Estimating Network Economics in Retail Chains: A Revealed Preference Approach

Paul B. Ellickson* Duke University

Stephanie Houghton† Duke University

Christopher Timmins Duke University and NBER

October 2, 2007

Abstract

Many industries are characterized by a small number of firms making many entry decisions over a large choice set. Nowhere is this more true than in the case of the discount retail sector, where Target, Wal-Mart, and Kmart compete in a national market. Traditional discrete choice models of firm entry are ill-suited to this high dimensional choice problem. We instead draw upon recent innovations in the application of maximum score estimators to models of revealed preference (Bajari and Fox (2006), Fox (2007)). Unlike previous work on this sector, our approach allows us to consider any number of potential rivals, any number of stores per location, the endogeneity of the distribution network, and unobserved (to the econometrician) location attributes that might cause firms to cluster their stores. Moreover, we show how recent innovations in set identification and inference can be used to separate the role of these unobservables from observed location attributes like population. We find that all firms (especially Target and Kmart) find it advantageous to cluster stores around distribution centers. Conditional upon that clustering, however, they find it costly to locate their stores in close proximity to one another (i.e., an "own business stealing" effect). All firms (especially Kmart) find it even more costly to locate in close proximity to a rival. Both of these strategic effects are understated if unobservable market attributes are ignored. Using counterfactual simulations, we explore the role of the distribution network in determining the level of retail competition faced by consumers in markets of varying sizes.

*The authors would like to thank Arie Beresteanu, Tom Holmes, Shakeeb Khan, and all the participants in the Duke University Applied Microeconometrics Lunch for their helpful comments. Any remaining errors and omissions are our own

†Corresponding author: Stephanie Houghton (shoughto@econ.duke.edu)

1 Introduction

A cursory examination of the size distribution of firms in almost any industry reveals a similar pattern: dominance by a few extremely large players. Large, multinational corporations sell everything from chemicals to coffee. Instead of observing the birth of many new firms as markets grow, we mainly witness the continued growth of a handful of large incumbents. Although game theory provides a rich framework for analyzing strategic interactions between small numbers of players, the complexity of the objective function that such firms maximize is truly daunting. Furthermore, many industries are dominated by as few as two or three firms, begging the question of how one can estimate rich systems of parameters with the observed behavior of so few players. As a result, empirical models of strategic interaction have tended to focus on interactions between very small firms in isolated geographic markets or narrowly defined product categories.[1] While analytically clean, this approach misses important features of modern competition: scale and scope. In this paper, we aim to develop a simple empirical method for tackling such large scale problems.

Our motivating example is the decision by a retail chain of where to locate its stores. While our empirical analysis will focus on the discount store industry, the same insights apply to almost every area of retail trade. Due to the increasing importance of information technology, distribution systems, and volume purchasing, virtually every retail industry is now dominated by a handful of powerful chains. While a few industries (like supermarkets) still feature some strong regional chains, most retail markets are effectively controlled by just two or three national players. Discount stores are a prominent example, with three national chains accounting for nearly three quarters of total sales. Each of these chains must solve a complex optimization problem: build a network of outlets that moves products to consumers as efficiently as possible, recognizing that their rivals are trying to do the same. While this optimization problem is naturally viewed as a discrete game, it has an extremely complex structure. Firms must choose the optimal design of a vast network of both stores and distribution centers in the face of fierce competition. Moreover, the combination of network and competitive effects implies that spillovers between stores could be either positive or negative. From an estimation standpoint, even if a tractable method for

[1] A notable exception is Jia's (2006) model of retail store location, which we will discuss in detail below.

solving an equilibrium were available, the solution would not likely be unique. Nonetheless, we argue that the parameters that govern such complex interactions can be estimated in a relatively simple manner using the logic of revealed preference and the tools of maximum score and set inference.

By relying on a revealed preference argument, we consider a large number of "local" perturbations to the observed network structure that involve swapping pairs of stores owned by rival firms between matched pairs of markets. The structure of these swaps serves two purposes. First, the deviations from the observed equilibrium must be payoff reducing, generating preference inequalities on which we base estimation. This revealed preference structure also eliminates the need to solve for an equilibrium, mitigating concerns over multiplicity of equilibria and greatly reducing the computational burden of the estimation procedure. Second, by considering only matched swaps of rival stores between pairs of markets, these inequalities mimic the structure of a difference-in-differences design, allowing us to eliminate a common, market level unobservable, which is key for obtaining unbiased estimates of either congestion effects or economies of density. Pairwise maximum score allows the researcher to estimate multinomial choice problems using arbitrarily chosen pairwise comparisons, provided the structure preserves a rank-ordering property.

Intuitively, our estimation strategy exploits the complexity of the decision space to offset the small number of market participants. Even though the number of players is small (in this case three), the number of alternative configurations (and pairwise deviations) is immense, allowing us to exploit asymptotics in the number of choices, rather than the number of agents. However, the small number of players does impose a cost when it comes to recovering and decomposing the fixed effects we have differenced away in the maximum score procedure. Employing another set of inequalities based this time on the first order conditions from the profit maximization problem, we recover ranges in which these fixed effects can lie, and then project them on additional covariates using set inference techniques developed in Beresteanu and Molinari (2006). Set identification follows from the limited number of inequalities available given the small number of players observed in this industry.

Our paper builds on and extends a large and growing IO literature on the estimation of discrete games that started with a series of seminal papers by Bresnahan and Reiss (1987, 1990, 1991). In three companion papers, the authors considered entry by single store firms into isolated geographic markets, quantifying the impact of price competition by identifying the threshold market sizes at which firms choose to enter a given market. Mazzeo (2002) and Seim (2006) extended this analysis to include product and spatial differentiation respectively, but maintained the focus on single unit firms in isolated markets. While Seim (2006) exploited the benefits of incomplete information to purify equilibria, Bajari et al. (2006) used this information structure to adapt the two step methods utilized in dynamic discrete games to the static settings considered here. The first structural papers to address the chain store network problem directly are Jia (2006) and Holmes (2006). Holmes (2006) examines the spatial structure of Wal-Mart's national network from the perspective of dynamic discrete choice, using Wal-Mart's sequential decisions over where to open additional stores to infer the importance of economies of density. However, he does not model (or control for) the choice of distribution network or reactions to (or by) Wal-Mart's rivals. The latter simplification allows for the possibility of mistakenly attributing density economies to the observation of firms clustering their stores to avoid rivals. Jia (2006) examines the network decision of both Kmart and Wal-Mart from a strategic perspective, using a full solution method that exploits a lattice representation of the two player problem. However, her elegant, lattice-based solution mechanism comes at a cost, restricting her analysis to only two national players and locations that contain only a single outlet per firm. She notes that, without the latter restriction, her finding of positive chain effects (i.e., density economies) would be less likely to result. In our analysis, we consider the full set of potential locations, allow for any number of competing firms, and place no restrictions on the number of stores per location while controlling for proximity to endogenously placed distribution centers. Aside from finding a strong incentive to cluster around distribution centers, we find no evidence that firms benefit from locating their stores in close proximity to one another – rather, we find strong evidence of own business-stealing.

Our analysis also draws extensively on recent extensions and applications of maximum score. While our pairwise maximum score procedure is based on Fox (2007), the first stage of our estimation framework is closest in structure to Bajari and Fox (2006), who consider competitive bidding for packages of mobile phone "spectrum" licenses. The double differencing procedure used here is also similar to one of the identification strategies used by Pakes, Porter, Ho and Ishii (2006) to eliminate the structural component of the composite error utilized in their moment inequality based approach.

2 Model

We focus on the discount store industry. While this segment of retail was once quite fragmented, in the past few decades it has come to be dominated by three main players: Wal-Mart, Kmart, and Target. While each firm is essentially national in scope, there are some fairly obvious distinctions in the types of markets they choose to serve. Some of this is driven by the types of consumers they target. Consistent with its rural beginnings and choice of merchandise, Wal-Mart clearly favors rural locations and smaller cities, avoiding the major cities almost completely (not always by choice). Target, on the other hand, clearly prefers urban locations, consistent with its more "up-market" focus. Finally, Kmart is much less focused, having stores both in major cities and rural towns (which may explain its lackluster performance). In addition to proximity to consumers, location choice is also constrained by logistics - stores have to be stocked with merchandise from a regional distribution center, and may further benefit from being close to one another (economies of density). In other words, all three firms are designing an optimal network of stores, balancing economies of scale and distribution against the idiosyncratic preferences of individual consumers.

In retail industries, location choice essentially dictates the types of consumers you are going to serve, as individuals tend to sort themselves into reasonably homogeneous neighborhoods. While such sorting could (and probably does) occur on a very local level, from the researcher's perspective, there is a trade-off between choosing a fine grid and allowing for meaningful correlated unobservables. With this trade-off in mind, we chose Core Based Statistical Areas (CBSAs) as our basic building block. CBSA refers collectively to metropolitan statistical areas and smaller micropolitan statistical areas, which we will call a market or location. Firms then choose which locations to enter and how many stores to build in each. While this is somewhat restrictive in that it ignores more nuanced aspects of spatial differentiation, it will allow us to account for correlated, market level, unobservables, which are key to correctly identifying network and congestion effects. In this section, we outline the estimation algorithm used to recover a firm's payoff function, which describes the determinants of its entry decisions.

Consistent with the determinants of behavior described above, we model the *per store* payoff to firm $f = \{T, K, W\}$ of each store in market j as:

$$\begin{aligned}
\pi_j^T &= \beta^{T,Own} \ln\left(N_j^T + 1\right) + \beta^{T,Other} \ln\left(N_j^W + N_j^K + 1\right) + \beta^{T,DC} DC_j^T + \beta^{T,X} X_j + \theta_j \\
\pi_j^W &= \beta^{W,Own} \ln\left(N_j^W + 1\right) + \beta^{W,Other} \ln\left(N_j^T + N_j^K + 1\right) + \beta^{W,DC} DC_j^W + \beta^{W,X} X_j + \theta_j \\
\pi_j^K &= \beta^{K,Own} \ln\left(N_j^K + 1\right) + \beta^{K,Other} \ln\left(N_j^W + N_j^T + 1\right) + \beta^{K,DC} DC_j^K + \theta_j
\end{aligned}$$

where N_j^f is the total number of stores firm f operates in market j, DC_j^f is the distance from market j to firm f's nearest distribution center, X_j is a vector of exogenous attributes of market j (e.g. income, population, household size, regional dummies), β is a vector of parameters $\left\{\beta^{f,Own}, \beta^{f,Other}, \beta^{f,DC}, \beta^{f,X}\right\}$ with $\beta^{f,K} = 0$, and

$$\theta_j = \gamma^X X_j + \xi_j$$

where ξ_j is an unobserved (to the econometrician) attribute of market j (assumed to be scalar and common across firms). Note that $\gamma^X X_j$ describes the way in which the X_j variables affect Kmart's profits. The impact of X_j on Wal-Mart's profits, on the other hand, is represented by $(\gamma^X + \beta^{W,X})X_j$, while the effect of X_j on Target's profits is found from a comparable expression. $\beta^{f,Own} > 0$ would indicate that the per-store profit for firm f is increasing in the number of stores it has a particular location (i.e., economies of density). $\beta^{f,Other} < 0$ indicates that firm f's profits per store are falling in the number of its competitors" stores in the same location.

Since per store profits are the same at every store in a given market, total firm level profits are given by

$$\Pi^f = \sum_j N_j^f \times \pi_j^f,$$

which is, ultimately, the quantity that firms are maximizing. An important feature of our modeling framework is the inclusion of the unobserved market attribute, ξ_j, which serves as our structural error. From a practical standpoint, it is unlikely that our vector of observable market attributes (X_j) will capture everything about a market that is important in driving profitability. If we ignore these unobservables (i.e., forcing the idiosyncratic store-specific error, which we have not yet explicitly included, to control for them), we will likely arrive at biased estimates of a number of parameters – particularly those associated with N_j^f.[2] Of course, explicitly including ξ_j as part of the error structure creates potential endogeneity problems since N_j^f is a local attribute that is determined in equilibrium. It is, therefore, a direct function of ξ_j as well as an indirect function of $\xi_k \, \forall \, k \neq j$. In solving this problem, we rely on three identification assumptions – in particular, (1) endogenous attributes of each market (i.e., the number of stores of one's own firm, the number of stores of other firms, and the distance to the closest distribution center) are firm specific (i.e., distance to closest Wal-Mart distribution center is relevant only for Wal-Mart stores), (2) the unobserved attribute ξ_j is common across firms, and (3) after controlling for the common unobservable ξ_j, any remaining idiosyncratic error is uncorrelated with N_j^f, N_j^{-f}, and DC_j^f. The first two assumptions are common in both the sorting and discrete choice (demand) literature. The final assumption is sensible if the idiosyncratic error represents measurement error or a realization of information learned after entry has taken place.[3] For this reason, it is important that we include amongst our observable attributes of each market variables that might be differently valued by different firms. One obvious candidate is simply geography. The locations of corporate headquarters can explain much of Wal-Mart's prominence in the South and the ubiquitousness of Target in the upper Midwest. We therefore include regional dummy variables in the X_j vector.

We model the equilibrium locations of all Targets, Wal-Marts, and Kmarts as they appeared in 2006.[4] The locations of each firm's distribution centers are important determinants of these siting decisions, but are certainly not exogenously determined. There is a potentially complicated model of firms' decisions about where to place these distribution centers that we do not attempt to model. Instead, we deal econometrically with the fact that DC_j^f is endogenous, allowing us to recover unbiased estimates of firm preferences.[5]

Our estimator is based on a revealed preference approach that uses pairwise comparisons

[2] See Bayer and Timmins (2007) for relevant Monte Carlo evidence.

[3] Pakes et al. (2006) employ a similar error decomposition.

[4] Following the standard practice in the static entry literature, we treat the entire networks of each firm as being determined simultaneously in a one-shot game. In contrast, Holmes (2006) models the dynamics of store diffusions, albeit from the perspective of only a single firm.

[5] The upside of this approach is that it allows us to recover the role played by distribution centers in the firm entry decision in a very simple framework. We find that proximity to distribution centers matters a lot. The downside of this approach is that it does not allow us to predict new spatial distributions of distribution centers under a counterfactual scenario. We return to this issue below.

between the observed location decisions made by firms and specific "single-store" deviations. The assumption is that a single-store deviation (i.e., taking a single store and moving it to a new location, holding the location decisions of other firms fixed) is a deviation from the observed Nash equilibrium and is, therefore, payoff reducing for the firm. Recall that the per store profit of a given firm (say Target) in market j is given by

$$\pi_j^T = \beta^{T,Own} \ln\left(N_j^T + 1\right) + \beta^{T,Other} \ln\left(N_j^W + N_j^K + 1\right) + \beta^{T,DC} DC_j^T + \beta^{T,X} X_j + \theta_j$$

Since the per store profit is the same for every store in a given market, Target's total profit in market j can be written

$$N_j^T \cdot \left(V_j^T\left(N_j^T\right) + \theta_j\right)$$

where

$$V_j^T\left(N_j^T\right) = \beta^{T,Own} \ln\left(N_j^T + 1\right) + \beta^{T,Other} \ln\left(N_j^W + N_j^K + 1\right) + \beta^{T,DC} DC_j^T + \beta^{T,X} X_j$$

Note that N_j^W and N_j^K are always held at their observed values in $V_j^T\left(N_j^T\right)$, as we only consider unilateral deviations. The logic of the estimator is to consider "swaps" of a single store from one market to another. For example, consider moving a single Target store from market a (e.g. Minneapolis), which currently contains N_a^T Target stores, to market b (e.g. Chicago), which currently contains N_b^T Target stores. Since the observed configuration $\left(N_a^T, N_b^T\right)$ is part of an equilibrium, Target's total profits must be higher under the observed configuration than under the proposed counterfactual configuration $\left(N_a^T - 1, N_b^T + 1\right)$. Note that since all "spillovers" (i.e. congestion effects or economies of density) are assumed to occur within, but not across, markets (i.e. profits are additively separable across markets) the change in total firm profits only depends on the incremental changes associated with the two markets exchanging stores, yielding the following relatively simple inequality

$$\begin{aligned} & N_a^T \cdot V_a^T\left(N_a^T\right) + N_a^T \cdot \theta_a + N_b^T \cdot V_b^T\left(N_b^T\right) + N_b^T \cdot \theta_b \\ > \; & \left(N_a^T - 1\right) \cdot V_a^T\left(N_a^T - 1\right) + \left(N_a^T - 1\right) \cdot \theta_a + \left(N_b^T + 1\right) \cdot V_b^T\left(N_b^T + 1\right) + \left(N_b^T + 1\right) \cdot \theta_b \end{aligned}$$

Simplifying this expression yields

$$\widetilde{\Delta V}_a^T\left(N_a^T, N_a^T - 1\right) + \widetilde{\Delta V}_b^T\left(N_b^T, N_b^T + 1\right) + \left(\theta_a - \theta_b\right) > 0 \tag{1}$$

where the $\widetilde{\Delta V}_j^f(\cdot,\cdot)$ notation represents the decrease (or increase) in profits associated with removing a single store from (or adding a single store to) a particular market. A difficulty in using this inequality to recover the structural parameters of Target's payoff function is the difference in fixed effects, $(\theta_a - \theta_b)$. In addition to the valuation of the exogenous features of the market which are common to all players (represented here as Kmart's preferences over population, for example), these fixed effects capture the parameters ξ_a and ξ_b, which are common knowledge of the firms, but unobserved to the econometrician. To obtain unbiased estimates of the profit parameters, we must eliminate these common unobservables from the preference inequality. Therefore, we consider another hypothetical store movement that will allow us to difference these fixed effects away. In particular, consider moving one Kmart store from b to a. This yields a similar inequality to the one above, but with an offsetting difference in the fixed effect term:

$$\widetilde{\Delta V}_a^K\left(N_a^K, N_a^K+1\right) + \widetilde{\Delta V}_b^K\left(N_b^K, N_b^K-1\right) + (\theta_b - \theta_a) > 0 \tag{2}$$

Subtracting equation (2) from equation (1) yields another inequality that is free from this problematic term:

$$\widetilde{\Delta V}_a^T\left(N_a^T, N_a^T-1\right) + \widetilde{\Delta V}_b^T\left(N_b^T, N_b^T+1\right) - \widetilde{\Delta V}_a^K\left(N_a^K, N_a^K+1\right) - \widetilde{\Delta V}_b^K\left(N_b^K, N_b^K-1\right) > 0 \tag{3}$$

Using the same logic, we construct many "offsetting" swaps of stores between the three retailers. By considering many such "minor perturbations" to the observed spatial network of stores, we are able to construct a pairwise maximum score objective function with which to estimate the β parameters (the remaining (common) parameters (γ) are estimated in a second step). Manski (1985) shows that point identification of the β parameters requires a special regressor with continuous support. This covariate essentially serves to break ties between choices and therefore must differ across choices by a sufficient margin. Distance to the nearest distribution center, which will always vary across locations and players is a natural candidate since we only consider switching stores from one market to another. Owing primarily to the non-smoothness of the objective function, the pairwise maximum score procedure does not yield analytic standard errors (or asymptotic normality). Therefore, standard errors will be obtained using the bootstrap with sub-sampling.

Finally, it is important to note what has happened with the endogenously determined local attributes ($N_j^{OTHER,f}, N_j^{OWN,f}, DC_j^f$) in this system. We would expect each of these to be correlated with ξ_j (i.e., places that are desirable owing to unobserved factors are likely to have more stores in them (both OWN and OTHER). Locations surrounded by locations with desirable unobservables are, similarly, more likely to be close to a distribution center (assuming distribution centers are placed with the goal of servicing a large number of attractive markets). Key to our estimation strategy, ξ_j no longer appears in our objective function when it comes time to estimate the parameter vector β. However, the other elements of θ_j (apart from ξ_j) are important in giving a meaningful economic interpretation to several of the parameters recovered in the first stage. The common, market specific terms were differenced away. This is an issue in any procedure that involves differencing out common unobservables. An important contribution of our framework lies in the ability to recover these additional parameters, rather than simply differencing them away.

2.1 Decomposing the Market Level Fixed Effects

Crucial to our understanding of firm entry behavior is the recovery of the parameters γ^X, which are included in θ_j, the fixed effect terms that were differenced out of equation (3). Recall from above that these parameters describe how the demographic variables included in X_j affect Kmart's profits. The impact of X_j on Wal-Mart's and Target's profits are represented by $(\gamma^X + \beta^{W,X})X_j$ and $(\gamma^X + \beta^{T,X})X_j$, respectively. Although we have already obtained estimates of $\beta^{W,X}$ and $\beta^{T,X}$ in the first stage, we clearly need to know Kmart's values (captured by $\gamma^X X$) in order to determine the overall value that any firm places on an attribute like per-capita income or population density. However, because we were concerned that ξ_j would be correlated with all the N's, we included it in the fixed effect (θ_j) and differenced it out in the first stage. We will now use our first stage estimates, along with an assumption on firm behavior, to recover (set valued) estimates of the fixed effects. We can then project these interval estimates onto X using Beresteanu and Molinari's (2006) techniques for set valued random variables.

With only three firms and without making explicit distributional assumptions, we are unable to recover point estimates of θ_j. There is simply not enough information in data describing the store siting decisions of just three firms to identify the precise values of this many parameters. However, we can recover ranges in which those parameters must lie and then use these intervals to set-identify γ^X. In order to do so, we make use of an additional assumption about firm behavior similar to that employed by Pakes, Porter, Ho, and Ishii (2007). In particular, we begin with each firm's marginal profitability of stores in each market j:

$$\begin{aligned}
\frac{\partial \Pi_j^T}{\partial N_j^T} &= \beta^{T,Own} \ln\left(N_j^T + 1\right) + \beta^{T,Own} \frac{N_j^T}{N_j^T + 1} + \beta^{T,Other} \ln\left(N_j^W + N_j^K + 1\right) + \beta^{T,DC} DC_j^T + \beta^{T,X} X_j + \theta_j \\
\frac{\partial \Pi_j^W}{\partial N_j^W} &= \beta^{W,Own} \ln\left(N_j^W + 1\right) + \beta^{W,Own} \frac{N_j^W}{N_j^W + 1} + \beta^{W,Other} \ln\left(N_j^T + N_j^K + 1\right) + \beta^{W,DC} DC_j^W + \beta^{W,X} X_j + \theta_j \\
\frac{\partial \Pi_j^K}{\partial N_j^K} &= \beta^{K,Own} \ln\left(N_j^K + 1\right) + \beta^{T,Own} \frac{N_j^K}{N_j^K + 1} + \beta^{K,Other} \ln\left(N_j^T + N_j^W + 1\right) + \beta^{K,DC} DC_j^K + \theta_j
\end{aligned}$$

Note that we will not attempt to recover the determinants of the number of total stores that each firm builds. This will be affected by capital constraints and cash reserves, long-run business plans, access to foreign suppliers, and so forth. Instead, we assume only that each firm allocated whatever stores it did build in an optimal fashion – that is, they added stores to each market until the marginal profitability of additional stores was equalized across markets. In particular, for each firm f, there is a value c^f such that observed marginal profits in each market with at least one store must be greater than c^f, while the marginal profits from one additional store must be less than c^f:

$$\beta^{f,Own} \ln\left(N_j^f + 1\right) + \beta^{f,Own} \frac{N_j^f}{N_j^f + 1} + \beta^{f,Other} \ln\left(N_j^{-f} + 1\right) + \beta^{f,DC} DC_j^f + \beta^{f,X} X_j + \theta_j \geq c^f \quad (4)$$

$$\beta^{f,Own} \ln\left(N_j^f + 2\right) + \beta^{f,Own} \frac{N_j^f + 1}{N_j^f + 2} + \beta^{f,Other} \ln\left(N_j^{-f} + 1\right) + \beta^{f,DC} DC_j^f + \beta^{f,X} X_j + \theta_j < c^f \quad (5)$$

where $\beta^{f,X} = 0$ if $f = K$. Note that in markets where the firm has zero stores, it is at a corner solution and only Equation (5) must hold. Define the following components of firm f's marginal profits in market j

$$\begin{aligned}
\psi_{1,j}^f &= \beta^{f,Own} \ln\left(N_j^f + 1\right) + \beta^{f,Own} \frac{N_j^f}{N_j^f + 1} + \beta^{f,Other} \ln\left(N_j^{-f} + 1\right) + \beta^{f,DC} DC_j^f + \beta^{f,X} X_j \\
\psi_{2,j}^f &= \beta^{f,Own} \ln\left(N_j^f + 2\right) + \beta^{f,Own} \frac{N_j^f + 1}{N_j^f + 2} + \beta^{f,Other} \ln\left(N_j^{-f} + 1\right) + \beta^{f,DC} DC_j^f + \beta^{f,X} X_j
\end{aligned}$$

Since we have estimated all the β parameters in the first stage (and observe everything else), we can treat this component of marginal profits as known.[6] Because payoffs are only identified up to an additive scale parameter in our discrete choice framework, we have one free normalization. Setting the value of θ_j equal to zero for some base location, we have[7]

$$\psi^f_{2,base} < c^f \leq \psi^f_{1,base}$$

With these bounds on c^f for each firm, we can now use equations (4 & 5) to bound θ_j for each firm:

$$\begin{aligned} \psi^f_{1,base} - \psi^f_{2,j} < \theta_j \leq \psi^f_{2,base} - \psi^f_{1,j}, \qquad & \text{if } N^f_j > 0 \\ \theta_j \leq \psi^f_{2,base} - \psi^f_{1,j}, \qquad & \text{if } N^f_j = 0 \end{aligned}$$

Assuming that the θ_j's are the same for all firms, the intersection of these bounds describes the set of admissible values that θ_j can take for each market j.[8] Note that if we had point estimates of the θ_j's, we could simply project them onto X using a linear regression. However, because the θ_j's are set identified, our dependent variable now comes in interval form: $\theta_j \in [\theta_{Lj}, \theta_{Uj}]$. Therefore, we rely on the methods developed by Beresteanu and Molinari (2006), which use a transformed Minkowski average of the data to recover bounds and confidence sets for γ^X. Specifically, we estimate the parameters for the best linear predictor of θ in the interval $[\theta_L, \theta_U]$ conditional on X. Suppose X were to consist of two variables, X_1 and X_2. Then the population set-valued best linear predictor is defined as

$$\Gamma = \left\{ \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} : \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} E(X_1^2) & E(X_1 X_2) \\ E(X_2 X_1) & E(X_2^2) \end{bmatrix}^{-1} \begin{bmatrix} E(X_1\theta) \\ E(X_2\theta) \end{bmatrix}, \theta \in [\theta_L, \theta_U] \right\}$$

The estimate $\widehat{\Gamma}$ is obtained by using the sample analogs of the above expectations. The third term in brackets can simply be estimated by $\hat{\Sigma}^{-1} = \left[\frac{1}{J} X'X\right]^{-1}$. The last term in

[6] We will correct for the fact that these are estimates in the second stage standard errors.

[7] In practice, the base location used is the Albany, Georgia Metropolitan Area. Kmart, Target, and Wal-Mart each have one store there. The area's population is around the 75th percentile and the population density is close to the median value of all 912 CBSAs in our sample.

[8] Note that we drop from the θ_j decomposition exercise around fifty locations where the sets of θ_j's are disjoint (i.e. where no value of θ_j can satisfy all three firms inequalities). This is an indication of model mis-specification. Because this only happens with a small fraction of our overall locations, we do not make any correction for possible sample selection in the decomposition of θ_j.

brackets can be written as $\mathbb{E}(G)$, where G is a set-valued random variable reflecting all possible values of $X'\theta$, given that θ is bounded by θ_U and θ_U:

$$G = \left\{ \begin{bmatrix} X_1\theta \\ X_2\theta \end{bmatrix} : \theta \in [\theta_L, \theta_U] \right\}$$

The sample analog to this expectation is given by the Minkowski average $\bar{G}_J = \frac{1}{J}\bigoplus_{j=1}^{J} G_j$. That is, for each observation j, G_j is a line segment with endpoints at $(X_{j1}\theta_{jL}, X_{j2}\theta_{jL})$ and $(X_{j1}\theta_{jU}, X_{j2}\theta_{jU})$. The Minkowski sum of G_j, from $j = 1..J$, adds all these segments to form a many-sided polygon. This polygon is then transformed by $(X'X)$ to obtain an estimate of the set $\Gamma = [\gamma_1, \gamma_2]$. Therefore,

$$\widehat{\Gamma} = \hat{\Sigma}^{-1}\bar{G}_J \tag{6}$$

In general, if there are K covariates in X, this estimated set $\widehat{\Gamma}$ will be a K-dimensional polytope. In our application, the X variables include a constant term and various location-specific characteristics, such as the log of population and median income, the average household size, the percent of the area's population living in urban areas, the population density, and regional dummies. It is difficult (if not impossible) to Minkowski sum such a high-dimensional set-valued random variable. However, we can estimate a subset of the parameters of the best linear predictor using the same logic as the Frisch-Waugh-Lovell Theorem for partial regression in point-identified models. For example, to estimate the set of possible coefficients on $log(median\ income)$ and $pctUrban$, first obtain the residuals from a linear regression of $log(median\ income)$ on all other X variables besides $pctUrban$ and the residuals from a linear regression of $pctUrban$ on all other X variables besides $log(median\ income)$. Denote these $\tilde{X}_1$ and $\tilde{X}_2$ and use in place of X_1 and X_2 above.

By the same partial regression logic, we can also estimate one-dimensional projections of the identification region. For example, the identification region of the best linear predictor coefficient for a single γ_k can be estimated by the interval [9]

$$\hat{\gamma}_k = \frac{1}{\sum_{j=1}^{J}\tilde{X}_{j,k}^2}\left[\sum_{j=1}^{J}\min\left\{\tilde{X}_{j,k}\theta_{Lj}, \tilde{X}_{j,k}\theta_{Uj}\right\}, \sum_{j=1}^{J}\max\left\{\tilde{X}_{j,k}\theta_{Lj}, \tilde{X}_{j,k}\theta_{Uj}\right\}\right] \tag{7}$$

where $\tilde{X}_{j,k}$ is the residual from the regression of $X_{j,k}$ on the other $X_{j,-k}$, including the constant term.

[9] This simple result comes from Corollary 4.5 of Beresteanu and Molinari (2006).

For both the one-dimensional and two-dimensional projections, confidence sets can be formed by bootstrapping and computing the Hausdorff distance between the estimated set for the original sample and the estimated set for each bootstrapped sample, $H(\Gamma^b, \hat{\Gamma})$. Under standard regularity conditions, Beresteanu and Molinari (2006) show that $r = \sqrt{n}H(\Gamma^b, \hat{\Gamma})$ is asymptotically normally distributed. Therefore, we can use the 95th percentile of the empirical distribution of r to construct bounds on the collection of all sets that, when specified as the null hypothesis for the true value of the population identification region for Γ, cannot be rejected at a 95% confidence level. This is analogous to forming a 95% confidence interval for a point estimate.

3 Data

The data for the discount store industry are taken primarily from Trade Dimension's Retail Tenant Database for 2006. This proprietary dataset contains all 6,150 Wal-Mart, Kmart, and Target stores in operation in the continental United States as of August 2006. These include both pure discount stores that carry general merchandise and newer supercenter formats that also carry full grocery lines.[10] Because of the additional difficulty in modeling a firm's choice of store format, we do not distinguish between these two types of stores in our current application.

Stores are assigned to markets based on the 2005 Census definitions for Core Based Statistical Areas (CBSAs), a term that refers collectively to metropolitan statistical areas and smaller micropolitan statistical areas. These statistical areas contain from one to twenty-eight counties (on average, 1.9 counties) and include both a core urban area and any adjacent counties that are closely linked economically and socially. Metro areas are those with a core urban area with a population of at least 50,000, and micro areas are those with a core urban area with a population from 10,000 to 50,000. These areas account for roughly 93% of the U.S. population and over 90% of the discount stores in our dataset. We exclude from our analysis any stores that are located in the isolated spaces outside these statistical areas.[11] We believe that using CBSA's as our market definition is more appropriate than

[10] We exclude Sam's Club stores, owned by Wal-Mart, because these warehouse clubs do not compete directly with Kmart and Target stores.

[11] We exclude 4 Target stores, 67 Kmart stores, and 431 Wal-Mart stores.

using counties since they typically define a more natural "shopping area" by grouping small adjacent counties together. The same criticism that applies to using large counties (e.g., Los Angeles County) to measure markets also applies here – all stores in the largest metro areas probably do not compete equally with one another.

We have collected demographic and economic data on each CBSA from the U.S. Census Bureau. The most recent population estimates at the CBSA level are available from the Census Population Division for July 2005. Household median income and average household size are aggregated from the county level from the 2000 Census. We also include the percentage of the metro or micro area population that is located in an urban area, collected from the website of the Missouri Census Data Center. The most recent data on retail sales are available at the county level from the 2002 Economic Census, which we then aggregate to the CBSA level.

In addition to these local characteristics, an important determinant of a firm's store location choice is the distance to suppliers, so a unique feature of our data is the availability of detailed information about distribution centers. The Trade Dimension's database contains the locations of each of the 113 distribution centers that serve the stores in our analysis. Wal-Mart is vertically-integrated, owning all 70 distribution centers that supply its stores. Kmart and Target operate most of the distribution centers that supply their stores, 13 and 24 centers, respectively. However, a small number of their stores rely on third-party distributors. We include the locations of one Merchants Distributors center and five SuperValu distribution centers that serve Kmart and Target stores, respectively. Using the Haversine formula, we calculate the distance from the population-weighted centroid of each CBSA to the CBSA of the closest distribution center for each firm.[12]

Summary statistics for these markets are provided in Tables 1 and 2. Of the 912 markets that contain at least one discount store, 358 are metro areas and 554 are micro areas. Population estimates for 2005 range from 11,638 (in Pecos, TX) to over 18.7 million (in New York-Northern New Jersey-Long Island). The number of discount stores per market ranges from 1 to 161, but this distribution is highly skewed. The median number of stores per market is 2, and 75% of markets have 5 or fewer stores.

[12] For the 8 Wal-Mart distribution centers located outside a CBSA, we assign them to the closest nearby CBSA. Since these distribution centers are often in an adjacent county, this amounts to a displacement of on average 35 miles.

Wal-Mart has a presence in 98% of the markets, compared to 57% and 41% of the markets for Kmart and Target, respectively. In part, this is expected because there are twice as many Wal-Mart stores included in the data (2913) as there are Kmart and Target stores (1295 and 1439, respectively). However, it also appears that these firms follow quite different strategies when locating stores. This can especially be seen in Table 2, which summarizes the location characteristics for each store. Over 95% of Target's stores are located in metro areas, compared to only 68% of Wal-Mart's stores and 78% of Kmart's stores.

4 Results

Our main empirical results are presented in Tables 3 and 4 for several alternative specifications of the model. Table 3 reports the parameters for the firm-specific components of per-store profits, which were estimated in the first stage of our two-step procedure. The results from the second stage procedure are presented in Table 4. Focusing first on the results from the first stage, several features are worth noting.

First, with respect to density economies, the negative coefficients on distance to distribution centers indicate that there are clear scale economies induced by the distribution network. All three firms exhibit a distaste for being far from their distribution centers. The effects are similar in magnitude for Kmart and Target, while the estimates for Wal-Mart's distaste is generally smaller than -1.[13] This is consistent with other evidence of Wal-Mart's highly-efficient logistics system. Second, in all of our specifications, we find clear evidence of the congestive effect of competition both among rival firms ($\beta^f_{OTHER} < 0$) and among stores owned by the same firm due to business stealing ($\beta^f_{OWN} < 0$). While the competition effect is to be expected, the dominance of the business stealing effect over local density economies unrelated to distribution is somewhat surprising given the results reported by both Jia (2006) and Holmes (2006). At least two factors are likely to be in play here. First, we consider larger markets than either Jia (2006) or Holmes (2006), and these markets contain many more stores. Therefore, the chain/business stealing effect we are finding is occurring within markets, as opposed to across markets. Cannibalization within markets is

[13]In all specifications, we normalized the coefficient on distance to distribution center in Kmart's profit function (β^K_{DC}) to -1.

likely to be much stronger than cannibalization across markets. Second, we have included a market level fixed effect that accounts for unobserved market features that are likely to be correlated with the clustering of stores.[14] Column V, which contains the results from a specification in which the market level fixed effects are dropped[15], illustrates the consequences of ignoring this heterogeneity. As expected, both the competition and business stealing results are biased downward, as own and rival stores are now proxying for unobserved spikes in the relative desirability of particular markets. Nonetheless, we still do not find agglomeration/chain effects.[16] The importance of the unobservable characteristic can also be seen by comparing the value of the score function under the different specifications. When the unobservable coefficient excluded in Column V, the estimator correctly matches approximately 67% of the sampled inequalities. This score jumps to 91% when the unobservable is included in Columns I-IV. We also note that the models which include more of the location-specific covariates have slightly higher scores, as expected.

These own and rival effects also shed light on the relative strengths and weaknesses of each firm. By comparing the magnitude of these coefficients, it appears that Kmart suffers the most from the introduction of a rival firm in the same market, while the impact of rivals on store profits is lowest for Wal-Mart. This indicates that Wal-Mart is the most able to insulate itself from competition. In fact, the impact of an additional store on per-store profit is almost the same whether the entrant is a competing chain or another Wal-Mart, suggesting that Wal-Mart is fairly saturated in its existing markets. The estimated business stealing effect for Kmart is small, suggesting that if Kmart were to open another store in an existing market, it would attract new customers rather than drawing them away from the existing store. Again, this is likely due to relative saturation, there are more places for Kmart to open stores. However, Kmart would pay a steep penalty were it to locate in the same market as Target or Wal-Mart.[17] The important role played by the unobserved

[14] Given the cross-market spillovers in Jia (2006) and Holmes (2006), the analog would be the inclusion of a nest/cluster fixed effect that is constant over the region in which the spillovers occur.

[15] By eliminating this unobservable, we no longer have to combine the matched inequalities for firm pairs. Instead, we use each of these inequalities (Equations (1) and (2)) separately and estimate the set of parameters that maximize the score function.

[16] We are in the process of repeating the analysis at the county level and allowing for cross-market spillovers. By including cluster-specific fixed effects, we should be able to control for correlated unobservables and isolate the pure chain/business stealing effect.

[17] Quick calculations based on Column III: The loss of a rival (changing from 2 stores total to 1 store)

location attributes is evident in the own- and rival-store effects in column V (where those unobservables are ignored). The effect of attractive unobservables is to make all firms more likely to place stores in particular locations. When these unobserved attributes are ignored, this effect is instead captured by the own- and rival-store effects, biasing those effects towards zero.

Note that this model allows firms to have different preferences over observable characteristics of the market. Target, which is located primarily in metropolitan areas, has the strongest preference for population relative to both Kmart and Wal-Mart, while Wal-Mart has the weakest. This is consistent with conventional wisdom regarding each firm's preferred demographic. (Right now, the coefficients suggest that Target likes median income less than Kmart, while Wal-Mart likes it more.)

The second stage estimates are reported in Table 4. The bounds and 95% confidence sets for the coefficient for population are positive in all our specifications, indicating, as expected, that larger markets are more profitable. We also note that, while the first-stage estimates indicate Wal-Mart's preference for population is lower than Kmart's or Target's, the net effect (captured by $\gamma_{pop} + \beta^{W}_{pop}$) is positive. The set estimates for the other market characteristics are generally neutral; that is, they cover both positive and negative values. It is therefore difficult to draw conclusions about Kmart's location preferences, after controlling for the impact of high values of population.

5 Counterfactuals

Because we did not explicitly model the decisions of box store retailers about how many distribution centers to build and where to place them, we are limited in the counterfactual scenarios that we can consider. In particular, we are constrained to consider either (1) small variations under which it would be reasonable to hold the number and location of distribution centers fixed, or (2) the situation in which distance to the distribution center is not a factor in firms' payoffs. The former is useful for discerning the role of particular local attributes in firms' decisions. The latter is useful for discerning the role of distribution

has a positive impact on Kmart's profit equal to the addition of 136,523 people in the market. (measured from the median population of 74,150.) For Target and Wal-Mart, this impact would be equivalent to a population increase of 82,250 and 56,450 respectively. These calculations use the midpoint of the interval $\gamma_{ln(pop)}$ estimated in the second stage.

costs in determining the spatial distribution of retail centers. This may have important implications for pricing – some locations may face little competition because of their distance form distribution center locations, leading consumers there to face higher prices.

In our simulation, we "turn off" the effect of distribution centers (i.e., $\beta^{f,DC} = 0 \forall f$) and determine the new equilibrium number of each firm's stores in each location.

- Comment on equilibrium properties (i.e., existence and uniqueness).
- Show old and new equilibria.
- Determine number of stores per capita by size category of MSA.
- Do people in smaller MSA's get more stores per capita under the counterfactual scenario? Given the cost of operating distribution centers, we would expect that firms strategically place the so as to allow them to service a large number of stores. We use our estimated model to determine what impact this motivation has on the level of competition faced by firms in various markets. Our interest is specifically in the impact on consumers. Are "isolated" locations served by fewer firms because of the desire on firms' part to reduce distribution distances? If so, we would expect that, from consumers' point of view, cost savings to firms arising from density economies could be offset by higher prices from reduced competition.[18] Is the distribution network problem causing certain groups to face higher prices?

6 Conclusions

[18] Footnote explaining that we are only looking at competition between big box stores – not competition between these stores and the "little guy", who should thrive in isolated locations. Did we want to try to add data on the little guys?

References

Bajari, P. and J. Fox (2006). "Complementarities and Collusion in an FCC Spectrum Auction" Working Paper, University of Chicago.

Bajari, P., J. Fox, and S. Ryan (2006). "Evaluating Wireless Carrier Consolidation Using Semiparametric Demand Estimation" Working Paper, University of Chicago.

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2006). "Estimating Static Models of Strategic Interactions" *Working Paper: University of Minnesota*

Bayer, P. and C. Timmins (2007). "Estimating Equilibrium Models of Sorting Across Locations" *Economic Journal*.

Beresteanu, A. and F. Molinari (2006)." Asymptotic Properties for a Class of Partially Identified Models " Working Paper: Cornell University.

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

Berry S. (1994). "Estimating Discrete Choice Models of Product Differentiation" *RAND Journal of Economics*. 25: 242-262.

Berry, S., J. Levinsohn, and A. Pakes (1995). "Automobile Prices in Market Equilibrium" *Econometrica*. 63(4):841-890.

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

Bresnahan, T. and P. Reiss (1991). "Entry and Competition in Concentrated Markets" *Journal of Political Economy*

Bresnahan, T. and P. Reiss (1987). "Do Entry Conditions Vary Across Markets?" Brookings Papers on Economic Activity: Microeconomics

Fox, J. (2006). "Estimating Matching Games with Transfers" Working Paper, University of Chicago.

Fox, J. (2007). "Semiparametric Estimation of Multinomial Discrete Choice Models Using a Subset of Choices" Forthcoming in the *RAND Journal of Economics*

Holmes, T. (2006). "The Diffusion of Wal-Mart and Economies of Density" Mimeo, University of Minnesota.

Jia, P. (2005). "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retail Industry" Working Paper: MIT.

Manski, C. (1975). "Maximum Score Estimation of the Stochastic Utility Model of Choice" *Journal of Econometrics*, 3: 205-228.

Manski, C. (1985). "Semiparametric Analysis of Discrete Response" *Journal of Econometrics*, 27: 313-333.

Manski, C. and E. Tamer (2002) "Inference on Regressions with Interval Data on a Regressor or Outcome" *Econometrica*. 70(2) pp. 519-46.

Mazzeo, M. (2002). "Product Choice and Oligopoly Market Structure" *Rand Journal of Economics*, 33(2), pp. 221-242

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006). "Moment Inequalities and Their Application" Working Paper: Harvard University.

Seim, K. (2006). "An Empirical Model of Firm Entry with Endogenous Product-Type Choices" Forthcoming in the *Rand Journal of Economics*.

Tamer, E. (2003) "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria" *Review of Economic Studies*, 70, pp. 147–65.

Sinnott, R.W. (1984). "Virtues of the Haversine" *Sky and Telescope* 68:159.

Table 1: Summary Statistics by Market

	Obs	Mean	Std Dev	Min	Max
Population in 2005	912	301,159	1,025,530	11,638	18,747,320
Land area (in sq. miles)	912	1716.63	2226.74	143.35	27259.87
Population density	912	152.61	225.52	1.78	2787.35
Average household size	912	2.55	0.19	2.18	3.75
Median Income	912	36,498.77	6,805.47	16,504.00	73,874.46
Percentage urban population	912	0.61	0.18	0.12	0.99
2002 retail sales (in $1000s)	912	3,169,755	10,632,448	73,768	183,700,000
Metro area	912	0.39	0.49	0	1
Northeast	912	0.10	0.30	0	1
Midwest	912	0.30	0.46	0	1
South	912	0.42	0.49	0	1
West	912	0.17	0.37	0	1
Distance to closest Kmart DC (in miles)	912	194.14	121.67	12.06	661.70
Distance to closest Target DC (in miles)	912	146.49	104.12	8.75	682.44
Distance to closest Wal-Mart DC (in miles)	912	92.64	64.14	8.75	506.67
Number of discount stores	912	6.19	14.12	1	161
Number of Kmart stores	912	1.42	3.43	0	43
Number of Target stores	912	1.58	5.44	0	72
Number of Wal-Mart stores	912	3.20	6.37	0	85

Table 2: Summary Statistics by Store

Variable	Kmart	Target	Wal-Mart
Population in 2005	2353013 (3975324)	3492545 (4282348)	1886391 (3105332)
Land area (in sq. miles)	3624.8 (3839.3)	4931.8 (4202.6)	3787.7 (3823.9)
Population density	534.9 (639.5)	696.0 (716.0)	414.5 (510.7)
Average household size	2.6 (0.2)	2.6 (0.2)	2.6 (0.2)
Median income	42083.2 (7411.4)	45536.4 (7474.9)	41024.4 (7562.3)
Percentage urban population	0.78 (0.18)	0.86 (0.12)	0.76 (0.18)
2002 retail sales (in $1000s)	24765258 (40070260)	36653720 (42806332)	19898874 (31626914)
Distance to closest own DC (in miles)	155.3 (118.7)	123.9 (89.8)	79.2 (49.3)
Metro area	0.78 (0.42)	0.95 (0.21)	0.68 (0.47)
Northeast	0.18 (0.39)	0.12 (0.33)	0.12 (0.33)
Midwest	0.30 (0.46)	0.28 (0.45)	0.24 (0.43)
South	0.33 (0.47)	0.34 (0.47)	0.47 (0.50)
West	0.19 (0.39)	0.26 (0.44)	0.16 (0.37)
	1295	1439	2914

Table 3: First Stage Estimates of Firm-Specific Coefficients

	I		II		III		IV		V	
	β	s.e.	β	s.e.	β	s.e.	β	s.e.	β	s.e.
Kmart-Specific										
$ln(N_{own}+1)$	-4.399	*0.346*	-4.931	*0.238*	-4.812		-4.663	*0.472*	-0.159	0.129
$ln(N_{other}+1)$	-58.059	*1.328*	-58.357	*0.993*	-59.580		-57.748	*1.106*	-10.147	0.428
Distance to DC	-1		-1				-1		-1	
Target-Specific										
$ln(N_{own}+1)$	-20.133	*0.794*	-21.397	*0.631*	-21.272		-20.964	*0.731*	-3.437	0.523
$ln(N_{other}+1)$	-44.770	*1.143*	-43.892	*0.710*	-46.221		-44.434	*0.934*	-9.614	0.498
Distance to DC	-0.952	*0.267*	-0.657	*0.258*	-1.068		-0.921	*0.247*	-1.113	0.405
ln(Median Income)	-1.810	*2.071*	2.0851†	*0.652*	-1.301					
ln(Population)	2.419	*0.911*	2.0851†	*0.652*	3.212		3.113	*3.642*	5.946	0.413
Avg. Household Size	-1.573	*1.325*	0.680	*0.854*						
% Urban	8.976	*5.001*	5.644	*3.145*						
Population Density										
Northeast	2.275	*0.780*	2.239	*0.598*	2.419		2.320	*0.986*	-5.646	1.246
Midwest	-2.747	*0.726*	-2.002	*0.621*	-2.447		-2.490	*0.918*	0.397	0.605
West	-4.600	*0.854*	-3.938	*0.537*	-4.288		-3.888	*0.846*	-0.327	0.381
Wal-mart-Specific										
$ln(N_{own}+1)$	-30.709	*0.751*	-31.598	*0.729*	-32.091		-31.040	*0.709*	-4.538	0.294
$ln(N_{other}+1)$	-29.438	*1.206*	-31.317	*0.755*	-30.412		-29.438	*0.843*	-10.290	0.513
Distance to DC	-0.210	*0.534*	-1.397	*0.400*	-0.722		-0.760	*0.501*	-0.212	0.574
ln(Median Income)	2.012	*2.977*	-0.9726†	*0.550*	0.195					
ln(Population)	-2.459	*1.115*	-0.9726†	*0.550*	-2.439		-2.444	*2.721*	3.289	0.418
Avg. Household Size	-1.637	*1.608*	0.562	*1.148*						
% Urban	0.057	*3.906*	1.169	*2.074*						
Population Density										
Northeast	4.037	*0.658*	3.491	*0.554*	4.575		4.435	*0.959*	-3.056	1.079
Midwest	-2.250	*0.828*	-1.515	*0.635*	-1.755		-1.635	*0.883*	-0.360	0.967
West	-3.101	*0.888*	-2.554	*0.648*	-3.223		-3.009	*0.716*	-5.069	1.036
Common										
ln(population)	‡		‡		‡		‡		6.175	0.103
Northeast									-0.927	0.759
Midwest									-0.561	0.427
West									-0.121	0.346
Number of Comparisons	15000		15000		15000		15000		30000	
Score	13656		13635		13632		13635		20208	

In all models, the coefficient on distance to Kmart's closest distribution center is normalized to -1.
† In this specification, the coefficients on ln(Median Income) and ln(Population) were constrained to be equal.
‡ These common components of the profit function are differenced out as part of the location fixed effect. They are estimated in a separate second stage, the results of which are reported in Table 4. In Specification V, because we have dropped the unobservable location-specific characteristic, there is no need to use the differencing approach; all parameters can be recovered in a single stage.

Table 4: Second Stage Set Estimates for the Components of the Common Fixed Effect

	I(a)	I(b)	IV (a)	IV (b)	IV (c)
ln(Median Income)	(-55.311, 31.107)			[-55.422, 33.014]	
	[-66.269, 42.064]			[-66.218, 43.810]	
ln(Population)	(28.814, 43.929)		(33.770, 43.949)	(33.785, 45.839)	
	[26.582, 46.161]		[31.989, 45.729]	[31.864, 47.759]	
ln(Median Income * Pop)		(25.510, 38.197)			(26.798, 39.100)
		[23.516, 40.191]			[24.852, 41.047]
Avg. Household Size	(-34.809, 21.736)	(-32.677, 23.164)		(-33.781, 23.442)	(-34.015, 23.138)
	[-43.592, 30.519]	[-42.244, 32.731]		[-42.198, 31.858]	[-43.211, 32.334]
% Urban	(-20.539, 73.130)	(-17.679, 76.260)			(-18.474, 77.290)
	[-31.350, 83.941]	[-28.725, 87.306]			[-29.696, 88.512]
NE	(-25.994, 7.923)	(-29.750, 1.550)	(-27.781, 3.286)	(-28.167, 5.698)	(-31.072, 0.737)
	[-31.553, 13.482]	[-35.573, 7.373]	[-32.685, 8.190]	[-33.444, 10.974]	[-36.523, 6.187]
Midwest	(-12.376, 23.303)	(-18.929, 14.156)	(-10.487, 22.164)	(-10.268, 25.231)	(-18.575, 14.644)
	[-15.947, 26.873]	[-22.047, 17.274]	[-13.564, 25.241]	[-13.775, 28.738]	[-21.768, 17.837]
West	(-19.238, 16.984)	(-24.002, 9.905)	(-15.900, 14.841)	(-16.555, 19.533)	(-24.500, 9.932)
	[-23.481, 21.228]	[-28.395, 14.298]	[-19.865, 18.805]	[-20.791, 23.769]	[-28.853, 14.285]
Constant	(-534.167, -213.251)	(-590.286, -373.147)	(-508.999, -382.095)	(-562.434, -245.895)	(-603.315, -388.664)
	[-572.481, -174.936]	[-626.871, -336.562]	[-528.890, -362.205]	[-601.258, -207.070]	[-639.935, -352.044]
Number of Observations	853	853	866	866	866

Specifications I(a) and I(b) use the first-stage estimates from specification I in Table 3, in which Target- and Wal-Mart-specific coefficients were estimated on all the listed covariates. Similarly, specifications IV(a)-IV(c) use the first-stage estimates from specifications IV in Table 3, which only estimated firm-specific coefficients on ln(population) and the regional dummies. The intervals appearing in square brackets represent the largest 95% confidence collection, i.e. the largest set that, when specified as a null hypothesis for the true value of the population identification region, cannot be rejected by a 95% test.

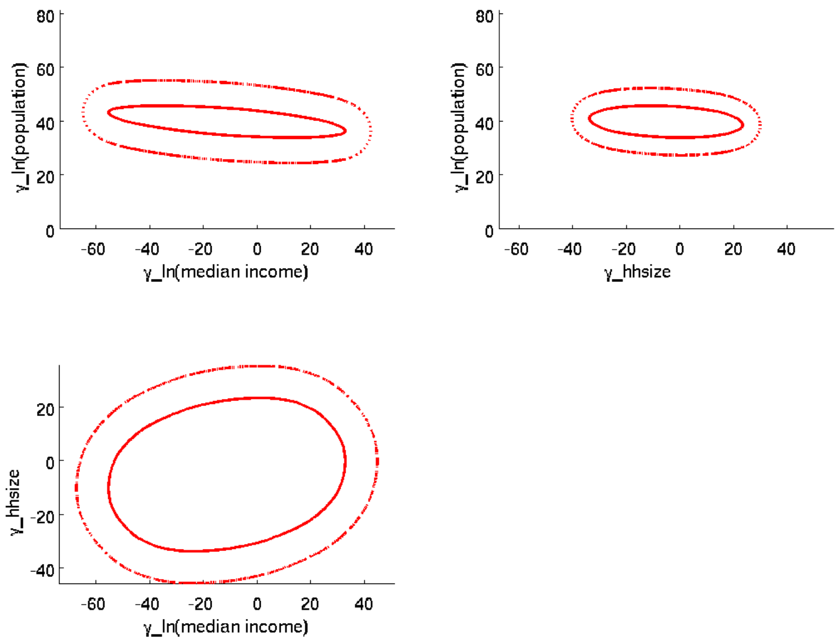

Figure 1: Second Stage Set Estimates - 2D Projections