ^m indicate the heterogeneity of market-specific factors across these five metropolitan areas, and the smaller estimate for bars than that of liquor stores (σ^{on} at 3.55 and σ^{off} at 6.71) may suggest that the heterogeneity across markets for bars is smaller than that of liquor stores.

5. Concluding Remarks

This paper examines the geographical location choices of two types of retailers, on-site retailers and off-site retailers, in a game-theory framework. The two types of retailers have different incentives to cluster: off-site retailers offer a homogeneous product and need to spatially differentiate from each other, whereas on-site retailers can differentiate their products and thus have a weaker incentive to engage in spatial differentiation. We estimate the Bayesian Nash equilibrium of this location game, and quantify the two contradictory effects of clustering and competition in affecting their geographical location choices. This is the first paper in the literature in pursuing a game-theoretic approach to empirically separate the firms' competing and clustering behaviors.

Estimation results are consistent with the theoretical model, suggesting the existence of a positive and significant clustering effect for on-site retailers, in addition to competition effect. Previous studies often cannot separate firms' strategic clustering incentives from the observed clustering due to factors such as population or topographic feasibility. Since detailed information of these factors is usually not available, it is hard to correctly identify the strategic clustering incentives if we focus on only a single industry. Thus in this paper we examine two closely related industries which face similar demand structure and topographic and zoning constraints, the on-site alcohol retailers (bars) and the off-site alcohol retailers (liquor stores). By obtaining different estimates of clustering pattern for these two industries from the same structural model, we show that the on-site alcohol retailers have a stronger strategic clustering incentive than off-site alcohol retailers. Our results support the previous theoretical literature which suggests the existence of firms' strategic clustering behavior.

References

- [1] Berry, Steve, 1994. "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, Vol. 25, pp. 242-262.
- [2] Berry, Steve, James Levinsohn, and Ariel Pakes, 1995. "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, pp. 841-890.
- [3] Caballero, Ricardo J., and Adam B. Jaffe, 1993. "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," *NBER Macroeconomics Annual*, eds. Blanchard, Olivier .J. and Stanley Fischer, MIT Press, pp. 15-74.
- [4] D'Aspremont, Claude, Jean Jaskold Gabszewicz, and Jacques-Francois Thisse, 1979. "On Hotelling's stability in competition," *Econometrica* 47 (5), 1145–1150.
- [5] Deng, Yi, 2008. "The Value of Knowledge Spillovers in the U.S. Semiconductor Industry," *International Journal of Industrial Organization*, Vol. 26, Issue 4, pp. 1044-1058.
- [6] Duranton, Gilles and Henry G. Overman, 2008. "Exploring The Detailed Location Patterns Of U.K. Manufacturing Industries Using Microgeographic Data," *Journal of Regional Science*, Blackwell Publishing, Vol. 48(1), pages 213-243.
- [7] Eaton, Curtis and Richard Lipsey, 1989. "Product differentiation," in: Schmalensee, Richard and Robert Willig (eds.), *Handbook of Industrial Organization*, Vol.1. Elsevier, Amsterdam, pp. 723–768. Ch. 12.
- [8] Ellison, Glenn and Edward Glaeser, 1997. "Geographic concentration in U.S. manufacturing industries: a dartboard approach," *Journal of Political Economy* 105 (5), 889–927.
- [9] Fisher, Jeffrey H. and Joseph E. Harrington, 1996. "Product Variety and Firm Agglomeration," *Rand Journal of Economics*, 27(2), 281-309.
- [10] Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg, 2005. "Market Value and Patent Citations," *Rand Journal of Economics*, Vol. 36, No. 1.
- [11] Hotelling, Harold, 1929. "Stability in competition," *Economic Journal* 39 (153), 41–57.

- [12] Irmen, Andreas and Jacques-Francois Thisse, 1998. "Competition in multi-characteristics spaces: Hotelling was almost right," *Journal of Economic Theory* 78, 76–102.
- [13] Jaffe, Adam and Manuel Trajtenberg, 1996. "Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations over Time and Across Institutional and Geographic Boundaries," *Proceedings of the National Academy of Sciences*, 93: 12671-12677, November 1996.
- [14] Netz, Janet S. and Beck A. Taylor, 2000. "Maximum or Minimum Differentiations? Location Patterns of Retail Outlets," *Review of Economics and Statistics*, February 2002, 84(1): 162-175.
- [15] Pakes, Ariel, Michael Ostrovsky, and Steven Berry, 2007. "Simple Estimators For the Parameters of Discrete Dynamic Games (with Entry/Exit Examples)," *Rand Journal of Economics*, Vol. 38, No.2.
- [16] Picone, Gabriel A., David B. Ridley, and Paul A. Zandbergen, 2009. "Distance Decreases With Differentiation: Strategic Clustering by Retailers," *International Journal of Industrial Organization*, Vol. 27, pp. 463-473.
- [17] Pinkse, Joris and Margaret E. Slade, 1998. "Contracting in space: An application of spatial statistics to discrete-choice models," *Journal of Econometrics*, Elsevier, Vol. 85(1), pages 125-154, July.
- [18] Seim, Katja, 2006. "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," *Rand Journal of Economics*, Vol. 37, No. 3, pp. 619-640.
- [19] Stavins, Joanna, 1995. "Model Entry and Exit in a Differentiated-Product Industry: The Personal Computer Market," *Review of Economics and Statistics*, Vol. 77 (1995), pp. 574-581

Table 1: Characteristics of Five Metropolitan Areas

	Birmingham, AL	Oakland, CA	Tampa, FL	Chicago, IL	Metropolitan Area Minneapolis-St. Paul, MN
Number of census-designated places	43	59	93	273	142
Total number of census tracts in each metropolitan area	202	603	675	2279	798
Largest number of census tracts in each place	19	38	53	428	72
Smallest number of census tracts in each place	1	1	1	1	1
Mean population of designated places (in thousands)	22.804	47.845	26.357	30.789	19.370
Highest population of designated places (in thousands)	68.146	185.974	171.854	998.660	217.109
Lowest population of designated places (in thousands)	4.141	3.736	2.810	1.177	1.842
Mean population in each census tract (in thousands)	4.430	4.892	4.410	4.420	3.874
Median population in each census tract (in thousands)	4.149	4.599	4.145	4.174	3.619
Median non-hispanic white ratio	0.6866	0.4737	0.8024	0.7604	0.8530
Median youth (age 25-34) ratio	0.1270	0.1521	0.1213	0.1405	0.1443
Median household income	39.930	61.507	39.189	61.604	58.467

Table 2: The Observed Clustering Pattern of Bars and Liquor Stores

	Bars	Liquor stores
Total number in five metro areas	2515	1939
Number of census tracts that have at least one bar or liquor stores	1149	1130
Average store number in one tract (conditional on there is at least one store in the tract)	2.19	1.48
Number of census tracts that have two or more stores in the same tract	497	180

Table 3: Model Parameter Estimates

	Single-industry game		Two-industry game
	Bars	Liquor Stores	
$\beta_{population}^{on}$	0.26 (2.97)	—	0.25 (2.71)
β_{whites}^{on}	0.01 (0.44)	—	0.01 (0.34)
β_{youths}^{on}	1.08 (5.36)	—	1.77 (7.12)
β_{income}^{on}	-0.01 (-0.92)	—	-0.01 (-1.37)
$\beta_{population}^{off}$	—	0.44 (9.37)	0.49 (9.50) 0.51 (10.17)
β_{whites}^{off}	—	0.52 (1.83)	0.49 (1.67) 0.23 (1.54)
β_{youths}^{off}	—	0.21 (1.10)	0.21 (1.08) 0.09 (0.75)
β_{income}^{off}	—	-0.00 (-0.57)	-0.00 (-0.55) -0.00 (-0.34)
d_p^{on}	0.21 (2.27)		0.26 (3.07)
d_r^{on}	1.45 (4.93)		1.48 (3.11)
d_p^{off}		1.38 (3.77)	1.14 (3.74) 1.07 (3.05)
d_r^{off}		0.46 (2.04)	0.45 (1.98) 0.45 (3.03)
$\gamma_1^{on_on}$	0.16 (3.10)		0.14 (3.01)
$\gamma_1^{on_off}$	—		0.06 (2.14)
$\gamma_1^{off_on}$	—		0.07 (1.87)
$\gamma_1^{off_off}$	—	0.40 (6.64)	0.32 (5.26) 0.27 (2.87)

Table 3: Model Parameter Estimates (Continued)

	Single-industry game		Two-industry game
	Bars	Liquor Stores	
γ_2	0.61 (3.88)	0.09 (0.63)	0.77 (2.79)
θ	0.77 (2.73)	0.12 (0.47)	0.85 (3.00)
μ^{on}	-4.71 (-3.63)	—	2.74 (4.87)
μ^{off}	—	-6.17 (-5.84)	-7.76 (-6.59)
σ^{on}	4.30 (3.47)	—	3.55 (4.28)
σ^{off}	—	4.64 (4.23)	5.40 (4.92)

Note: t -statistics of point estimates are reported in parantheses.

Table 4: Marginal Effect of Exogenous Variables on Location Probabilities

	Simulated changes	Single-industry Bars	Location Game Liquor Stores	Two-industry game Bars	Liquor Stores
Population	+ 10 percent	0.57%	1.21%	0.65%	1.46%
Median non-hispanic whites ratio	+ 10 percentage points	0.01%	0.26%	-0.00%	0.09%
Median youths (age 25-34) ratio	+ 10 percentage points	0.43%	0.10%	0.13%	0.02%
Median household income	+ 10 percent	-0.24%	-0.02%	-0.32%	-0.02%

Table 5: Marginal Effect of One Additional Retailer on Location Probabilities

	Single-industry Location Game		Two-industry game	
	One Additional Bar	One Additional Liquor Store	One Additional Bar	One Additional Liquor Store
Competition Effect	-0.80%	-0.57%	-0.71%	-0.61%
Clustering Effect	4.85%	—	6.58%	—
Net Effect	4.05%	-0.57%	5.87%	-0.61%

