The role of captive consumers in retailers’ location choice

Anett Erdmann*
Universidad Carlos III de Madrid
aerdmann@eco.uc3m.es

First draft: January 2014, This draft: November 2014

Abstract

This paper investigates empirically the effect of anticipated price competition and distribution costs in the firms’ location choice within an oligopolistic market. I set up a static location-price game of incomplete information in which retailers choose their locations based on (firm-)location specific characteristics, expected market power and the expected degree of price competition. In particular I tie the firms’ strategic location incentives to the population distribution using the concept of captive consumers. This approach is in line with theoretical spatial price competition models and does not require price or quantity data. I address the computational difficulties of the estimation using mathematical programming with equilibrium constraints. Applied to the supermarket industry, the model confirms the existence of benefits of spatial differentiation for firm profits and provides evidence of firms anticipating price competition and distribution costs in their site selection.

Keywords: spatial competition, location choice, price competition, retail competition, discrete games, constrained optimization.

JEL classification: L13, L20, D43, R12, C51.

*I gratefully acknowledge that this research was supported by grant ECO2012-31358 from the Spanish Minister of Education. I thank my advisors, Jesús Carro for his support and guidance during this project and Makoto Watanabe for his advice and many instructive discussions about competitive strategies. I am also grateful to Tobias Kretscher, Matilde Machado, Ulrich Wagner, Jorge Pena Izquierdo, Alex Perez, Helene Perrone and seminar participants at the UC3M Student Seminar, the 2nd WIPE Workshop at Universitat Rovira i Virgili and the XXIX Jornadas de Economía Industrial for their comments and suggestions. Any remaining errors are my own.
1 Introduction

In this paper, I study the site selection of retail firms competing in prices. In particular I analyze whether firms anticipate price wars and distribution costs in their location decision and how the population distribution within a market determines the firms’ location incentives.

I propose a static discrete choice location model under incomplete information where firms compete in locations and prices within local duopolistic markets. The model is formalized as a simultaneous move location game. Since in many cases we observe only the final location decision of the firms, without price or quantity data, I exploit the information inherit in location data in a reduced form profit function using geospatial analysis. In particular, I propose the percentage of ‘captive consumers’ in the firm’s trade area as a new empirical measure of market power under spatial differentiation. With the notion captive I refer to consumers who have access to one firm but traveling to the other firm to do the shopping is out of scale. This is a concept from theoretical spatial competition models, that to the best of my knowledge has so far not received any explicit attention in econometric models. In a similar way, I use the difference in ‘captive consumers’ between rivals as a proxy for the degree of price-competition between the firms. The latter allows to identify whether firms anticipate price competition in order to attract consumers in overlapping market areas. Additionally the model accounts for cost aspects, in terms of the proximity of a store to the firm’s distribution center, which on the one hand serves for the model identification and on the other hand allows to estimate the effect of endogenized fixed distribution costs in the location choice. For the estimation I use a Maximum Likelihood approach and address the computational difficulties inherit in the game theoretic setting through the reformulation of the optimization problem as a mathematical program with equilibrium constraints (MPEC) as suggested by Su and Judd (2012).

I apply the econometric model to study the strategic location determinants for the supermarket industry. To be precise, I use POI (point of interest) data for traditional supermarket stores as a novel type of free available datasets and process the data with the Geographic Information System tool ArcGIS. I find that on average 13% of the firm’s trade area are captive consumers. The location model identifies an incentive of generating local market power through spatial differentiation and firms anticipating price competition as well as distribution costs when choosing a location for their store. Leaving other rivals unconsidered, I find that the second effect of a change in captive consumers, denoted as price-competition effect, only has a negative profit effect if the percentage of captive consumers in the firm’s trade area is small enough (<60%). However, considering the market presence of rivals of a superior format weakens the monopoly power of the firms and I find a clear negative price-competition effect that becomes stronger as the competitive region becomes relatively more important for the firm.
The main contribution of the paper is the explicit consideration of strategic aspects of price competition in a spatial competition model based on observed location data. In particular, I tie the firms’ strategic behavior to the population distribution which has been discussed by Davis (2006) to be long recognized as important link in order to evaluate any policy interest.

The literature of competition models without price and quantity data goes back to Bresnahan and Reiss (1990), who use the fact that under quantity competition a la Cournot, the reduced profit function can be expressed in terms of the number of firms in a market. A latent profit specification is used to estimate a discrete choice market entry model. Katja Seim (2006) extended the model to an entry-location game where firms additionally choose their location within a market. In her approach, the ‘measure of competition’ is the effect of an additional firm in a certain concentric ring (‘donut’) around the store location, that is the corresponding location incentive is independent of the population distribution. In the last years, her model has been extended to differentiation in more than one dimension (Datta and Sudhir, 2013) or allowing for asymmetries in competitive interaction (Zhu and Singh, 2009). However, applying this models to industries where price competition seems more reasonable (e.g. supermarkets), the implicit assumption of quantity competition or a fixed exogenous market price doesn’t allow to identify the appropriate location incentives. The two crucial limitations of this kind of ‘donut-models’ are the following: First, the strong assumption that a rival locating within a certain distance (ring) of the firm has a ‘ring-uniform-competition effect’ disregards the population distribution within a ‘distance ring’. In other words, considering two potential locations of the rival which are at the same distance of the firm but differ in the associated population density, this papers assume that a rival locating at a sparse populated location exercises the same competitive pressure on the firm as if he were located at the dense populated location. Second, the model can only estimate the ‘net effect’ of competition but not the incentives that lead to the observed market structure. While the latter is also discussed in Datta and Sudhir (2011) stating that this structural models are “incapable of separating the ‘net effect’ of competitors into a volume effect and a price competition effect”, in their proposed solution using additionally revenue and price data they rely again on the critical assumption of a ring-uniform-competition effect.

My model offers a way of determining strategic incentives of price competition without additional data and at the same time getting rid of the ‘ring-uniform-competition-effect’ assumption. Incorporating the effect of the population distribution within a trade area, this approach builds on implicit distances instead of accounting for the distance to the competitor straightaway in the profit function.

Methodologically, the paper contributes to the application of recently developed computational methods for the estimation of structural models. So far, the MPEC approach has been shown to be applicable to the structural estimation of dynamic discrete choice
models (Su and Judd, 2012), BLP demand estimation (Fox and Su, 2012) as well as the estimation of static games (Su, 2012). While Vitorino (2012) provides the first application of the MPEC approach to an empirical static binary choice model of market entry, this paper provides an application to a multinomial location choice model.

The paper is organized as follows. First the reader is introduced to some key elements of price competition under spatial differentiation (theory) which will be used in the model. Second, I set up the econometric location model and subsequently explain the estimation method and computation strategy. Finally, I present the data used for the application and report the results. The paper finishes with a comparison of the main findings to alternative approaches in the literature and comments on possible extensions of my framework.

2 Economic intuition

To get an economic intuition about the firms’ strategic price setting behavior under spatial differentiation, let us make use of the Hotelling (1929) framework, the workhorse of theoretical spatial analysis, to highlight some key aspects and to identify observable strategic elements of price competition in space that will be considered later in the empirical model. Consider two firms, A and B, located at the extremes of a linear market. Consumers are uniformly distributed over the line. Additionally, at each extreme, where the firms are located, lives a consumer mass $X_A$ and $X_B$, respectively. Consumers have a unit demand, face displacement costs and decide at which firm to buy maximizing their utility. In addition to this textbook framework, consumers face an exogenous restriction on the travel distance ($D_{\text{max}}$) which reflects their time constraint for shopping.

Figure 1 sketches this toy model. If $D_{\text{max}}$ is large enough, the market area between the two firms can be partitioned into a ‘captive area of firm A’, a ‘competition area’ and a ‘captive area of firm B’. In the following, the notion of captive refers to areas where consumers have only access to one firm since the cost of displacement to another firm is out of scale. The demand of firm A is given as the sum of the consumers that the firm draws from the competitive region and the firm’s captive consumers. Since the firms cannot identify from which region the consumers come when visiting the store and consumers have a unit demand, the firms have to set uniform prices. Solving the simultaneous profit maximization problem, the measure of captive consumers of a firm plays an ambiguous role. An increase in captive consumers causes an increase in the equilibrium price of the firm which reflects the market power effect. But since consumers in the competitive area are rational, buying from the firm that minimizes the overall cost of price and transportation disutility, an increase in the difference of captive consumers with respect to the rival decreases the demand drawn from the competitive area. Thus, for a given number...
of captive consumers of firm B, an exogenous increase in captive consumers of firm A induces the firm to exercise this market power in setting a higher price but the positive effect on the revenues is mitigated through a decrease in the number of consumers drawn from the overlapping market area where competition takes place.

Since the purpose of this toy model is to tell a story of price competition in space, I briefly summarize the main insights here (Appendix A gives an outline of the maths.)

1. An increase in the number of captive consumers of A increases the firm’s price setting power.
   This profit enhancing effect of captive consumers is mitigated through a negative quantity effect on the demand from the competitive area.

2. If the difference in captive consumers between the firms (normalized by the consumers in the competitive area) is small enough, in equilibrium both firms can draw demand from the competitive region.
   But, if the reservation value of the consumers is high enough, there exists a critical percentage of captive consumers in the trade area such that for a higher fraction of captive consumers the firm is better off restricting the demand to the captive area setting the monopoly price.

3. If the number of captive consumers of A is sufficiently high with respect to the captive consumers of B, an increase in captive consumers reduces the revenues from
the competitive area (operating in the elastic section of the demand curve from the competitive area).

An increase in captive consumers of A always increases the total revenues of the firm.

In the following section I transfer this idea to a real geography discrete location-price game assuming firms to anticipate the role of captive consumers when choosing among a finite number of locations to maximize their profits.

3 An econometric spatial location-price game

Analyzing firms optimal location choice empirically, it would be ideal to have access to prices and sales data at the firm level to model the demand side (e.g. Davis, 2006). Unfortunately, this firm specific data are in general not available neither for the researcher nor for the rival firm or any third party (e.g. anti-trust-organization, local government). Inspired by Seim (2006)’s seminal work, I provide a model that exploits the information inherit in the observed location decision of the firms but set up the model in such a way that I fully exploit the population distribution within the market in order to reveal the firms’ location incentives.

3.1 The model

Consider a spatial market $m$ of any polynomial shape with a finite number of equally spaced discrete locations $L_m$ and a corresponding discrete consumer distribution $F_m(X)$ as illustrated in Figure 2.

![Figure 2: Discrete locations in a polynomial market](image)

There are two firms with one store each in the market and each firm faces a discrete choice problem to identify the optimal location which maximizes it’s profit, anticipating the subsequent price competition with the other firm. Assume further that consumers buy from the store for which the price plus the travel cost is the lowest and let them face a maximum exogenous travel distance (radius $D_{max}$) which determines the potential trade
area of a firm.¹ Let us define a matrix $A_m$ of size $L_m \times L_m$ with elements $a_{ll'}$ taking the value 1 if $l'$ is within the trade area of a store at $l$ and 0 otherwise. Assuming that the market within the range of the stores is covered, we can distinguish three scenarios: both firms located at the same location (Bertrand competition), differentiation with overlapping range of influence of the stores (Differentiation with captive consumers) and the case of captive consumers only (Full monopolization). Figure 3 illustrates the most interesting case of differentiation with captive consumers.

Figure 3: Price-competition and Market power in space

![Figure 3: Price-competition and Market power in space](image)

The light gray area depicts the overlapping market range, denoted as the 'area of competition', and the dark gray area illustrates the 'captive consumers' of firm A. The dashed line depicts the analog to the indifferent consumer in the linear model depending on the price setting of the firms. While the total potential demand of a store is the sum of consumers in the distance ring around the store location, the realized demand of A are only the consumers below the dashed line.

Hence, in the simplest framework, the optimal location choice for the store is determined through the potential demand, the market power and the strength (or dominance) of price competition in the competitive region. The intuition for the economic mechanism follows the example from the previous subsection. A higher fraction of captive consumers increases the price setting power and hence the profit per unit sold. But, for a given number of captive consumers of the rival, an increase in captive consumers increases the price difference with respect to the rival and hence decreases the demand drawn from the competitive region. Hence, I expect to find a positive market power effect of captive consumers but a negative quantity effect for the revenues drawn from the competitive area.

¹Defining an exogenous cap on the shopping distance is standard in the empirical literature, e.g. Seim (2006), Datta and Sudhir (2013), Holmes (2011).
However, whether this logic is reflected in the firms’ behavior is an empirical question.

For the econometric specification of the firm’s profit function I follow a reduced form approach. In order to differentiate between the two effects of captive consumers I use two different strategic variables, the absolute number of captive consumers and the difference with respect to the rival. In the simplest way, the profit function of a store of firm $F$ for each location $l = \{1, 2, ..., L_m\}$ is defined as follows: 

\begin{align*}
(I) \quad \pi_{Fl}^I &= \beta_1 \bar{X}_l + \frac{f_{2Fl}(d_{-F})}{\bar{X}_l} X_{\text{captive}Fl} + \frac{f_{3Fl}(d_{-F})}{\bar{X}_l} \Delta X_{\text{captive}Fl} + \delta Z_{Fl} + \omega_{Fl}
\end{align*}

where $\bar{X}_l$ indicates the potential population that can be reached by a store at location $l$ and $Z_{Fl}$ is a firm-specific cost shifter indicating the distance from location $l$ to the closest distribution center of firm $F$. The function $f_{2Fl}(d_{-F})$ indicates the number of captive consumers of firm $F$ located at $l$ for a given location of the rival. The division by the population within the trade area turns the variable into the percentage of captive consumers within the the trade area and hence a measure of market power on the interval $[0, 1]$. The function $f_{3Fl}(d_{-F})$ measures the difference in captive consumers with respect to the rival, as an indicator for the strength of price competition in the competitive area.

Both variables depend on the location structure of the market which is the outcome of the decision of firm $F$ locating at $l$ given that the rival $-F$ is located at $k$. The rival’s location is indicated as a vector $d_{-F}$ of dimension $L_m \times 1$ with elements $d_{-Fk}$ being dummy variables that take the value one if store $-F$ chooses location $k$ and 0 otherwise.

Note that specification (I) assumes a constant marginal effect of the difference in captive consumers. However it seems more reasonable to assume that the competition effect becomes more severe in the location choice as the percentage of consumers in the competitive area increases. Hence, a second specification allows for an interaction effect between the difference of captive consumers and the percentage of consumers within the competitive area.

\begin{align*}
(II) \quad \pi_{Fl}^{II} &= \pi_{Fl}^I + \beta_4 \frac{\Delta X_{\text{captive}Fl}}{\bar{X}_l} \left(1 - \frac{X_{\text{captive}Fl}}{\bar{X}_l}\right)
\end{align*}

The unobservables at the firm-location level $\omega_{Fl}$ are private information of the decision taking firm, captured in the vector $\omega_{mF}$ of dimension $L_m \times 1$. The realization is neither known by the rival nor by the researcher but it is common knowledge that for each market,

\textsuperscript{2}The model abstracts from the outside option for consumers to buy from small mom-and-pop stores. However, in a sensitivity analysis I consider possible rivals of a superior format.
each $\omega_{mF_i}$ is independently, identically distributed extreme value. Hence, considering the information structure of all agents, notice that we as researchers are as informed as the least informed party of the location game.

The information set of firm $F$ when making it’s location decision in market $m$ is $\mathcal{I}_m^F = (X_m, Z_m, \omega_{mF})$, with $(X_m, Z_m)$ being common knowledge among firms and researchers and $\omega_{mF}$ being private knowledge of the firm.

Conditional on $\mathcal{I}_m^F$, the firm forms its belief about the location choice of it’s rival and makes it’s location decision based on expected profits. In the following I use for the beliefs of firm $F$ about it’s rival’s behavior the notation $BP_m^{-F}$, a $L_m \times 1$ dimensional vector of Bayesian probabilities for each possible location $l$. Analog, $BP_m^F$ denotes the beliefs of $-F$ about the location choice of a firm $F$.

Given the profit specification from above, the introduced uncertainty about the rival’s strategy implies forming expectations about $f_2(\cdot)$ and $f_3(\cdot)$. Omitting again the market subscript, the expected profit of specification (I) and (II) respectively is the following:

\begin{align*}
\pi_{Ie}^{F_i} &= \beta_1 \bar{X}_l + \beta_2 \frac{E[X_{\text{captive}F_i}]}{X_l} f_2^{BP-F} \bar{X}_l + \beta_3 \frac{E[\Delta X_{\text{captive}F_i}]}{X_l} f_3^{BP-F} \bar{X}_l + \delta Z_{F_i} + \omega_{F_i} \\
\pi_{IIe}^{F_i} &= \pi_{Ie}^{F_i} + \beta_4 \frac{E[\Delta X_{\text{captive}F_i}]}{X_l} \left(1 - \frac{E[X_{\text{captive}F_i}]}{X_l}\right) + \delta Z_{F_i} + \omega_{F_i}
\end{align*}

with the expectations of firm $F$, based on it’s beliefs about the rival, being the following:

\begin{align*}
f_2^{BP-F}(BP^{-F}) &= \sum_{l'} A(l, l') \cdot (1 - \phi_{l'}^{-F}) \cdot X_{l'} \\
f_3^{BP-F}(BP^{-F}) &= \sum_{l'} (A(l, l') - \phi_{l'}^{-F}) \cdot X_{l'}
\end{align*}

where $\phi_{l'}^{-F}$ is the probability that location $l'$ is covered by the rival firm, that is $P(\text{covered}_{F'-F}=1) = \sum_{k} A_{kl'} \cdot BP_{k}^{-F}$. For a detailed calculation of the structural variables see appendix B.

Note that the profit specification (1) is linear in parameters as well as in beliefs while specification (2) is non-linear in the beliefs. Furthermore note that while the market structure in terms of captive consumers enters directly in the profit equation of both firms, firm-specific variables like the distribution distance have only an indirect effect on the rival’s profit through it’s beliefs.

As can be deduced from the profit equation above, for a profit maximizing firm, it’s best
response depends on the firms believes about the rival’s choice probabilities. The solution concept of the location game is the Bayesian Nash equilibrium such that the equilibrium conditions are the following:

\[
BP^F_l = \Psi^F_l(BP^{-F}, X, Z; \beta, \delta) \quad \forall l
\]

\[
BP^{-F}_l = \Psi^{-F}_l(BP^F, X, Z; \beta, \delta) \quad \forall l
\]

where \( \Psi^F_l \) is a function that defines the choice probability of location \( l \) for a store of firm \( F \), which has to be equal to the beliefs of the rival for any possible location. The analog holds for the rival.

Given the latent profit equations (1) and (2), the choice probability for a profit maximizing firm \( F \) of choosing location \( l \), conditional on being two firms in the market, can be written as follows:

\[
\Psi^F_l \equiv P(d_{Fl} = 1 | BP^{-F}, X, Z, \beta, \delta) = P(\bar{\pi}^e_{Fl} + \omega_{Fl} \geq \bar{\pi}^e_{Fl'} + \omega_{Fl'} \quad \forall l' \neq l)
\]

and under the assumption of \( \omega_{Fl} \) being EV type I distributed:

\[
\Psi^F_l = \frac{\exp \{ \bar{\pi}^e_{Fl}(BP^{-F}, X, Z; \beta, \delta) \}}{\sum_{l'=1}^L \exp \{ \bar{\pi}^e_{Fl'}(BP^{-F}, X, Z; \beta, \delta) \}} \quad (2)
\]

The analog holds for the rival firm \(-F\).

### 3.2 Maximum Likelihood Estimation approach

The estimation of static games with incomplete information implies two main challenges. Once we have chosen an estimation approach, we have to find a way how to solve the game computationally. Second, if there is a chance of multiple equilibria in the model, this has consequences for the computation as well as the identification of the parameters that we aim to estimate based on only one observed equilibrium.

#### 3.2.1 Computational methodologies

As outlined previously, the choice probabilities in an incomplete information game depend on the beliefs about the rivals strategy \( (\Psi^F(BP^{-F})) \). This implies that the likelihood function to be maximized depends on the unknown Bayesian probabilities, a fixed point problem that arises from the equilibrium condition of the game, which makes an iteration on the parameters infeasible without solving at some point for the equilibrium of the game. I will briefly outline the different methodologies that have been developed to address this issue and discuss why I choose the MPEC approach for this problem.
The first computational methodology to address this issue was the Nested Fixed Point (NFXP) algorithm developed by Rust (1987) with a suggested application to static games in Rust (1994). The algorithm solves in each iteration on the parameters for the fixed point of the game providing a full-solution approach. However, the computational burden of this methodology is not only the CPU time but more importantly the trouble in the presence of multiple equilibria. While, based on an assumption about the competitive effect, Seim (2006) could proof the existence of a unique equilibrium for her model and successfully implement the NFXP approach, in the presented model as in many other application this is not the case which implies two problems of this approach: First, if the number of equilibria is unknown there is no way to guarantee that in each iteration all possible equilibria have been found. Second, the number of equilibria may change for different parameter sets which can cause jumps in the likelihood function.

This complications have motivated the development of alternative Maximum Likelihood Methodologies as the two-step methods, going back originally to the dynamic single agent model of Hotz and Miller (1993). This method is based on the idea of estimating in a first step non-parametrically the Bayesian probabilities. In a second step the estimates are used as variables for the beliefs so that the coefficients of the profit function can be estimated using a standard probit or logit model. In other words, the parameters are estimated such that the choice probability is as close as possible to the first stage estimates. Conditioning in the second stage on the equilibrium probabilities from the first stage, which are apparently ‘played by the observed data’, addresses the multiplicity problem and at the same time implies getting rid of the fixed-point problem. However, an important requirement of this method is a consistent estimate at the first stage which is problematic in many applications dealing with small samples and in the present model in particular since the number of possible choices of the stores differs across markets.

Picking up the advantages of this two approaches, Aguirregabiria and Mira (2002) suggest the Nested Pseudo Likelihood (NPL) estimator, which analog the two-step method uses an initial estimate (or guess) of choice probabilities but after estimating the structural parameters computes new choice probabilities and goes on with the iteration on the choice probabilities until convergence is achieved, i.e. swapping the order of the nests of the NFXP algorithm. If the model has more than one equilibria, the authors suggest to use different starting values and choose the outcome with the largest pseudo-likelihood. However, as discussed in Pesendorfer and Schmidt-Dengler (2010), a required assumption to achieve convergence are stable best-response equilibria. Especially they state that already a slight asymmetry in the firms’ payoffs makes it difficult to verify the stability of all the possible equilibria, and this is just the case in my model inherent in the firm-specific distribution distances and implies that this approach cannot guarantee to find the
equilibrium of the model. ³
For a more detailed discussion on the general pros and cons of this three methods for the estimation of discrete games see for example Ellickson and Misra (2011).

In this paper I make use of the recent advances in this field, reformulating the econometric model as a mathematical problem with equilibrium constraints (MPEC) as suggested by Su and Judd (2012).⁴ Their idea is clear and simple: constrained optimization problems are present in many economic applications (e.g. utility maximization subject to budget constraint; transportation problems etc.) but so far, optimization problems in econometrics (regression models) have used unconstrained optimization approaches. The authors show that treating the equilibrium choice probabilities together with the structural parameters as a vector of parameters to be estimated, provides a way of formulating the maximum likelihood approach as a constrained optimization problem that can be solved with any state-of-the-art nonlinear constrained optimization solver (e.g. KNITRO). Consequently there is no need to compute repeatedly equilibria, the stability property of an equilibrium is not an issue and it is relatively easy to implement.

Implementation of the MPEC approach:

Formulating the model as a constrained optimization problem on the joint parameter space \((\beta, \delta, BP)\) can be written as follows:

\[
\begin{align*}
\max_{(\beta, \delta, \{BP_m^F, BP_m^{-F}\}_{m=1}^{M})} & \sum_{m=1}^{M} \sum_{l \in L_m} \left[ d_{ml} \cdot \log(BP_{ml}^F) + d_{m-l} \cdot \log(BP_{ml}^{-F}) \right] \\
\text{s.t.} & \quad BP_{ml}^F = \Psi_{ml}^F(BP^{-F}, X, Z; \beta, \delta) \quad \forall l, m \\
& \quad BP_{ml}^{-F} = \Psi_{ml}^{-F}(BP^F, X, Z; \beta, \delta) \quad \forall l, m \\
& \quad 0 \leq BP_{ml}^F \leq 1 \quad \forall l, m, F
\end{align*}
\]

Note that I assume that the parameters \((\beta, \delta)\) are the same for all markets but the Bayesian Nash equilibrium \((BP_m^F, BP_m^{-F})\) is solved separately for each market.

Given the smooth and concave likelihood function and the fact that the choice probabilities

³Although in a static framework the stability concept may be considered as different from the discussed dynamic framework, note that static games are just a special case setting the discount factor zero. Hence, whenever the initial guess doesn’t exactly coincide with the true equilibrium, a small perturbation is enough to make it impossible for the algorithm to reach that equilibrium if it is an unstable one.

⁴An example for a static discrete-choice game of market entry is provided by Su (2012) and a first real application by Vitorino (2012).
of potential locations are strictly bounded on \([0 + \epsilon, 1]\), for any parameter vector \((\beta, \delta)\), the existence of an equilibrium is guaranteed by Browers Fixed Point Theorem.

To solve this optimization problem taking into account the high dimensionality of the problem, I use the KNITRO solver through MATLAB.

3.2.2 Multiple equilibria and Identification

While the existence of an equilibrium is guaranteed, let us consider the potential multiplicity of equilibria. The multiplicity can come either from the identity of the firms or the distribution of location characteristics within a market.

First, contrary to Seim (2006)’s approach, in the present model, I don’t assume firms to be completely symmetric so that the identity of a firm that chooses a certain location matters. Both firms face an analog problem but the distance to the closest distribution center is firm specific and so are the equilibrium choice probabilities. However, using for the estimation a maximum likelihood approach, through the maximization of the overall likelihood, this location-firm specific characteristics serve as a kind of implicit equilibrium selection rule respective the identity of the firms.\(^5\) Hence, the availability of firm-specific location characteristics becomes a necessary data requirement to deal with the multiplicity inherit in the firm identity (Data requirement 1).

Second, for some distributions of location-characteristics and the true parameters \((\beta^*, \delta^*)\), there may be more than one local equilibrium but I observe only one in each market. In this aspect, we follow the standard assumption in the literature that for markets with the same (exogenous) observable characteristics, firms coordinate on the same equilibrium (Assumption 1). That is, I admit the possible existence of multiple equilibria but assume that there are no multiple equilibria played in the data such that the multiplicity issue doesn’t hinder the identification of the equilibria.

As commented earlier for the NFXP approach, the multiplicity of equilibria also goes along with computational challenges, in particular inherit in the repeated solving of the game. Using the MPEC approach, I optimize on the joint parameter space of structural estimates and beliefs solving the game only once which overcomes the problems associated with repeatedly solving the game (for a detailed discussion see Su (2012)). However, analog other numerical optimization algorithms, this approach can only find a local optimum which does not need to coincide with the global one so that the challenge of finding all equilibria remains. In order to increase the probability of finding the best equilibrium in terms of the highest log-likelihood, I use many different initial values.

Respective the identification of the parameters in the model, I exploit the variation of

\(^5\)Zhu and Singh (2009) discuss the usage of firm-specific variables, like the distance to the closest distribution center, in another context. They set up a model with firm-specific parameters and make use of distances to firm-specific facilities as exclusion restrictions to guarantee parameter identification.
general location-characteristics, firm-specific location attributes within markets and the variation in the distribution of the characteristics across markets together with the observed store locations. With respect to the strategic effects, we need the identification requirement that the markets are large enough or $D_{\text{max}}$ small enough such that $A_{kl} = 0$ for at least one $l$, $\forall k$, $\forall m$ (Data requirement 2). In other words, there is no location from where the firm can serve the whole market. A weak requirement that prevents any collinearity problem between the strategic variables.

Furthermore, I make the strong assumption that any kind of market effects are uncorrelated with the market structure as well as the population distribution and such that non-negative profits for the firms are guaranteed. Accounting for this unobserved heterogeneity across markets is at the moment considered as computationally too expensive.

### 3.2.3 Coherence with the theory

Considering the coherence of the estimates with the theoretical intuition outlined initially, the arguments go as follows. First, if there were no interaction between the firms, the only profit determinants would be the potential consumers within the trade area and the cost structure. Second, if firms competed for market shares and prices were exogenously given, additionally the number of 'captive consumers' should enter positively in the profit function, but the difference in captive consumers, as a proxy for price differences, should be irrelevant in this context. Third, if firms anticipated price competition in their location choice, a positive difference of captive consumers with respect to the rival is supposed to decrease the demand drawn from the competitive region and hence should enter with a negative sign in the profit equation.

### 4 Data description

In my application I consider the location choice of the two strongest (traditional) supermarket chains in the U.S., Kroger and Safeway, whenever they encounter each other in a local market. This example has been chosen since statistically, both firms seem to target the same type of geographic markets and consumers and sell similar grocery products. Hence, abstracting from some preferences over one or another private label which is not part of this paper, the products of the firms can be assumed to be perfect substitutes. To set up the necessary dataset for the analysis, I use four types of data sets: observed store locations, locations of distribution centers, spatial administrative units for the market definition and spatial subunits (smaller than the market definition) with associated population characteristics, that I combine using the geographical information system ArcGIS. First, taking advantage of the advances in consumer services for GPS users, I use POI datasets ('Point of interests') for GPS users to identify the store locations of the two firms.
as well as their primary rivals of a superior format, Walmart and Target. The advantage of this type of data source is that locations are already geocodified to an eight-digit latitude/longitude format and can directly be imported into the geographic information system that I use for the analysis and hence avoids any type of matching problems.

A second data set identifying the locations of regional distribution centers is constructed using information from the firms’ website. Making use of the GIS North American Address Locator I geocode the street addresses of the distribution centers in a latitude and longitude format analog the store dataset.

A third type of dataset, which is provided by the GIS online library, contains borderline definitions (in polygon format) of different administrative spatial units which I use for the market definition. Using the insights from my previous paper where I find that 90% of all stores of the considered firms are located within Urban Areas, densely populated regions, I use Urban areas as the market definition (for a more detailed discussion see my paper "Hotelling meets Holmes").

Finally, a fourth data set contains all Census Block Groups in the US, as the smallest available geographic unit for which associated population characteristics are available. This data set is available from the U.S. Census Bureau and provided in a shape file format with associated demographic characteristics by GIS. By construction, the Block Groups capture contrary to larger spatial units, relatively homogeneous population clusters.

Furthermore, I need to know the maximum radius from where a store draws consumers (range of influence), which is taken from the Kroger Fact book, stating that it’s supermarkets "typically draw customers from 2.0-2.5 mile radius”. I use the upper bound setting $D_{max} = 2.5$ miles and assume that for Safeway supermarkets holds a similar range of influence.\(^6\)

Before combining this available information, I project each of the four data sets on a x-y Cartesian coordinate system (Albers Equal Area Conic Projection) which builds the reference system for the spatial analysis. Furthermore in this paper, I restrict the analysis to ‘Urban Areas’ which are sufficiently far from each other (‘isolated’) to guarantee that consumers patronize only stores in the own market.\(^7\) Given this database, I conduct the discretization of the locations. First, I discretize the potential store locations inside a market defining over each market a grid of equally sized cells of 1.0x1.0 square mile, which is small enough to fulfill the identification assumption of the model and has the convenience that the population in each cell corresponds to the population density of the

\(^6\)Note that the construction of the variables rest on the definition of the exogenous radius of influence. Since this is the case in most of the empirical spatial competition models and for a small perturbation of the radius I do not expect any change in the conclusions, I rely on the information by Kroger and abstain from testing alternative values.

\(^7\)With ‘sufficiently far’ I refer to markets for which the range of influence of each location contains exclusively locations from the own market.
Figure 4: Data visualization for some sample markets

(a) Discretized locations

(b) Population distribution

(c) Trade areas
associated BG which is measured in pop/sqmi. Then I define the centroids of the cells as possible locations. For computational reasons I exclude markets with more than 500 potential locations. This data set deals with a set of 70 isolated markets with the presence of both firms, however I center my analysis on the 31 urban markets with two competing stores, one of each firm. On average those markets consist of 34 potential locations, with the smallest market counting 12 locations and the largest 112.

Now I augment the discretized market dataset combing each location with the associated block group characteristics, the observed store locations and compute the Euclidean distance of each location to the closest distribution center of each firm and to the closest big-box store, considering Walmart and Target. Figure 4a visualizes the discretized structure for three example markets and Figure 4b the associated population distribution and observed store location(s) of Kroger and Safeway as dot and triangle respectively. Note that the sales potential is not uniformly distributed within the neighborhood of the store, which motivates my approach of constructing a strategic variable that depends on the population distribution rather than defining a uniform radial competition effect and accounting for the total population within the trade area only as a covariate in the profit equation.

As defined by the model, the construction of the strategic variables relies on delimited trade areas of a 2.5 miles radius. Hence, I construct a distance matrix that measures the Euclidean distance from each location to any other location within the same market, which is than used to construct the feasibility matrix $A_m$ for each market. Figure 4c illustrates the feasibility of consumer locations for given store locations using distance rings of radius 2.5 miles to define the trade area. Then I export the dataset of discrete locations and it’s associated variables as well as the distance matrix to Matlab. Table I provides the descriptive statistics of the variables of interest at the observed store locations for the set of markets with one store per firm.

Considering the exogenous variables of the model, $\bar{X}$ defines the total population within a 2.5 miles radius of the store measured in 1000. The variable $Z$ indicates the distribution distances to the closest distribution center of the respective firm measured as the Euclidean distance in 100 miles. $BB_{distance}$ is the distance to the closest big-box store of either Walmart or Target, measured in 100 miles from the store location. $av.Age$ and $av.HHsize$ are the average age and the average household size of the population within the store’s trade area.

Respective the endogenous variables of the model, $X_{captive}/\bar{X}$ indicates the fraction of captive consumers for the store and $\Delta X_{captive}/\bar{X}$ defines the difference in captive consumers with respect to the rival, normalized by the population of the trade area of the

---

8Choosing the size of the cells yields a trade-off between accuracy and speed of the algorithm. Vincentini(2012) follows in his dynamic model a similar approach dividing the city of Greensboro into cells of 2.25 square miles (1.5x1.5).
firm.

The data indicate that for the considered firms, on average 13% and 14% respectively of the population in the trade area are captive. Note that we observe complete monopolization as well as markets with firms located at the same location. Considering the difference in captive consumers with respect to the competitor, there is no statistical difference in the means of the two firms implying that statistically there is no dominance in terms of captive consumers of one or the other player.

Since this statistics are the outcome of the location decision but the decision taking is modeled as an incomplete information game, note that the domain of the expected number of captive consumers (corresponding to the range of \( f_2 \)) is \((0, 1)\), which is due to the positive choice probabilities for each location alternative and the data requirement 2. Additionally, I check the correlation between the number of captive consumers of the two firms providing Pearson’s linear correlation coefficient \( \rho_{\text{captive}} \). This is necessary for identification. If for example one firm always established in the city center and the other one situated closer to the border the profit specification (2) would suffer from multicollinearity. However, I find that there is no such significant correlation in the data.

Appendix C provides some summary statistics about the population distribution within markets. Note that for some locations, due to urban restrictions (e.g. parks), the population can be zero, but as can be seen from Table I, the population of a trade area is never zero so that the endogenous variables will be always defined.
<table>
<thead>
<tr>
<th>Variables</th>
<th>observed outcomes</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kroger</td>
<td>Safeway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>st.dev.</td>
<td>min</td>
<td>max</td>
<td>mean</td>
<td>st.dev.</td>
<td>min</td>
<td>max</td>
<td></td>
</tr>
<tr>
<td>exogenous variables:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>15.0625</td>
<td>(10.7861)</td>
<td>0.8196</td>
<td>44.3463</td>
<td>15.2931</td>
<td>(11.2544)</td>
<td>0.8196</td>
<td>44.3463</td>
<td></td>
</tr>
<tr>
<td>$Z$</td>
<td>1.1495</td>
<td>(0.7075)</td>
<td>0.2841</td>
<td>3.2698</td>
<td>1.1615</td>
<td>(0.6229)</td>
<td>0.2556</td>
<td>2.8767</td>
<td></td>
</tr>
<tr>
<td>$BB_{distance}$</td>
<td>0.0644</td>
<td>(0.1188)</td>
<td>0.0012</td>
<td>0.4672</td>
<td>0.0597</td>
<td>(0.1121)</td>
<td>0.0017</td>
<td>0.4242</td>
<td></td>
</tr>
<tr>
<td>$av_{Age}$</td>
<td>39.4228</td>
<td>(6.7055)</td>
<td>25.0114</td>
<td>54.4940</td>
<td>39.4624</td>
<td>(6.7941)</td>
<td>25.0114</td>
<td>54.5760</td>
<td></td>
</tr>
<tr>
<td>$av_{HHsize}$</td>
<td>2.3810</td>
<td>(0.2217)</td>
<td>1.9098</td>
<td>2.8331</td>
<td>2.3672</td>
<td>(0.2111)</td>
<td>1.9844</td>
<td>2.7899</td>
<td></td>
</tr>
<tr>
<td>endogenous variables:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{captive}/\bar{X}$</td>
<td>0.1280</td>
<td>(0.2613)</td>
<td>0</td>
<td>1</td>
<td>0.1392</td>
<td>(0.2474)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\Delta X_{captive}/\bar{X}$</td>
<td>-0.0756</td>
<td>(0.3196)</td>
<td>-1.5661</td>
<td>0.3780</td>
<td>0.0177</td>
<td>(0.2132)</td>
<td>-0.6078</td>
<td>0.6103</td>
<td></td>
</tr>
<tr>
<td>$\rho_{captive}$</td>
<td>0.1791</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>($p - value$)</td>
<td>(0.3349)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5 Estimation Results

Estimating the model as outlined in section 3, Table II reports the estimated parameters in the profit function of the firms for model specification (1) and (2). In order to evaluate the significance of the parameters and test the coherence with the theory, I use the Bootstrap percentile method. I generate for each specification 300 resamples with replacement from the original set of markets, solve the problem for each sample and calculate the percentile confidence intervals for the parameters. Appendix E provides the details on the bootstrap distributions.

Table II. Estimation Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>without rivals</th>
<th>with Big Box rivals</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>without rivals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (1)</td>
<td>0.2794**</td>
<td>0.3473*</td>
</tr>
<tr>
<td>model (2)</td>
<td>0.2909**</td>
<td>0.2649**</td>
</tr>
<tr>
<td>model (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X_captive</td>
<td>1.5977*</td>
<td>1.3320**</td>
</tr>
<tr>
<td>X_captive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (1)</td>
<td>0.3624**</td>
<td>0.1824**</td>
</tr>
<tr>
<td>model (2)</td>
<td>0.3548**</td>
<td>0.3469</td>
</tr>
<tr>
<td>model (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆X_captive</td>
<td>-0.2282**</td>
<td>-0.1824**</td>
</tr>
<tr>
<td>∆X_captive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (1)</td>
<td>-0.9113**</td>
<td>-0.8702**</td>
</tr>
<tr>
<td>model (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆X_captive × (1 - X_captive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆X_captive × (1 - X_captive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (1)</td>
<td>-1.5302*</td>
<td>-1.6297*</td>
</tr>
<tr>
<td>model (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model (3)</td>
<td>-1.6594*</td>
<td>-2.0040**</td>
</tr>
<tr>
<td>model (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>-1.928</td>
<td>0.5026**</td>
</tr>
<tr>
<td># Iterations</td>
<td>25</td>
<td>135</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-149.6538</td>
<td>-142.6685</td>
</tr>
<tr>
<td></td>
<td>-147.6514</td>
<td>-138.9203</td>
</tr>
</tbody>
</table>

* Significance at the 10% level. ** Significance at the 5% level.

Model (3) and (4) provide a robustness check of the results with respect to other rivals. Further robustness checks, with respect to the specification of the distribution costs and some demographic characteristics of the potential consumers, turned out to be worse in terms of the log-likelihood and the convergence properties (see appendix Table III).

The baseline model (other rivals disregarded):

Population distribution. The population within the trade area of the stores (measured in 1000) has a significant positive effect on the location choice of the firm which captures the attractiveness of dense populated areas.

Market power and price-competition effect. The positive effect of a high fraction of captive consumers in model (1) as well as in (2) captures the market power effect. The bootstrap analysis for model (1) suggest that we can be at least 95% certain that the structural estimates are consistent with the outlined economic intuition, i.e. a positive effect of the percentage of captive consumers and a negative profit-effect of the difference with respect to the rival. Given a certain population in the firm’s trade area, the higher the percentage of captive consumers, the larger is the profit of the firm, which can be justified by an
increased price-setting power of the firm. However, the negative effect of the difference of captive consumers with respect to the competitor, which captures the price difference of the firms, suggests that an advantage in captive consumers with respect to the rival has a negative effect on the firm’s profit. Where this effect comes from becomes more clear when considering model specification (2), which allows for an interaction effect with the percentage of consumers living in the competitive area, those who care about price differences when choosing at which store to buy. While the effect of the difference in captive consumers becomes positive, the interaction effect indicates that this effect decreases with the fraction of consumers in the competitive area. Considering the total effect of the difference in captive consumers, I find that if the fraction of consumers in the competitive region is above a threshold of 40% then an increase in the difference of captive consumers has a negative profit effect. That is, contrary to my expectations, I find that an increase in the strategic variable which captures the price difference between firms does not always have a negative profit effect but depends on the market structure. I will discuss this later with more detail.

Distribution costs. Considering the cost effect, as expected, I find a significant negative profit-effect of the distribution distance which is consistent with other retail-studies (e.g. Vitorino (2012), Zhu and Singh (2009)) and confirms the findings in Erdmann (2013).

Presence of other rivals:

An other important issue in the present competition analysis are other grocery retailers, rivals which are not on a par with the firm but are able to ‘steal’ a significant part of potential consumers of the firm. Safeway classifies its competitors in primary conventional supermarkets and other rivals like big-box stores and warehouses or discounters (Safeway Fact Book 2011). In my analysis I control for the market presence of the big-box stores Walmart and Target, as rivals of a superior format which have repeatedly be demonstrated to have an effect on the conventional supermarket competition (e.g. Jia (2008), Matsa (2011)). Model (3) and (4) account for the distance of a supermarket location to the closest big-box retailer. Considering model (3), the presence of this rivals is not significant. However, note that compared to model (1), the market power effect as well as the competition effect decrease a bit in absolute terms, which may suggests that the isolated analysis without the consideration of other rivals slightly overestimates the strength of competition between the two firms. Contrary, the distribution cost effect becomes slightly stronger in absolute terms.

---

9 Other conventional rivals, which are not operated under the Kroger Company, are stores operated by SuperValu or Cerberus. A shortcoming of my analysis is that I lack data on this rivals so that I focus on the two main supermarket operators with a significantly higher market share.
Model (4) in turn, which yields the largest log-likelihood, identifies a significant positive profit effect of the distance to the closest superstore. This implies that the competitive pressure of this format diminishes with the distance to the store. Note also that accounting for the presence of other rivals, the difference in captive consumers is not any more significant while the interaction term with the fraction of consumers in the competitive area remains negative significant. This results suggest that the threshold argument from model (2) doesn’t hold any more when I control for other rivals. In other words, taking into account the presence of other rivals, I find a clear negative price competition effect that becomes stronger as the competitive area becomes relatively more important for the firm.

Finally, considering all the identified profit determinants, note that the 'hunt for captive consumer’ can outweigh the attraction of dense populated locations but given a strong position of the rival in terms of captive consumers, locating close but in a less attractive area, the firm can gain a large fraction of the consumers in the competitive area which may be an attractive strategy if the competitive area is sufficiently densely populated. Taking both arguments together, the model can explain observed spatial segmentation as well as observed spatial closeness, for example with one firm in a high populated area and another one close-by.

6 Discussion

6.1 On the role of captive consumers

I have proposed a model that uses the measure of 'captive consumers' to draw inferences on the incentives that lead retailers to a certain location decision if anticipating price competition. The application to supermarket data of Kroger and Safeway has revealed that the percentage of captive consumers in a retailer’s trade area has a significant positive profit-effect. This is consistent with the initial reasoning that a high percentage of captive consumers induces the firm to set a higher price. I also find that the difference in captive consumers can have a negative profit effect which depends on the market structure. To interpret this finding which is not that straightforward, I make use of the toy model of a linear city. First, as argued previously, a higher price difference is supposed to induces a loss of consumers in the overlapping market area which explains the negative impact from model (1) but not the dependence on the market structure. My intuition is that the results of model (2) suggest that with an increasing percentage of captive consumers, the consumers in the competitive area become less important for the firm up to a point of ignorance focusing only on the market with captive consumers. Note that this is in line with the toy model of a linear city. As mentioned earlier, for a sufficiently high reservation price of the consumers, whenever the percentage of captive consumers exceeds
a certain threshold, it is more profitable for the firm to set the monopoly price in the captive market rather than competing in prices to draw consumers from the competitive area. However, the presence of other rivals provides an outside option for the consumers and hence debilitates the firms monopoly power that much that the price difference with respect to the rival is solely relevant within the overlapping market area.\(^\text{10}\)

Note that, while the 'market power effect' could also be justified under Cournot competition, the 'difference in captive consumers', which captures price differences, is characteristic to Bertrand competition. Hence, I take the results as evidence that the analyzed firms do compete in prices and anticipate this competition in their location choice.

6.2 Comparison to other studies

Comparing the results to other game-theoretic location studies, note that the notion of 'returns to spatial differentiation' is similar to the concept of 'percentage of captive consumers'. Hunting for captive consumers goes necessarily along with spatial differentiation but additionally accounts for the population distribution over space.

In order to contemplate the difference of the present approach to studies using uniform radial competition effects, let us consider the model by Datta and Sudhir (2013), an extension of Seim (2006), with endogenous location choice and endogenous types of stores. Although in my model the type (firm) is given exogenously and restricted to markets with one store per firm, I use this example to illustrate the missing feature when firms compete in prices. Simplifying the model to a market with two firms only and adapting the notation to the one used above, allows a direct comparison of their profit specification:

\[
\pi_{Fl}^e = \gamma_1 \bar{X}l + \gamma_2 E[N_{-F,b=1}|Fl] + \gamma_3 E[N_{-F,b=2}|Fl] + \delta Z_{Fl} + \zeta + \omega_{Fl}
\]

where \(h_2\) is the probability that the rival locates within a distance of up to \(D_2\) mile, \(h_3\) the probability that the rival locates within a distance of \(D_2 - D_3\) miles from firm \(F\) and \(\zeta\) is a market fixed effect. Note that this setting assumes that any rival location in a certain distance band of the store has the same competitive impact. If the neighborhood of a store location were characterized by local homogeneity in terms of the population distribution, this concentric ring approach would be unproblematic. However, as illustrated in Figure 4, in many geographic markets this is not the case. That is, competitors located at different potential locations within a certain distance of the store count a different number of captive consumers and consumers in the overlapping market area with the

\(^{10}\)Note that the identified threshold of 40% in model (2) is determined by the distance bound parameter that I have based on the firm information by Kroger. A sensitivity analysis with respect to this ad-hoc parameter may be appropriate. However, since controlling for other rivals this effect vanishes, I consider this as unnecessary for my results.
store and hence I expect them to exercise a different competitive pressure on the store. In other words, this specification ignores the effect of the population distribution on the price setting power of the firm. If you nevertheless prefer the 'donut-approach' over the model proposed in this paper, as an alternative to account for location specific competition effects, I suggest to define any measure of competitive pressure for each location in the respective donut and weight the expected number of stores within a donut by this competitive strength.

Note that the limitation of the radial approach comes from a direct transfer of the consumer behavior to the firm behavior which is not necessarily correct. Specifying a differentiated product demand model (e.g. Davis (2006), de Palma et al. (1994)), it is reasonable to assume that whenever products are only differentiated in their geographic location, that consumers indifference curves are concentric circles around their locations. However, when the firms are choosing locations which implies reaching some consumer locations and others not, their 'indifference curves' which are isoprofit curves are not necessarily concentric rings. This comes from the fact that for the firm the population distribution matters in it’s choice while when analyzing consumer behavior the individual decision is independent of the population distribution (unless in the case of network products).

Demand estimation literature The importance of accounting for the population distribution when empirically measuring strategic effects, has also be emphasized for the estimation of a structural demand models in space. Using firm locations and price data, Davis (2006) estimates a retail demand model under spatial differentiation using a BLP-approach. Beyond the typical BLP-instruments, employing product characteristics of the rival, he exploits the spatial structure of the demand using population counts in the close locality of the rival as valid instrument for the prices. Note that implicitly this idea is in line with the concept of captive consumers.

Likewise, the literature on gravity models allows a comparison to our results. For an overview see Anderson et al. (2009). This models go back to Reilly’s Law of Retail Gravitation and later Converse’s revision in order to define a breaking point between to retailers that defines the ‘indifferent consumer’. This approach defines the ability of a firm location to attract consumers from a third (competitive) area as a decreasing function of the distance and an increasing function of the population at the store location. Note that the latter contradicts our argument. Their argument which predicts larger ‘competitive demand’ for locations with a higher population is based on the ‘agglomeration’ principle. However, given the difference in retail pattern in metropolitan areas Mason and Mayer (1990) argue that Reilly’s model works well in rural areas but not in urban areas and propose to invert the breaking-point formula such that the demand drawn from the competitive area increases as the population density decreases. Note that this is in line with my findings with the difference that I base my arguments on a game theoretic framework.

24
6.3 Limitations and further research

My model has the following limitations. First, by the nature of the model and the computational methodology, I have identified a local maximum. Although I have run the model with many different starting points, I cannot guarantee that the equilibrium found is also global. Second, the study is limited to the competition between two firms operating one store each. My conjecture is that the main result is similar for markets with more than one store per firm but this generalization would require some additional information on the firm’s pricing practice across stores within a local geographic market. Firms can either follow a uniform pricing strategy, setting the same price for all stores within a geographic market, or practice price flexing, setting different prices across stores of the same chain. Depending on the strategy played by the firms, Krčál (2012) shows that the outcome in terms of firm locations and shopping costs incurred by consumers can differ substantially. Unfortunately for the application to the US supermarkets there is no evidence about the local pricing strategy of a supermarket operating more than one store within a market. Furthermore, I have focused on a covered trade area, which allows for a straightforward comparison to the modified Hotelling version to interpret the results. Relaxing this model assumption, specifying the consumer attraction as a decreasing function of the distance to the store, for instance using a retail gravity model, is not expected to change the results but may provide additional insights.

7 Conclusion

I have provided an econometric location model under price competition that can be estimated with publicly available location data and the population distribution at the smallest possible unit. In the application to supermarkets I find evidence on price competition, in particular that firms anticipate the degree of price competition in their location choice. I also find that firms consider distribution costs when choosing a location and confirm that geographic differentiation from the competitor can increase profits.

As policy implication, the local antitrust authorities may use the outlined mechanism to set up appropriate zoning restrictions in order to avoid excessive market power and promote a high degree of price competition.

Last but not least, I hope that my analysis motivates also the location analysis in the business practice to take the outlined strategic location determinants into account.
References


A Toy model

Here I provide some exercises and the main insights from using the Hotelling framework as a toy model, as illustrated in Figure 1. This exercise is especially relevant for the theoretical understanding of the firms’ strategy and will be useful for the interpretation of the empirical results.

Normalizing the competitive area to one, i.e. $\bar{A}B - 2a \equiv 1$ so that $\text{Comp} = 1$, $\hat{X}_A = \frac{X_A}{\text{Comp}}$, $\hat{a} = \frac{a}{\text{Comp}}$, $\hat{X}_B = \frac{X_B}{\text{Comp}}$, the demand of firm A is defined as $\hat{D}_A = \hat{X}_A + \hat{a} + (\hat{x} - \hat{a})$. The last term defines the demand drawn from the competitive region which is specified by the indifferent consumer as usual. But contrary to the standard Hotelling framework, we may have situations where only one firm draws demand from the competitive area. That is, $\hat{x} - \hat{a} = \frac{1}{2} - \frac{p_a - p_b}{2t}$ if $|\Delta p| \leq \frac{1}{2}$, $\hat{x} - \hat{a} = 0$ if $\frac{p_a - p_b}{2t} > \frac{1}{2}$ and $\hat{x} - \hat{a} = 1$ if $\frac{p_a - p_b}{2t} < -\frac{1}{2}$.

Suppose for a moment that both firms draw demand from the competitive area. That is, we may have situations where only one firm draws demand from the competitive area. By the indifferent consumer as usual. But contrary to the standard Hotelling framework, we may have situations where only one firm draws demand from the competitive area.

That is, $\hat{x} - \hat{a} = \frac{1}{2} - \frac{p_a - p_b}{2t}$ if $|\Delta p| \leq \frac{1}{2}$, $\hat{x} - \hat{a} = 0$ if $\frac{p_a - p_b}{2t} > \frac{1}{2}$ and $\hat{x} - \hat{a} = 1$ if $\frac{p_a - p_b}{2t} < -\frac{1}{2}$.

Suppose for a moment that both firms draw demand from the competitive area. Then, solving the firm’s optimization problem $Max \{p_a(\hat{X}_A + \hat{a} + \frac{1}{2} - \frac{p_a - p_b}{2t})\}$, maximizing over $p_a$ the best response of the firm is $p_a = t(\hat{X}_A + \hat{a} + \frac{1}{2}) + \frac{1}{2}p_b$ and analog for firm B.

Solving the simultaneous equation system, the optimal pricing strategy for firm A becomes $p_a^* = \frac{4}{3}t(\hat{X}_A + \hat{a}) + \frac{2}{3}t(\hat{X}_B + \hat{a}) + t$ and analog $p_b^* = \frac{4}{3}t(\hat{X}_B + \hat{a}) + \frac{2}{3}t(\hat{X}_A + \hat{a}) + t$ such that the prices are a function of the travel cost parameter $t$ and the number of captive consumers. Hence the demand that A draws from the competitive region becomes $\hat{x} - \hat{a} = \frac{1}{2} + \frac{1}{3}\Delta \hat{X}_A$. This implies that both firms target the competitive area iff $|\Delta \hat{X}_A| \leq \frac{3}{2}$ and generate profits from the competitive area of captive consumers in it’s as well as from the competitive area ($\pi_{A1}$), i.e. $\pi_A = \pi_{A1} + \pi_{A2} = p_a^*(\hat{X}_A + \hat{a}) + p_a^*(\frac{1}{2} - \frac{1}{3}\Delta \hat{X}_A)$. But, if $\Delta \hat{X}_A < -\frac{3}{2}$, firm A will receive the whole demand from the competitive area while B’s optimal strategy is generating revenues only from it’s captive consumers setting the monopoly price. Considering only the revenues generated from the competitive area, I calculate the demand elasticity of competitive consumers as $\epsilon(\hat{x} - \hat{a}) = -\frac{1}{3}\frac{\hat{X}_A + \hat{a}}{(\hat{x} - \hat{a})}$. Whenever $\hat{X}_A + \hat{a} > \frac{1}{2} + \frac{\hat{X}_B + \hat{a}}{2}$, the demand is elastic so that an increase in captive consumers reduces the revenues from the competitive area.

Figure 5 illustrates this situation.

Alternatively, normalizing the trade area of the firm to one, i.e. $X_A + a + \text{Comp} \equiv \bar{X}_A = 1$, allows to interpret the firm’s strategic behavior as a function of the percentage of captive consumers in it’s trade area $(X_A + a)/\bar{X}_A$ and the normalized difference in captive consumers $\Delta X_A/\bar{X}_A$ respectively. Under this normalization, I ask whether there exists a critical number of captive consumers for which the firm is better of setting the monopoly price for it’s captive consumers instead of engaging in price competition in the competitive area. This is equivalent to ask whether there is a solution to $\pi_A^M \geq \pi_{A1} + \pi_{A2}$.

Hence, denoting $R$ as the consumers reservation price and solving the game, the inequality becomes $R(X_A + a) \geq p_a^*(\frac{X_A + a}{X_A} + \frac{\hat{x}}{X_A})$ with $p_a^* = \frac{4}{3}t\frac{X_A + a}{X_A} + \frac{2}{3}t\frac{X_B + a}{X_A} + t(1 - \frac{X_A + a}{X_A})$ and $\hat{x}^* = \frac{(1 - \frac{X_A + a}{X_A})}{2} - \frac{1}{3}(\frac{X_A + a}{X_A} - \frac{X_A + a}{X_A})$. For any given number of captive consumers of the rival
there exists an upper bound on the percentage of captive consumers \((X_A + a)/\bar{X}_A,\)

such that for a sufficiently high reservation price of the consumers (=monopoly price),

the firm is better off focusing on the captive consumers to extract their surplus instead

of competing over the competitive area. For instance, suppose that \(X_B + a = 0,\) then

the inequality above can be written as a quadratic equation that has a solution if the
discriminant \(D = (\frac{1}{3} - R)^2 - 4 \cdot \frac{1}{18} t \cdot \frac{1}{2} \geq 0.\) Setting \(t = 1,\) a solution exists if \(R \geq \frac{2}{3}.\)

For example, setting \(R = 1\) implies that a fraction of captive consumers higher than 80

% induces firm A to set the monopoly price although for a fraction of captive consumers

less than 120 % both firms could draw a positive demand from the competitive area.

Figure 5: The effect of an increase in captive consumers

![Figure 5: The effect of an increase in captive consumers](image)
B Detailed calculations of the structural variables

Since the model with one store of each firm is just a special case of the extension to markets with $N$ stores, I provide here the calculation for the general case with $s(F)$ denoting a store with firm affiliation $F$ and assuming that prices are set at the market-firm level while the location choice takes place at the store level.

The number of consumers within a maximal travel distance $D_{\text{max}}$ who may patronize the store at $l$ is the same for all stores and independent of the rivals’ choice.

$$X_l = \sum_{l':d(l,l') \leq D_{\text{max}}} X_{l'}$$

In order to compute the expected number of captive consumers and the difference to the competitor, let us first consider those variables under full information.

The total number of captive consumers for chain $F$ (cannot reach any store of the rival chain) can be written as follows

$$X_{\text{captive}_F} = \sum_l \text{captive}_F l \cdot X_l \equiv f(d_F, d_{-F}, A, X)$$

with

$$\text{captive}_F l = \begin{cases} 
1 & \text{if } \sum_{s(F)} \sum_{k=1}^L d_{s(F)k} A_{kl} > 0 \\
0 & \text{otherwise}
\end{cases}$$

where $\text{captive}_F l$ is a dummy variable taking the value one if location $l$ is captive for firm $F$ and zero otherwise. Whether a location is captive or not is computed using a symmetric feasibility matrix $A$ of dimension $L \times L$ with elements $A_{kl}$ taking the value one if a store at $k$ can reach consumers at $l$ and zero otherwise. Considering the firm’s location vector $d_{s(F)}$, if any of it’s stores reaches location $l$, then I declare this location as ‘covered by $F$’. If additionally, the location is ‘not covered’ by the rival, then I call it a captive location.

Under asymmetric information about the rival’s location, we need the expectations over the number of captive consumers.

**Proposition 1.** If a store $s(-F)$ locates at $l$, the probability that location $l'$ is covered by any $F$-store is given as follows

$$\phi^{ENF}_{l'} \equiv \Pr(\text{covered}_{l'} = 1) = 1 - (1 - \sum_k A_{kl'} \cdot B_{s(F)}^{a(F)} )^{NEF},$$

with $B_{s(F)}^{a(F)}$ being the beliefs about the location choice of a store with chain affiliation $F$.

The proof is provided on page 33.
In the following I will show that based on $\phi_{F,N,F}^{F,\cdot}$, it is straightforward to determine the number of consumers in competitive areas and captive regions, at the store level as well as at the firm level.

The expected number of ‘competitive consumers’ at the store level is just the expected number of consumers within the feasible market range that are ‘covered’ by the rival:

$$E[X_{\text{comp}_s}|s(F)l] = \sum_{l':d(l,l') \leq D_{\text{max}}} \phi_{F,N,F}^{-F,\cdot} \cdot X_{l'} \equiv g_3(BP)$$

However, since the prices are set at the firm level, when allowing for more than one store per firm, what matters is the total number of consumers in the competitive areas so that the expectations considering all stores are the following:

$$E[X_{\text{captive}_F}|s(F)l] = f_{2s(F)l}(d_{s(F)}, BP_{s(F)}, BP_{s(-F)}, A, X)$$

$$= \sum_{l'} \left[ \phi_{F,N,F}^{F,N,F}-1(1 - A(l, l')) + A(l, l') \right] \cdot (1 - \phi_{F,N,F}^{-F,\cdot}) \cdot X_{l'}$$

$$E[X_{\text{captive}_{\cdot,F}}|s(F)l] = f_{2s(-F)l}(d_{s(F)}, BP_{s(F)}, BP_{s(-F)}, A, X)$$

$$= \sum_{l'} \left[ (1 - \phi_{F,N,F}^{-F,\cdot})(1 - A(l, l')) \right] \cdot \phi_{F,N,F}^{-F,\cdot} \cdot X_{l'}$$

$$E[\Delta X_{\text{captive}_{\cdot,F}}|s(F)l] = f_{3s(F)l}(d_{s(F)}, BP_{s(F)}, BP_{s(-F)}, A, X) = f(\cdot) - g(\cdot)$$

$$= \sum_{l'} \left[ \phi_{F,N,F}^{F,N,F}-1(1 - A(l, l')) + A(l, l') - \phi_{F,N,F}^{-F,\cdot} \right] \cdot X_{l'}$$

Note that for the particular case with two stores, one of each chain ($N_F = N_{-F} = 1$), the structural variables are linear in the beliefs.
Proof of Preposition 1.

\[ E[X_{\text{captive}}] = E[f(d_F, d_{-F}, A, X)] \]

\[ = E[\sum_i X_i | \text{captive}] = \sum_i E[X_i | \text{captive}] \]

\[ = \sum_i E[I\{\sum_k d_{sk} A_{k} > 0\} \cdot (1 - I\{\sum_s d_{-F_s, A_{sl}} > 0\})|s(-F)] \cdot X_i \]

\[ = \sum_i P(\text{captive} \rightarrow \text{F}) \cdot X_i \]

\[ = \sum_i P(N_{\text{covered}} \geq 1 \cap N_{\text{covered}_{-F}} = 0) \cdot X_i \]

by Conditional Independence Assumption:

\[ = \sum_i P(N_{\text{covered}} \geq 1) \cdot [1 - P(N_{\text{covered}_{-F}} \geq 1)] \cdot X_i \]

(1.) for \(N_F = 1\):

\[ P(\text{covered}_{s(F)}) = E[I\{\sum_{k=1}^L d_{s(F)k} A_{kl} > 0\}] \]

\[ = P(d_{s(F)1} A_{1l} = 1 \cup d_{s(F)2} A_{2l} = 1 \cup ... \cup d_{s(F)L} A_{Ll} = 1) \]

by Mutually Exclusive Choices:

\[ = \sum_k P(d_{s(F)k} A_{kl} = 1) \]

\[ = \sum_k A_{kl} \cdot P(d_{s(F)k} = 1) \]

\[ = \sum_k A_{kl} \cdot E P_{s(F)} = \phi_s(BP_s(F)) \quad \text{result how firms form their expectations} \]

(2.) for \(N_F \geq 1\):

\[ P(N_{\text{covered}} \geq 1) = E[I\{\sum_{s(F)} d_{s(F)k} A_{kl} > 0\}] \]

\[ = E[1 - I\{\sum_{s(F)} d_{s(F)k} A_{kl} = 0\}] \]

since \(\text{covered}_{s(F)} \sim \text{Bernoulli}(\phi_s^F)\)

\[ \Rightarrow N_{\text{covered}} \sim \text{Binomial}(N_F, \phi_s^F) \]

\[ = 1 - P(N_{\text{covered}} = 0) \]

\[ = 1 - (1 - \phi_s^F)^{N_F} = 1 - (1 - \sum_k A_{kl} \cdot BP_k s(F))^{N_F} = \phi_s^{F,N_F}(BP_s(F)) \]
## Markets

Table A1. Discrete population distribution within markets

<table>
<thead>
<tr>
<th>UA/UC (ID)</th>
<th>Name (State)</th>
<th>L</th>
<th>Av. pop/loc*</th>
<th>Std.dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>847</td>
<td>Alamosa (CO)</td>
<td>12</td>
<td>0.3997</td>
<td>(0.2297)</td>
<td>0.0235</td>
<td>0.7275</td>
</tr>
<tr>
<td>955</td>
<td>Albany (OR)</td>
<td>41</td>
<td>1.2223</td>
<td>(1.5334)</td>
<td>0.0451</td>
<td>7.3970</td>
</tr>
<tr>
<td>3547</td>
<td>Astoria (OR)</td>
<td>28</td>
<td>0.2072</td>
<td>(0.2512)</td>
<td>0.0000</td>
<td>0.9583</td>
</tr>
<tr>
<td>5302</td>
<td>Barstow (CA)</td>
<td>37</td>
<td>0.6823</td>
<td>(1.1847)</td>
<td>0.0062</td>
<td>5.2514</td>
</tr>
<tr>
<td>11431</td>
<td>Bullhead City (AZ)</td>
<td>55</td>
<td>0.8104</td>
<td>(1.1928)</td>
<td>0.0245</td>
<td>6.5929</td>
</tr>
<tr>
<td>13267</td>
<td>Canon City (CO)</td>
<td>37</td>
<td>0.7244</td>
<td>(1.1034)</td>
<td>0.0038</td>
<td>5.000</td>
</tr>
<tr>
<td>14158</td>
<td>Carson City (NV)</td>
<td>60</td>
<td>0.9702</td>
<td>(1.7270)</td>
<td>0.0179</td>
<td>9.7462</td>
</tr>
<tr>
<td>14401</td>
<td>Casa Grande (AZ)</td>
<td>41</td>
<td>1.1651</td>
<td>(15624)</td>
<td>0.0211</td>
<td>6.7474</td>
</tr>
<tr>
<td>17020</td>
<td>The Dalles (OR)</td>
<td>30</td>
<td>0.7394</td>
<td>(1.4726)</td>
<td>0.0060</td>
<td>6.2118</td>
</tr>
<tr>
<td>20368</td>
<td>Cortez (CO)</td>
<td>19</td>
<td>0.3867</td>
<td>(0.5963)</td>
<td>0.0301</td>
<td>2.500</td>
</tr>
<tr>
<td>20557</td>
<td>Cottonwood (AZ)</td>
<td>35</td>
<td>0.6224</td>
<td>(0.8030)</td>
<td>0.0049</td>
<td>2.5872</td>
</tr>
<tr>
<td>20827</td>
<td>Craig (CO)</td>
<td>15</td>
<td>0.7369</td>
<td>(1.2891)</td>
<td>0.0112</td>
<td>4.7438</td>
</tr>
<tr>
<td>23230</td>
<td>Delta (CO)</td>
<td>15</td>
<td>0.4674</td>
<td>(0.6490)</td>
<td>0.0054</td>
<td>1.9948</td>
</tr>
<tr>
<td>26983</td>
<td>Ellensburg (WA)</td>
<td>16</td>
<td>1.3426</td>
<td>(2.3313)</td>
<td>0.0042</td>
<td>7.9152</td>
</tr>
<tr>
<td>30034</td>
<td>Florence (OR)</td>
<td>15</td>
<td>0.8056</td>
<td>(0.8199)</td>
<td>0.0162</td>
<td>2.1213</td>
</tr>
<tr>
<td>32491</td>
<td>Galveston (TX)</td>
<td>32</td>
<td>1.4669</td>
<td>(2.9546)</td>
<td>0.0000</td>
<td>10.660</td>
</tr>
<tr>
<td>33652</td>
<td>Glenwood Springs (CO)</td>
<td>30</td>
<td>0.1801</td>
<td>(0.3235)</td>
<td>0.0020</td>
<td>1.2994</td>
</tr>
<tr>
<td>36001</td>
<td>Ginnison (CO)</td>
<td>16</td>
<td>0.2738</td>
<td>(0.4421)</td>
<td>0.0013</td>
<td>1.2394</td>
</tr>
<tr>
<td>46747</td>
<td>Lake Havasu City (AZ)</td>
<td>51</td>
<td>0.8841</td>
<td>(0.9946)</td>
<td>0.0012</td>
<td>2.9632</td>
</tr>
<tr>
<td>59437</td>
<td>Morro Bay (CA)</td>
<td>35</td>
<td>0.7277</td>
<td>(1.2567)</td>
<td>0.0000</td>
<td>49571</td>
</tr>
<tr>
<td>62839</td>
<td>Newport (OR)</td>
<td>18</td>
<td>0.4912</td>
<td>(0.8911)</td>
<td>0.0000</td>
<td>3.1125</td>
</tr>
<tr>
<td>63514</td>
<td>North Bend (WA)</td>
<td>31</td>
<td>0.5433</td>
<td>(0.7387)</td>
<td>0.0066</td>
<td>3.7391</td>
</tr>
<tr>
<td>75367</td>
<td>Riverton (WY)</td>
<td>18</td>
<td>0.7144</td>
<td>(0.8853)</td>
<td>0.0050</td>
<td>3.0414</td>
</tr>
<tr>
<td>76339</td>
<td>Roseburg (OR)</td>
<td>53</td>
<td>0.7172</td>
<td>(1.0468)</td>
<td>0.0181</td>
<td>4.2613</td>
</tr>
<tr>
<td>77527</td>
<td>St. Helens (OR)</td>
<td>43</td>
<td>0.4857</td>
<td>(0.6949)</td>
<td>0.0307</td>
<td>3.4343</td>
</tr>
<tr>
<td>80686</td>
<td>Sequim (WA)</td>
<td>24</td>
<td>0.6345</td>
<td>(0.6736)</td>
<td>0.0000</td>
<td>2.7324</td>
</tr>
<tr>
<td>81415</td>
<td>Shelton (WA)</td>
<td>36</td>
<td>0.3390</td>
<td>(0.5840)</td>
<td>0.0000</td>
<td>2.7229</td>
</tr>
<tr>
<td>81901</td>
<td>Sierra Vista (AZ)</td>
<td>121</td>
<td>0.4313</td>
<td>(1.0137)</td>
<td>0.0000</td>
<td>5.2250</td>
</tr>
<tr>
<td>84682</td>
<td>Steamboat Springs (CO)</td>
<td>25</td>
<td>0.2402</td>
<td>(0.3711)</td>
<td>0.0209</td>
<td>1.5894</td>
</tr>
<tr>
<td>89920</td>
<td>Vail (CO)</td>
<td>11</td>
<td>0.0827</td>
<td>(0.0513)</td>
<td>0.0451</td>
<td>0.1624</td>
</tr>
<tr>
<td>97966</td>
<td>Yucca Valley (CA)</td>
<td>45</td>
<td>0.4428</td>
<td>(0.4960)</td>
<td>0.0047</td>
<td>1.7297</td>
</tr>
</tbody>
</table>

* population density in 1000
D Knitro problem specification and outcome

As specified in section 3, I formalize the equilibrium conditions of the game as nonlinear equality constraints, I leave the structural parameters unrestricted and define the choice probabilities as bounded on the interval \([0.00001, 1]\). The lower bound assumes that the selection probabilities are positive for all alternatives, which implies little loss of generality since empirically a probability of zero can not be distinguished from such a small probability (McFadden, 1974). The upper bound is a hypothetical constraint that would be active only if there were only one possible location in the market which is ruled out by the identification requirement 2 Note that this setting provides a closed and bounded set for the choice probabilities. As initial values for the beliefs I use a uniform distribution over all the locations within a market. For the structural parameters I use many different initial values, with the guess for the population coefficient and distribution distance based on the results from Datta and Sudhir (2013). For the implementation, I use numerical derivatives (first-difference approximation). I am aware of the efficiency improvement providing analytical derivatives but given the complexity of the constraints which makes the hand-coded Jacobian error-prone I could not yet code it correctly for the entire model and hence use numerical derivatives at the cost of a higher CPU time to avoid unnecessary bugs.

The output below provides the Knitro results for the baseline model specification (1), including the single iteration steps and the final statistics.
Notation: iteration number (Iter), cumulative number of function evaluations (fCount), value of the negative log-likelihood function (Objective), feasibility violation and the violation of the Karush-Kuhn-Tucker first order condition of the respective iterate (FeasError, OptError), distance between the new iterate and the previous iterate (Step), number of Projected Conjugate Gradient iterations required to compute the step (CGits).

**E Bootstrap distribution**

To determine the significance of the estimates I use the Bootstrap percentile method. Since the justification of this method rests on an approximately normal distribution of the parameters, let us exemplarily have a detailed look at the bootstrap distribution of the parameters of model specification (1), providing the non-parametric density functions for the structural parameters. Note that the distribution could be approximated by a normal distribution.

Hence, based on the bootstrap estimates that reported convergence (approx. 80%) I calculate for each model specification and each parameter the 90% and 95% confidence interval. Table B1-B4 indicate the quantiles of interest and the probability of a negative coefficient.
Table B1. Bootstrap distribution model (1)

<table>
<thead>
<tr>
<th></th>
<th>( \beta_1^{(1)} )</th>
<th>( \beta_2^{(1)} )</th>
<th>( \beta_3^{(1)} )</th>
<th>( \delta^{(1)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5% percentile</td>
<td>0.1280</td>
<td>-0.4312</td>
<td>-1.5947</td>
<td>-9.7423</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.1549</td>
<td>0.2677</td>
<td>-1.4286</td>
<td>-4.6864</td>
</tr>
<tr>
<td>95% percentile</td>
<td>1.0299</td>
<td>3.2601</td>
<td>0.1650</td>
<td>0.0258</td>
</tr>
<tr>
<td>97.5% percentile</td>
<td>1.2311</td>
<td>3.2779</td>
<td>-0.0758</td>
<td>0.0145</td>
</tr>
<tr>
<td>prob. ( \beta \leq 0 )</td>
<td>0.00</td>
<td>0.04</td>
<td>0.99</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Table B2. Bootstrap distribution model (2)

<table>
<thead>
<tr>
<th></th>
<th>( \beta_1^{(2)} )</th>
<th>( \beta_2^{(2)} )</th>
<th>( \beta_3^{(2)} )</th>
<th>( \beta_4^{(2)} )</th>
<th>( \delta^{(2)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5% percentile</td>
<td>-0.2392</td>
<td>0.7939</td>
<td>0.0883</td>
<td>-2.2971</td>
<td>-2.1172</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.1245</td>
<td>0.8266</td>
<td>0.1141</td>
<td>-2.2231</td>
<td>-2.1141</td>
</tr>
<tr>
<td>95% percentile</td>
<td>0.7417</td>
<td>2.9457</td>
<td>2.1987</td>
<td>-0.5243</td>
<td>-0.1710</td>
</tr>
<tr>
<td>97.5% percentile</td>
<td>0.8301</td>
<td>3.0037</td>
<td>2.3412</td>
<td>-0.4211</td>
<td>0.0599</td>
</tr>
<tr>
<td>prob. ( \beta \leq 0 )</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table B3. Bootstrap distribution model (3)

<table>
<thead>
<tr>
<th></th>
<th>( \beta_1^{(3)} )</th>
<th>( \beta_2^{(3)} )</th>
<th>( \beta_3^{(3)} )</th>
<th>( \delta^{(3)} )</th>
<th>( \gamma^{(3)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5% percentile</td>
<td>0.0791</td>
<td>-0.662</td>
<td>-1.4754</td>
<td>-2.5058</td>
<td>-0.4613</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.1921</td>
<td>0.0625</td>
<td>-1.3074</td>
<td>-2.1274</td>
<td>-0.3814</td>
</tr>
<tr>
<td>95% percentile</td>
<td>0.9927</td>
<td>3.0351</td>
<td>-0.1576</td>
<td>-0.2073</td>
<td>1.9518</td>
</tr>
<tr>
<td>97.5% percentile</td>
<td>1.0465</td>
<td>3.2201</td>
<td>-0.1225</td>
<td>0.3161</td>
<td>1.9956</td>
</tr>
<tr>
<td>prob. ( \beta \leq 0 )</td>
<td>0.00</td>
<td>0.04</td>
<td>0.99</td>
<td>0.95</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table B4. Bootstrap distribution model (4)

<table>
<thead>
<tr>
<th></th>
<th>( \beta_1^{(4)} )</th>
<th>( \beta_2^{(4)} )</th>
<th>( \beta_3^{(4)} )</th>
<th>( \beta_4^{(4)} )</th>
<th>( \delta^{(4)} )</th>
<th>( \gamma^{(4)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5% percentile</td>
<td>0.3313</td>
<td>0.9437</td>
<td>-0.0824</td>
<td>-1.0440</td>
<td>-1.5000</td>
<td>0.5681</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.2943</td>
<td>1.5397</td>
<td>-0.0495</td>
<td>-0.6354</td>
<td>-0.9063</td>
<td>1.1688</td>
</tr>
<tr>
<td>95% percentile</td>
<td>0.2426</td>
<td>0.5568</td>
<td>1.8205</td>
<td>-1.8995</td>
<td>-1.7471</td>
<td>0.2442</td>
</tr>
<tr>
<td>97.5% percentile</td>
<td>0.3616</td>
<td>2.2118</td>
<td>1.9206</td>
<td>-1.7789</td>
<td>-1.8498</td>
<td>0.2548</td>
</tr>
<tr>
<td>prob. ( \beta \leq 0 )</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
F Further robustness checks

Model (5) allows for a linear-quadratic shape of the distribution costs and model (6) controls for average consumer characteristics, like household size and age, within the trade area of the firm. Given the large number of iterations necessary to achieve convergence, I abstain from the computationally intensive bootstrap analysis and report only the equilibrium results which have to be interpreted with caution.

Table III. Further robustness checks

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{X}$</td>
<td>0.2794</td>
<td>0.2833</td>
<td>0.2556</td>
</tr>
<tr>
<td>$X_{\text{captive}}$</td>
<td>1.5977</td>
<td>1.5282</td>
<td>1.6168</td>
</tr>
<tr>
<td>$\Delta X_{\text{captive}}$</td>
<td>-0.2282</td>
<td>-0.3842</td>
<td>-0.5297</td>
</tr>
<tr>
<td>$\frac{\Delta X_{\text{captive}}}{X} \cdot (1 - \frac{\Delta X_{\text{captive}}}{X})$</td>
<td>-1.5302</td>
<td>-1.5635</td>
<td>-0.8151</td>
</tr>
<tr>
<td>$Z$</td>
<td>-1.5302</td>
<td>-1.5635</td>
<td>-0.8151</td>
</tr>
<tr>
<td>$Z^2$</td>
<td>0.3888</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB_distance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Av_Age</td>
<td>0.2039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Av_HHsize</td>
<td>0.6612</td>
<td></td>
<td></td>
</tr>
<tr>
<td># Iterations</td>
<td>25</td>
<td>328</td>
<td>2465</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-149.6538</td>
<td>-158.4787</td>
<td>-181.2045</td>
</tr>
</tbody>
</table>

Allowing for a more flexible form of the cost structure, including the squared distance, suggests a U-shaped pattern which confirms the results from my previous work. Note that the market power effect is robust to this functional variation of the cost structure while the price-competition effect becomes slightly stronger. This sensibility to the costs structure may be carefully be interpreted as distribution costs effecting partially also the marginal costs and hence the price setting.

The positive coefficients of age and household size suggest that traditional supermarkets are more likely to target ‘older’ people, which is to be understood in relative terms in the sense of families.

Further possible control variables would be the geographic income distribution or the social class of the households. However, given the data limitation at the disaggregated level of block groups I do not control for that and assume implicitly that the reservation price of all households for a standard food basket at a supermarket store is high enough.