What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry*

Panle Jia†

First version: November 2005. This version: July 2006

Abstract

In the past few decades multi-store retailers, especially those with a hundred or more stores, have experienced substantial growth. At the same time, there is widely reported public outcry over the impact of these chain stores on small retailers and local communities. This paper develops an empirical model to assess the impact of chain stores on the profitability and entry/exit decisions of small discount retailers and to quantify the size of the scale economies within a chain. The model has two key features. First, it allows for flexible competition patterns among all players. Second, for chains, it incorporates the scale economies that arise from operating multiple stores in nearby regions. In doing so, the model relaxes the commonly used assumption that entry in different markets is independent. The estimation exploits a unique data set that covers the discount retail industry from 1988 to 1997 and yields interesting results. First, Wal-Mart's expansion from the late 1980s to the late 1990s explains about fifty to seventy percent of the net change in the number of small discount retailers. Failure to address the endogeneity of the firms' entry decisions would result in underestimating this impact by fifty to sixty percent. Second, scale economies were important for both Kmart and Wal-Mart, but the magnitude did not grow proportionately with the chains' sizes. Finally, direct government subsidies to either chains or small retailers are unlikely to be effective in increasing the number of firms or the level of employment.

Keywords: Competition, Entry, Chain effect, Cross-sectional Dependence
JEL Classifications: L13, L81, L52, C13, C61

*This paper is a revision of chapter one of my thesis. I am deeply indebted to my committee members – Steven Berry, Penny Goldberg, Hanming Fang, and Philip Haile – for their continual support and encouragement. Special thanks go to Pat Bayer and Alvin Klevorick, who have been very generous with their help. I have also benefited from constructive discussions with Donald Andrews, Donald Brown, Judy Chevalier, Yuichi Kitamura, Herbert Scarf, and my colleagues at Yale University.

†Department of Economics, M.I.T. Email: pjia@mit.edu. Comments are welcome.

"Bowman's [in a small town in Georgia] is the eighth 'main street' business to close since Wal-Mart came to town.... For the first time in seventy-three years the big corner store is empty." Archer and Taylor, *Up against the Wal-Mart*.

"There is ample evidence that a small business need not fail in the face of competition from large discount stores. In fact, the presence of a large discount store usually acts as a magnet, keeping local shoppers....and expanding the market...." Morrison Cain, Vice president of International Mass Retail Association.

1 Introduction

The landscape of the U.S. retail industry has changed considerably over the past few decades as the result of two closely related trends. One is the rise of discount retailing; the other is the increasing prevalence of large retail chains. In fact, the discount retailing sector is almost entirely controlled by chains. In 1997, the top three chains (Wal-Mart, Kmart, and Target) accounted for 72.7% of total sector sales and 54.3% of the discount stores.

Discount retailing is a fairly new concept, with the first discount stores appearing in the 1950s. The leading magazine for the discount industry, *Discount Merchandiser*, defines a modern discount store as a departmentalized retail establishment that makes use of self-service techniques to sell a large variety of hard goods and soft goods at uniquely low margins.[1,2] Over the span of several decades, the sector has emerged from the fringe of the retail industry and become part of the mainstream.[3] From 1960 to 1997, the total sales revenue of discount stores, in real terms, increased 15.6 times, compared with an increase of 2.6 times for the entire retail industry.

As the discount retailing sector continues to grow, opposition from other retailers, especially small ones, begins to mount. The critics tend to associate discounters and other big retailers with small-town problems caused by the closing of small firms, such as the decline of downtown shopping districts, eroded tax bases, decreased employment, and the disintegration of closely

[1] See the annual report "The True Look of the Discount Industry" in the June issue of *Discount Merchandiser* for the definition of the discount retailing, the sales and store numbers for the top 30 largest firms, as well as the industry sales and total number of discount stores.

[2] According to *Annual Benchmark Report for Retail Trade and Food Services* published by the Census Bureau, from 1993 to 1997, the average markup for regular department stores was 27.9%, while the average markup for discount stores was 20.9%. Both markups increased slightly from 1998 to 2000.

[3] In 1997, the discount retailing sector accounted for 15% of total retail sales. The other retail sectors are: building materials, food stores, automotive dealers, apparel, furniture, eating and drinking places, and miscellaneous retail.

knit communities. Partly because tax money is used to restore the blighted downtown business districts and to lure the business of big retailers with various forms of economic development subsidies, the effect of big retailers on small firms and local communities has become a matter of public concern.[4] My first goal in this paper is to quantify the impact of national discount chains on the profitability and entry and exit decisions of small retailers from the late 1980s to the late 1990s.

The second salient feature of retail development in the past several decades, including in the discount sector, is the increasing dominance of large chains. In 1997, retail chains with a hundred or more stores accounted for 0.07% of the total number of firms, yet they controlled 21% of the establishments and accounted for 37% of sales and 46% of retail employment.[5] Since the late 1960s, their share of the retail market more than doubled. In spite of the dominance of chain stores, few empirical studies (except Holmes (2005) and Smith (2004)) have quantified the potential advantages of chains over single-unit firms, in part because of the modeling difficulties.[6] In entry models, for example, the store entry decisions of multi-unit chains are related across markets. Most of the literature assumes that entry decisions are independent across markets and focuses on competition among firms within each local market. My second objective here is to extend the entry literature by relaxing the independence assumption, and to quantify the advantage of operating multiple units by explicitly modeling chains' entry decisions in a large number of markets.

The model has two key features. First, it allows for flexible competition patterns among all retailers. Second, it incorporates the potential benefits of locating multiple stores near one another. Such benefits, which I group as "the chain effect," can arise through several different channels. For example, there may be significant scale economies in the distribution system. Stores located near each other can split advertising costs or employee training costs, or they can share knowledge about the specific features of local markets.

The chain effect causes profits of stores in the same chain to be spatially related. As a result, choosing store locations to maximize total profit is complicated, since with N markets there are 2^N possible location choices. In the current application, there are more than 2,000 markets and the number of possible location choices exceeds 10^{600}. When several chains compete against each other, solving for the Nash equilibrium becomes further involved, as firms balance the

[4] See *The Shils Report (1997): Measuring the Economic and Sociological Impact of the Mega-Retail Discount Chains on Small Enterprises in Urban, Suburban and Rural Communities.*

[5] See the 1997 Economic Census Retail Trade subject series *Establishment and Firm Size (Including Legal Form of Organization)*, published by the US Census Bureau.

[6] I discuss Holmes (2005) in detail below. Smith (2004) estimates the demand cross-elasticities between stores of the same firm and finds that mergers between the largest retail chains increase the price level by up to 7.4%.

gains from the chain effect against competition from rivals. I tackle this problem in several steps. First, I transform the profit maximization problem into a search for the fixed points of the necessary conditions. This transformation shifts the focus of the problem from a set with 2^N elements to the set of fixed points of the necessary conditions. The latter has a much smaller dimension, and is well-behaved with easy-to-locate minimum and maximum points. Having dealt with the problem of dimensionality, I take advantage of the supermodularity property of the game to search for the Nash equilibrium. Finally, in estimating the parameters, I adopt the econometric technique proposed by Conley (1999) to address the issue of cross-sectional dependence.

The analysis exploits a unique data set I collected that covers the entire discount retailing industry from 1988 to 1997, during which the two major national chains were Kmart and Wal-Mart.[7] The results indicate that Wal-Mart's expansion from the late 1980s to the late 1990s explains about fifty to seventy percent of the net change in the number of small discount retailers. Unobserved market-level profit shocks induce a positive correlation among the entry decisions of chains and small firms; failure to address this endogeneity issue would underestimate Wal-Mart's impact on small firms by fifty to sixty percent. Scale economies were important to both Wal-Mart and Kmart, but their importance did not grow proportionately with the size of the chains. Finally, government subsidies to either chains or small firms in this industry are not likely to be effective in increasing the number of firms or the level of employment.

The paper complements a recent study by Holmes (2005), which analyzes the diffusion process of Wal-Mart stores. Holmes quantifies the economies of density, defined as the cost savings from locating stores close to one another, a concept similar to the chain effect in this paper. The central insight in his paper is that markets vary in quality; in the absence of economies of density, Wal-Mart would open stores in the most profitable markets first and gradually expand to less profitable ones. Since profitable markets do not necessarily cluster, Wal-Mart should open stores erratically across regions. The actual opening process, however, displayed a regular pattern of diffusion from the South, where Wal-Mart's headquarters are, to other regions. Due to the complexity of the dynamics, with the state space growing exponentially with the number of markets and time periods, it is extremely difficult to solve Wal-Mart's optimization problem. By abstracting from competition and focusing on Wal-Mart's single-agent maximization problem, Holmes is able to exploit a perturbation approach to estimate the economies of density. The findings suggest that these economies of density are important.

[7] During the sample period, Target was a regional store that competed mostly in the big metropolitan areas in the Midwest with few stores in the sample. See the data section for more details.

Holmes' approach is appealing because he derives the magnitude of the economies of density from the dynamic expansion process. In contrast, I identify the chain effect from the stores' geographic clustering pattern. My approach abstracts from a number of important dynamic considerations. For example, it does not allow firms to delay store openings because of credit constraints, nor does it allow for any preemption motive as the chains compete and make simultaneous entry decisions. A dynamic model that incorporates both the competition effects and the chain effect would be ideal. However, given the great difficulty of estimating the economies of density in a single agent dynamic model, as Holmes (2005) shows, it is infeasible to estimate a dynamic model that also incorporates the strategic interactions within chains and between chains and small retailers. Since one of my main goals is to analyze the competition effects and perform policy evaluations, I adopt a two-stage model in which all players make a once-and-for-all decision, with chains moving first and small retailers moving second. I estimate the model seperately for 1988 and 1997, and exploit the coefficient estimates from both years to analyze the impact of chains on small retailers. The extension of the current framework to a dynamic model is left for future research.

This paper contributes to the entry literature initiated by Bresnahan and Reiss (1990, 1991) and Berry (1992), where researchers infer the firms' underlying profit functions by observing their equilibrium entry decisions across a large number of markets. To the extent that retail chains can be treated as multi-product firms whose differentiated products are stores with different locations, this paper relates to several recent empirical entry papers that endogenize firms' product choices upon entry. For example, Mazzeo (2002) considers the quality choices of highway motels, and Seim (2005) studies how video stores soften competition by choosing different locations. Unlike these studies, in which each firm chooses only one product, I analyze the behavior of multi-product firms whose product spaces are potentially large.

This paper is also related to a large literature on spatial competition in retail markets, for example, Pinkse *et. al.* (2002), Smith (2004), and Davis (2005). All of these models take the firms' locations as given and focus on price or quantity competition. I adopt the opposite approach. Specifically, I assume a parametric form for the firms' reduced-form profit functions from the stage competition, and examine how they compete spatially by balancing the chain effect against the competition effect of rivals' actions on their own profits.

Finally, the paper is part of the growing literature on Wal-Mart, which includes Stone (1995), Basker (2005a, 2005b), Hausman and Leibtag (2005), and Neumark *et al* (2005).

The remainder of the paper is structured as follows. Section 2 provides background information about the discount retailing sector. Section 3 describes the data set, and section 4 discusses the model. Section 5 proposes a solution algorithm for the game between chains and small firms when there is a large number of markets. Section 6 explains the estimation approach. Section 7 presents the results. Section 8 concludes. The appendix outlines the technical details not covered in section 5.

2 Industry background

Discount retailing is one of the most dynamic sectors in the retail industry. Table 1 (A) displays some statistics for the industry from 1960 to 1997. The sales revenue for this sector, in 2004 US dollars, skyrocketed from 12.8 billion in 1960 to 198.7 billion in 1997. In comparison, the sales revenue for the entire retail industry increased only modestly from 511.2 billion to 1313.3 billion during the same period. The number of discount stores multiplied from 1329 to 9741, while the number of firms dropped from 1016 to 230.

Chain stores dominate the discount retailing sector, as they do other retail sectors. In 1970, the 39 largest discount chains, with twenty-five or more stores each, operated 49.3% of the discount stores and accounted for 41.4% of total sales. By 1989, both shares had increased to roughly 88%. In 1997, the top 30 chains controlled about 94% of total stores and sales.

The principal advantages of chain stores include the central purchasing unit's ability to buy on favorable terms and to foster specialized buying skills; the possibility of sharing operating and advertising costs among multiple units; the freedom to experiment in one selling unit without risk to the whole operation. Stores also frequently share their private information about local markets and learn from one another's managerial practices. Finally, chains can achieve economies of scale by combining wholesaling and retailing operations within the same business unit.

Until the late 1990s, the two most important national chains were Kmart and Wal-Mart. Each firm opened its first store in 1962. The first Kmart was opened by the variety-chain Kresge. Kmart stores were a new experiment that provided consumers with quality merchandise at prices considerably lower than those of regular retail stores. To reduce advertising costs and to minimize customer service, these stores emphasized nationally advertised brand-name products. Consumer satisfaction was guaranteed, and all goods could be returned for a refund or an exchange (See Vance and Scott (1994), pp32). These practices were an instant success, and Kmart grew rapidly in the 1970s and 1980s. By the early 1990s, the firm had more than 2200 stores nationwide. In the late 1980s, Kmart tried to diversify and pursued various forms of specialty retailing in pharmaceutical products, sporting goods, office supplies, building materials, etc. The attempt was unsuccessful, and Kmart eventually divested itself of these interests by the late 1990s. Struggling with its management failures throughout the

1990s, Kmart maintained roughly the same number of stores; the opening of new stores offset the closing of existing ones.

Unlike Kmart, which was initially supported by an established retail firm, Wal-Mart started from scratch and grew relatively slowly in the beginning. To avoid direct competition with other discounters, it focused on small towns in southern states where there were few competitors. Starting in the early 1980s, the firm began its aggressive expansion process that averaged 140 store openings per year. In 1991, Wal-Mart replaced Kmart as the largest discounter. By 1997, Wal-Mart had 2362 stores (not including the wholesale clubs) in all states, including Alaska and Hawaii.

As the discounters continue to grow, small retailers start to feel their impact. There are extensive media reports on the controversies associated with the impact of large chains on small retailers and on local communities in general. As early as 1994, the United States House of Representatives convened a hearing titled "The Impact of Discount Superstores on Small Businesses and Local Communities." Witnesses from mass retail associations and small retail councils testified, but no legislation followed, partly due to a lack of concrete evidence. In April 2004, the University of California, Santa Barbara, held a conference that centered on the cultural and social impact of the leading discounter, Wal-Mart. In November 2004, both CNBC and PBS aired documentaries that displayed the changes Wal-Mart had brought to the society.

3 Data

The available data sets dictate the modeling approach used in this paper. Hence, I discuss them before introducing the model.

3.1 Data sources

There are three main data sources. The data on discount chains come from an annual directory published by Chain Store Guide Inc. The directory covers all operating discount stores of more than ten thousand square feet. For each store, the directory lists its name, size, street address, telephone number, store format, and firm affiliation.[8] The U.S. industry classification system changed from the Standard Industrial Classification System (SIC) to the North American Industry Classification System (NAICS) in 1998. To avoid potential inconsistencies in the

[8] The directory stopped providing store size information in 1997 and changed the inclusion criterion to 20,000 square feet in 1998. The store formats include membership stores, regional offices, and in later years distribution centers.

industry definition, I restrict the sample period to the ten years before the classification change. As first documented in Basker (2005), the directory was not fully updated for some years. Fortunately, it was fairly accurate for the years used in this study. See appendix 10 for details.

The second data set, the County Business Patterns, tabulates at the county level the number of establishments by employment size category for very detailed industry classifications. However, data disaggregated at the three-digit or finer SIC levels are unusable because of data suppression due to confidentiality requirements.[9] There are eight retail sectors at the two-digit SIC level: building materials and garden supplies, general merchandise stores (or discount stores), food stores, automotive dealers and service stations, apparel and accessory stores, furniture and home-furnishing stores, eating and drinking places, and miscellaneous retail. I focus on small general merchandise stores with nineteen or fewer employees, which are the direct competitors of the discount chains.

Data on county level population are downloaded from the websites of U.S. Census Bureau (before 1990) and the Missouri State Census Data Center (after 1990). Other county level demographic and retail sales data are from various years of the decennial census and the economic census.

3.2 Market definition and data description

In this paper, a market is defined as a county. Although the Chain Store Guide publishes the detailed street addresses for the discount stores, information about small firms is available only at the county level. Many of the market size variables, like retail sales, are also available only at the county level.

I focus on counties with an average population between 5,000 and 64,000 from 1988 to 1997. There are 2065 such counties among a total of 3140 in the U.S. According to Vance and Scott (1994), the minimum county population for a Wal-Mart store was 5,000 in the 1980s, while Kmart concentrated in places with a much larger population. 9% of all U.S. counties were smaller than 5,000 and were unlikely to be a potential market for either chain, while 25% of them were large metropolitan areas with an average population of 64,000 or more. These big counties typically included multiple self-contained shopping areas, and consumers were unlikely to travel across the entire county to shop. The market configuration in these big counties tended to be very complex with a large number of competitors and many market

[9] Title 13 of the United States Code authorizes the Census Bureau to conduct censuses and surveys. Section 9 of the same Title requires that any information collected from the public under the authority of Title 13 be maintained as confidential and that no estimates be published that would disclose the operations of an individual firm.

niches. For example, in the early 1990s, there were more than one hundred big discounters and close to four hundred small general merchandise stores in Los Angeles County, one of the largest counties. Modeling firms' strategic behavior in these markets requires geographic information more detailed than that provided by the county level data.

During the sample period, there were two national chains: Kmart and Wal-Mart. The third largest chain, Target, had 340 stores in 1988 and about 800 stores in 1997. Most of them were located in metropolitan areas in the Midwest, with on average fewer than twenty stores in the counties studied here. I do not include Target in the analysis.[10]

In the sample, only eight counties had two Kmart stores and forty-nine counties had two Wal-Mart stores in 1988; the figures were eight and sixty-six counties, respectively, in 1997. The current specification abstracts from the choice of the number of opening stores and considers only market entry decisions, as there is not enough variation in the data to identify the profit for the second store in the same market. The algorithm proposed in this paper can be applied with little modification to models that also incorporate the store-number choice.

Table 1 (B) presents summary statistics of the sample for the years 1988 and 1997. The average county population grew from 22,470 to 24,270, an increase of 8%. Retail sales per capita, in 1984 dollars, rose 10%, from \$3,690 to \$4,050. The average percentage of urban population was 30% in 1988 and increased to 33% in 1997. About one quarter of the counties was primarily rural with a small urban population, which is why the average across the counties seems somewhat low. 41% of the counties were in the Midwest (which includes the Great Lakes region, the Plains region, and the Rocky Mountain region, as defined by the Bureau of Economic Analysis), and 50% of the counties were in the southern regions (including the Southeast region and the Southwest region), with the rest in the Far West and the Northeast regions. Kmart had stores in 21% of the counties at the beginning of the sample period, and the number dropped slightly to 19% at the end. In comparison, Wal-Mart had stores in 32% of the counties in 1988 and in 48% of them in 1997. The average number of small firms decreased quite a bit over the same period, from 3.86 to 3.49. The median was three, with a maximum of twenty-five small firms in 1987, and nineteen in 1997. The percentage of counties with six or more small firms dropped from 22% to 18%, while the percentage of counties with at most one small firm increased from 18% to 22% over the sample period.

[10] The rest of the discount chains are much smaller and are all regional. They are not included in the analysis.

4 Modeling

4.1 Model setup

The model I develop is a two-stage game with complete information. In stage one, Kmart and Wal-Mart simultaneously choose store locations to maximize their total profits in all markets. In stage two, small firms observe Kmart's and Wal-Mart's choices and decide whether to enter the market.[11] Once the entry decisions are made, firms compete and profits are realized. All firms are fully rational with perfect knowledge of their rivals' profitability and the payoff structure. When Kmart and Wal-Mart make location choices in the first stage, they take into consideration the small retailers' reaction. There are no entry barriers; small firms enter the market until profit for an extra entrant becomes negative.

In reality, small retailers existed long before the era of the discount chains. As the chains emerge in the retail industry, small firms either continue their operations and compete with the chains or exit the market. This might suggest a three-stage model, with each stage corresponding to each of these events. However, given the nature of the retail industry, the sunk entry cost is unlikely to be significant for the small firms. In other words, their first stage decisions are irrelevant, and the small retailers respond to chain stores' entry decisions in the third stage.[12] In contrast, I implicitly assume that chains can commit to their entry decisions and do not further adjust after small firms enter. This is based on the observation that most chain stores enter with a long-term lease of the rental property, and in many cases they invest considerably in the infrastructure construction associated with establishing a big store.

4.2 The profit function

One way to obtain the profit function is to start from primitive assumptions of supply and demand in the retail markets, and derive the profit functions from the equilibrium conditions. Without any price, quantity, or sales data, and with very limited information on store characteristics, this approach is extremely demanding on data and relies heavily on the primitive assumptions. Instead, I follow the convention in the entry literature and assume that firms' profit functions from the stage competition take a linear form and that profits decline in the

[11] I have implicitly assumed that small firms, which are firms with one to nineteen employees, are single-unit stores.

[12] It is possible that small firm owners have invested considerably in their communities to establish a customer base and a good reputation. These kinds of investment is sunk if stores are closed. However, these investment can be partially recovered when the firm owners switch to other retail sectors to make use of the consumer goodwill.

presence of rivals.

Let $D_{i,m} \in \{0,1\}$ stand for chain i's strategy in market m, where $D_{i,m} = 1$ if chain i operates a store in market m and $D_{i,m} = 0$ otherwise. $D_i = \{D_{i,1}, ..., D_{i,M}\}$ is a vector indicating chain i's location choices for the entire set of markets. Let $D_{j,m}$ denote rival j's strategy in market m, and $N_{s,m}$ the number of small firms in market m. X_m, ε_m, and $\eta_{i,m}$ stand for a vector of observed market size variables, the market level profit shock, and firm i's private profit shock in market m, respectively. Finally, let Z_{ml} designate the distance from market m to market l in miles, and $Z_m = \{Z_{m1}, ..., Z_{mM}\}$.

The profit function for chain i in market m takes the following form:

$$\begin{aligned}\Pi_{i,m}(D_i, D_{j,m}, N_{s,m};\ X_m, Z_m, \varepsilon_m, \eta_{i,m}) \quad = \quad & D_{i,m} * [X_m\beta_i + \delta_{ij} D_{j,m} + \delta_{is}\ln(N_{s,m}+1) \\ & + \delta_{ii}\Sigma_{l\neq m}\frac{D_{i,l}}{Z_{ml}} + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}] \qquad (1)\end{aligned}$$

where $i, j \in \{k, w\}$, with "k" for Kmart and "w" for Wal-Mart. Note the presence of D_i in $\Pi_{i,m}(\cdot)$: profit in market m depends on the number of stores chain i has in other markets.

Profit from staying outside the market is normalized to 0. Chains maximize their total profits in all markets $\Sigma_m \Pi_{i,m}$. In equilibrium, the number of small firms is a function of Kmart's and Wal-Mart's first stage decisions: $N_{s,m}(D_{k,m}, D_{w,m})$. When making location choices, the chains take into consideration the impact of small firms' reactions on their own profits.

There are several components in chain i's profit $\Pi_{i,m}$ in market m: the observed market size $X_m\beta_i$ that is parameterized by demand shifters, like population, the extent of urbanization, etc.; the unobserved profit shock $\sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}$, known to the firms but unknown to the econometrician; the competition effects $\delta_{ij}D_{j,m} + \delta_{is}\ln(N_{s,m}+1)$, as well as the chain effect $\delta_{ii}\Sigma_{l\neq m}\frac{D_{i,l}}{Z_{ml}}$. Notice that the observed market size component $X_m\beta_i$ is allowed to differ for different players. X_m includes all factors that influence profits, and β_i picks up the factors that are relevant for firm i. For example, Kmart might have some advantage in the Midwest, while Wal-Mart stores might be more profitable in markets close to their headquarters.

The unobserved profit shock has two elements: ε_m, the market-level profit shifter that affects both chains and small firms, and $\eta_{i,m}$, a firm-specific profit shock. ε_m is assumed to be i.i.d. across markets, while $\eta_{i,m}$ is assumed to be i.i.d. across both firms and markets. $\sqrt{1-\rho^2}$ (with $0 \leq \rho \leq 1$) measures how important the market component is. In principle, it can differ for each chain and for small firms. For example, the market specific business environment – how developed the infrastructure is, whether the market has sophisticated shopping facilities, and the stance of the local community toward large corporations including big retailers – might matter more to chains than to small firms. In the baseline specification, I restrict ρ to be the same across all players. Relaxing it does not improve the fit much. $\eta_{i,m}$ incorporates the unobserved store level heterogeneity, including the management ability, the display style and shopping environment, employees' morale or skills, etc. As is standard in discrete choice models, the scale of the parameter coefficients and the variance of the error term are not separately identified. I normalize the variance of the error term to 1 by assuming that both ε_m and $\eta_{i,m}$ are standard normal random variables.

The competition effect from the rival chain is captured by $\delta_{ij} D_{j,m}$, where $D_{j,m}$ is one if rival j operates a store in market m. $\delta_{is} \ln(N_{s,m}+1)$ denotes the effect of small firms on chain i's profit. The addition of 1 in $\ln(N_{s,m}+1)$ is used to avoid $\ln 0$ for markets without any small firms. The log form allows the incremental competition effect to taper off when there are many small firms.

The last unexplained term in the bracket, $\delta_{ii}\Sigma_{l \neq m}\frac{D_{i,l}}{Z_{ml}}$, captures the chain effect, the benefit that having stores in other markets generates for the profitability in market m. δ_{ii} is assumed to be non-negative. Stores split the costs of operation, delivery, and advertising among nearby ones to achieve scale economies. They also share knowledge of the localized markets and learn from one another's managerial success. All these factors suggest that having stores nearby benefits the operation in market m, and that the benefit declines with the distance. Following Bajari and Fox (2005), I divide the spillover effect by the distance between the two markets Z_{ml}, so that profit in market m is increased by $\delta_{ii}\frac{D_{i,l}}{Z_{ml}}$ if there is a store in market l that is Z_{ml} miles away. This simple formulation serves two purposes. First, it is a parsimonious way to capture the fact that it might be increasingly difficult to benefit from stores that are farther away. Second, the econometric technique exploited in the estimation requires the dependence among the observations to die away sufficiently fast. I also assume that the chain effect takes place among counties whose centroids are within fifty miles, or roughly an area that expands seventy-five miles in each direction. Including counties within a hundred miles increases substantially the computing time with little change in the parameters.

This paper focuses on the chain effect that is "localized" in nature. Some chain effects are "global": for example, the gain that arises from a chain's ability to buy a large volume at a discount. The latter benefits affect all stores the same, and cannot be separately identified from the constant of the profit function. Hence, the estimates δ_{ii}, should be interpreted as a lower bound to the actual advantages enjoyed by a chain.

Profit for a small firm that operates in market m is:

$$\begin{aligned}\Pi_{s,m}(D_{k,m}, D_{w,m}, N_{s,m}; X_m, \varepsilon_m, \eta_{s,m}) &= X_m\beta_s + \Sigma_{i=k,w}\delta_{si}D_{i,m} + \delta_{ss}\ln(N_{s,m}) \\ &\quad + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{s,m}\end{aligned} \tag{2}$$

All small firms are symmetric with the same profit function $\Pi_{s,m}(\cdot)$. For markets with no small firms, the entry condition implies that profit for a single small firm is negative: $X_m\beta_s + \Sigma_{i=k,w}\delta_{si}D_{i,m} + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{s,m} < 0$. For markets with $N_{s,m}$ small firms, $\Pi_{s,m}(N_{s,m}) \geq 0$ and $\Pi_{s,m}(N_{s,m}+1) < 0$. The term $\delta_{ss}\ln(N_{s,m})$ captures the competition among small firms, while $\Sigma_{i=k,w}\delta_{si}D_{i,m}$ denotes the impact of Kmart and Wal-Mart on small firms. The static nature of the model does not allow separate identification of the different channels through which the competition effect takes place. For example, one can't tell how much of the competition effect is due to the forced exit of small firms, and how much is due to the preemption that reduces entry of small firms.

The market-level error term ε_m makes the location choices of the chain stores $D_{k,m}$ and $D_{w,m}$, and the number of small firms $N_{s,m}$ endogenous in the profit functions, since a large ε_m leads to more entries of both chains and small firms. I explicitly address the issue of endogeneity by solving chains' and small firms' entry decisions simultaneously within the model. To estimate only the competition effects of big retailers on small firms $\{\delta_{si}\}_{i=\{k,w\}}$, without analyzing the equilibrium consequences of policy changes, it suffices to regress the number of small stores on market size variables, together with the number of chain stores, and to use instruments to correct the OLS bias of the competition effect. However, valid instruments for the presence of each of the rivals are difficult to find. Researchers have experimented with distance to headquarters or stores' planned opening dates to instrument for Wal-Mart's entry decisions.[13] It is much more difficult to find good instruments for Kmart. The R^2 of regressing Kmart stores' locations on their distance to headquarters is less than 0.005. Another awkward feature of the linear IV regression is that the predicted number of small firms can be negative. Limited dependent variable estimation avoids this problem, but accounting for endogeneity in the discrete games requires strong assumptions about the nature of the endogeneity that are not satisfied by the current model. Perhaps the best argument for the current approach, besides the possibility of analyzing policy experiments and studying the spillover among the chain stores, is that the structural model can exploit the chain effect to help with identification. The chain effect drives entry decisions of chain stores, but is not related to small firms' entry decisions, and serves as a natural excluded variable in the identification of the chains' competition effects on small firms.

Note that the above specification allows very flexible competition patterns among all the possible firm-pair combinations. The parameters to be estimated are $\{\beta_i, \delta_{ij}, \delta_{ii}, \rho\}, i, j \in \{k, w, s\}$, and the central parameters are the competition effects $\delta_{ij},\ i, j \in \{k, w, s\},\ i \neq j$ and the chain effects $\delta_{ii},\ i \in \{k, w\}$.

[13] See Neumark (2005) and Basker (2005a).

5 Solution algorithm

The unobserved market level profit shock ε_m, together with the chain effect $\delta_{ii}\Sigma_{l\neq m}\frac{D_{i,l}}{Z_{ml}}$, renders all of the discrete variables $D_{i,m}$, $D_{j,m}$, $D_{i,l}$, and $N_{s,m}$ endogenous in the profit functions (1) and (2). Finding the Nash equilibrium of this game is complicated. I take several steps to address this problem. Section 5.1 explains how to solve each chain's single agent problem, section 5.2 derives the solution algorithm for the game between two chains, and section 5.3 adds the small retailers and solves for the Nash equilibrium of the full model.

5.1 Chain i's single agent problem

In this subsection, let us focus on the chain's single-agent problem and abstract from competition. In the next two subsections I incorporate competition and solve the model for all players.

For notational simplicity, I have suppressed the firm subscript i and used X_m instead of $X_m\beta_i + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}$ in the profit function throughout this subsection. Let M denote the total number of markets, and let $\mathbf{D} = \{0,1\}^M$ denote the choice set. An element of the set $\mathbf{D}$ is an M-coordinate vector $D = \{D_1, ..., D_M\}$. The profit-maximization problem is:

$$\max_{D_1,\ldots,D_M\in\{0,1\}} \Pi = \sum_{m=1}^{M}\left[D_m * (X_m + \delta\Sigma_{l\neq m}\frac{D_l}{Z_{ml}})\right]$$

The choice variable D_m appears in the profit function in two ways. First, it directly determines profit in market m: the firm earns $X_m + \delta\Sigma_{l\neq m}\frac{D_l}{Z_{ml}}$ if $D_m = 1$, and zero if $D_m = 0$. Second, the decision to open a store in market m increases the profits in other markets through the chain effect.

The complexity of this maximization problem is twofold: first, it is a discrete problem of large dimension. In the current application, with $M = 2065$ and two choices for each market (enter or stay outside), the number of possible elements in the choice set $\mathbf{D}$ is 2^{2065}, or roughly 10^{600}. The naive approach that evaluates all of them to find the profit-maximizing vector(s) is infeasible. Second, the profit function is irregular: it is neither concave nor convex. Consider the relaxed function where D_m takes real values, rather than integers {0, 1}. The Hessian of this function is indefinite, and the usual first-order condition does not guarantee an optimum.[14]

[14] A symmetric matrix is positive (negative) semidefinite iff all the eigenvalues are non-negative (non-positive). The Hessian of the profit function (1) is a symmetric matrix with zero for all the diagonal elements. Its trace, which is equal to the sum of the eigenvalues, is zero. If the Hessian matrix has a positive eigenvalue, it has to have a negative one as well. There is only one possibility for the Hessian to be positive (or negative) semidefinite, which is that all the eigenvalues are 0. This is true only for the zero matrix H=0.

Even if one could exploit the first-order condition, the search with a large number of choice variables is a daunting task.

Instead of solving the problem directly, I transform it into a search for the fixed points of the necessary conditions for profit maximization. In particular, I exploit the lattice structure of the set of fixed points of an increasing function and propose an algorithm that obtains an upper bound D^U and a lower bound D^L for the profit-maximizing vector(s). With these two bounds at hand, I evaluate all vectors that lie between them to find the profit-maximizing ones.

The set of profit maximizing vectors may not be a singleton. For example, in the case of two markets with $X_1 = -1, X_2 = -1$, $\delta = 1$, and $Z_{1,2} = Z_{1,2} = 1$, both $D^* = \{0, 0\}$ and $D^{**} = \{1, 1\}$ maximize the total profit. Here I assume there is only one solution. In appendix 9.5, I show that allowing multiple optimal solutions is a straightforward extension.

Throughout this paper, the comparison between vectors is coordinate-wise. A vector D is bigger than vector D' if and only if every element of D is weakly bigger: $D \geq D'$ if and only if $D_m \geq D'_m$ $\forall m$. D and D' are unordered if neither $D \geq D'$ nor $D \leq D'$. They are the same if both $D \geq D'$ and $D \leq D'$.

Let the profit maximizer be denoted $D^* = \arg\max_{D \in \mathbf{D}} \Pi(D)$. The optimality of D^* implies that profit at D^* must be (weakly) higher than the profit at any one-market deviation:

$$\Pi(D_1^*, ..., D_m^*, ..., D_M^*) \geq \ \Pi(D_1^*, ..., D_m, ..., D_M^*), \forall m$$

which leads to:

$$D_m^* = 1[X_m + 2\delta\Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}} \geq 0], \forall m \tag{3}$$

The derivation of equation (3) is left to appendix 9.1. These conditions have the usual interpretation that $X_m + 2\delta\Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}}$ is market m's marginal contribution to total profit. This equation system is not definitional; it is a set of necessary conditions for the optimal vector D^*. Not all vectors that satisfy (3) maximize profit, but if D^* maximizes profit, it must satisfy these constraints.

Define $V_m(D) = 1[X_m + 2\delta\Sigma_{l \neq m} \frac{D_l}{Z_{ml}} \geq 0]$, and $V(D) = \{V_1(D), ..., V_M(D)\}$. $V(\cdot)$ is a vector function that maps from $\mathbf{D}$ into itself: $V : \mathbf{D} \to \mathbf{D}$. It is an increasing function: $V(D') \geq V(D'')$ whenever $D' \geq D''$, as δ_{ii} is assumed non-negative. By construction, the profit maximizer D^* is one of $V(\cdot)$'s fixed points. The following theorem, proved by Tarski (1955), states that the set of fixed points of an increasing function that maps from a lattice into itself is a lattice and has a greatest point and a least point. Appendix 9.2 describes the basic lattice theory.

Theorem 1 *Suppose that $Y(X)$ is an increasing function from a nonempty complete lattice $\mathbf{X}$ into $\mathbf{X}$.*

(a) The set of fixed points of $Y(X)$ is nonempty, $\sup_{\mathbf{X}}(\{X \in \mathbf{X}, X \leq Y(X)\})$ is the greatest fixed point, and $\inf_{\mathbf{X}}(\{X \in \mathbf{X}, Y(X) \leq X\})$ is the least fixed point.

(b) The set of fixed points of $Y(X)$ in $\mathbf{X}$ is a nonempty complete lattice.

A lattice in which each nonempty subset has a supremum and an infimum is complete. Any finite lattice is complete. A nonempty complete lattice has a greatest and a least element. Since the choice set $\mathbf{D}$ is a finite lattice, it is complete, and Theorem 1 can be directly applied. Several points are worth mentioning. First, $\mathbf{X}$ can be a closed interval or it can be a discrete set, as long as the set includes the greatest lower bound and the least upper bound for any of its nonempty subsets. That is, it is a complete lattice. Second, the set of fixed points is itself a nonempty complete lattice, with a greatest and a smallest point. Third, the requirement that $Y(X)$ is "increasing" is crucial; it can't be replaced by assuming that $Y(X)$ is a monotone function. Appendix 9.2 provides a counterexample where the set of fixed points for a decreasing function is empty.

Now I outline the algorithm that delivers the greatest and the least fixed point of $V(D)$, which are, respectively, an upper bound and a lower bound for the optimal solution vector D^*. To find D^*, I rely on an exhaustive search among the vectors lying between these two bounds.

Start with $D^0 = \sup(\mathbf{D}) = \{1, ..., 1\}$. The supremum exists because $\mathbf{D}$ is a complete lattice. Define a sequence $\{D^t\} : D^1 = V(D^0)$, and $D^{t+1} = V(D^t)$. By the construction of D^0, we have: $D^0 \geq V(D^0) = D^1$. Since $V(\cdot)$ is an increasing function, $V(D^0) \geq V(D^1)$, or $D^1 \geq D^2$. Iterating this process several times generates a decreasing sequence: $D^0 \geq D^1 \geq ... \geq D^t$. Given that D^0 has only M distinct elements and at least one element of the D vector is changed from 1 to 0 in each iteration, the process converges within M steps: $D^{T-1} = D^T, T \leq M$. Let D^U denote the convergent vector. D^U is a fixed point of the function $V(\cdot) : D^U = V(D^U)$. To show that D^U is indeed the greatest element of the set of fixed points, note that $D^0 \geq D'$, where D' is an arbitrary element of the set of fixed points. Applying the function $V(\cdot)$ to the inequality T times, we have $D^U = V^T(D^0) \geq V^T(D') = D'$.

Using the dual argument, one can show that the convergent vector derived from $D^0 = \inf(\mathbf{D}) = \{0, ..., 0\}$ is the least element in the set of fixed points. Denote it by D^L. In appendix 9.3, I show that starting from the solution to a constrained version of the profit maximization problem yields a tighter lower bound. There I also illustrate how a tighter upper bound can be obtained by starting with a vector $\tilde{D}$ such that $\tilde{D} \geq D^*$ and $\tilde{D} \geq V(\tilde{D})$.

With the two bounds D^U and D^L at hand, I evaluate all vectors that lie between them and find the profit-maximizing vector D^*.

5.2 The maximization problem with two competing chains

The discussion in the previous subsection abstracts from rival-chain competition and considers only the chain effect. With the competition effect from the rival-chain, the profit function for chain i becomes: $\Pi_i(D_i, D_j) = \Sigma_{m=1}^{M}[D_{i,m} * (X_{im} + \delta_{ii}\Sigma_{l \neq m}\frac{D_{i,l}}{Z_{ml}} + \delta_{ij}D_{j,m})]$, where X_{im} contains $X_m\beta_i + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}$.

To address the interaction between the chain effect and the rival-chain competition effect, I invoke the following theorem from Topkis (1978), which states that the best response function is decreasing in the rival's strategy when the payoff function is supermodular and has decreasing differences. Specifically:[15]

Theorem 2 *If* $\mathbf{X}$ *is a lattice,* K *is a partially ordered set,* $Y(X,k)$ *is supermodular in* X *on* $\mathbf{X}$ *for each* k *in* K*, and* $Y(X,k)$ *has decreasing differences in* (X,k) *on* $\mathbf{X} \times K$*, then* $\arg\max_{X \in \mathbf{X}} Y(X,k)$ *is decreasing in* k *on* $\{k : k \in K, \arg\max_{X \in \mathbf{X}} Y(X,k)$ *is nonempty*$\}$.

$Y(X,k)$ has decreasing differences in (X,k) on $\mathbf{X} \times K$ if $Y(X,k'') - Y(X,k')$ is decreasing in $X \in \mathbf{X}$ for all $k' \leq k''$ in K. Intuitively, $Y(X,k)$ has decreasing differences in (X,k) if X and k are substitutes. In appendix 9.4, I verify that the profit function $\Pi_i(D_i, D_j) = \Sigma_{m=1}^{M}[D_{i,m} * (X_{im} + \delta_{ii}\Sigma_{l \neq m}\frac{D_{i,l}}{Z_{ml}} + \delta_{ij}D_{j,m})]$ is supermodular in its own strategy D_i and has decreasing differences in (D_i, D_j). From Theorem 2, chain i's best response correspondence $\arg\max_{D_i \in \mathbf{D_i}} \Pi_i(D_i, D_j)$ decreases in rival j's strategy D_j. Similarly for chain j's best response to i's strategy.

As the simple example in section 5.1 illustrates, given the rival's strategy D_j, $\arg\max_{D_i \in \mathbf{D}} \Pi_i(D_i, D_j)$ can contain more than one element. For the moment, assume that $\arg\max_{D_i \in \mathbf{D}} \Pi_i(D_i, D_j)$ is a singleton for any given D_j. Appendix 9.5 discusses the case in which $\arg\max_{D_i \in \mathbf{D}} \Pi_i(D_i, D_j)$ has multiple elements. The extension involves the concepts of set ordering and increasing (decreasing) selection, but is fairly straightforward.

The set of Nash equilibria of a supermodular game is nonempty and it has a greatest element and a least element.[16,17] The current entry game is not supermodular, as the profit

[15] The original theorem is stated in terms of $\Pi(D,t)$ having increasing differences in (D,t), and $\arg\max_{D \in \mathbf{D}} \Pi(D,t)$ increasing in t. Replacing t with $-t$ yields the version of the theorem stated here.

[16] See Topkis (1978) and Zhou (1994).

[17] A game is supermodular if the payoff function $\Pi_i(D_i, D_{-i})$ is supermodular in D_i for each D_{-i} and each player i, and $\Pi_i(D_i, D_{-i})$ has increasing differences in (D_i, D_{-i}) for each i.

function has decreasing differences in the joint strategy space $\mathbf{D} \times \mathbf{D}$. This leads to a non-increasing joint best response function, and we know from the discussion after Theorem 1 that a non-increasing function on a lattice can have an empty set of fixed points. A simple transformation, however, restores the supermodularity property of the game. The trick is to define a new strategy space for one player (for example, Kmart) to be the negative of the original space. Let $\widetilde{\mathbf{D}}_k = -\mathbf{D}_k$. The profit function can be re-written as:

$$\begin{aligned}\Pi_k(-D_k, D_w) &= \Sigma_m(-D_{k,m}) * [-X_{km} + \delta_{kk}\Sigma_{l\neq m}\frac{-D_{k,l}}{Z_{ml}} + (-\delta_{kw})D_{w,m}] \\ \Pi_w(D_w, -D_k) &= \Sigma_m D_{w,m} * [X_{wm} + \delta_{ww}\Sigma_{l\neq m}\frac{D_{w,l}}{Z_{ml}} + (-\delta_{wk})(-D_{k,m})]\end{aligned}$$

It is easy to verify that the game defined on the new strategy space $(\widetilde{\mathbf{D}}_k, \mathbf{D}_w)$ is supermodular, therefore a Nash equilibrium exists. Using the transformation $\widetilde{\mathbf{D}}_k = -\mathbf{D}_k$, one can find the corresponding equilibrium in the original strategy space. In the following paragraphs, I explain how to find the desired Nash equilibrium directly in the space of $(\mathbf{D}_k, \mathbf{D}_w)$ using the "Round-Robin" algorithm, where each player proceeds in turn to update its own strategy.[18]

To obtain the equilibrium most profitable for Kmart, start with the smallest vector in Wal-Mart's strategy space: $D_w^0 = \inf(\mathbf{D}) = \{0, ..., 0\}$. Derive Kmart's best response $K(D_w^0) = \arg\max_{D_k \in \mathbf{D}} \Pi_k(D_k, D_w^0)$ given D_w^0, using the method outlined in section 5.1, and denote it by $D_k^1 = K(D_w^0)$. Similarly, find Wal-Mart's best response $W(D_k^1) = \arg\max_{D_w \in \mathbf{D}} \Pi_w(D_w, D_k^1)$ given D_k^1, again using the method in section 5.1, and denote it by D_w^1. Note that $D_w^1 \geq D_w^0$, by the construction of D_w^0. This finishes the first iteration $\{D_k^1, D_w^1\}$.

Fix D_w^1 and solve for Kmart's best response $D_k^2 = K(D_w^1)$. By Theorem 2, Kmart's best response decreases in the rival's strategy, so $D_k^2 = K(D_w^1) \leq D_k^1 = K(D_w^0)$. The same argument shows that $D_w^2 \geq D_w^1$. Iterating the process generates two monotone sequences: $D_k^1 \geq D_k^2 \geq ... \geq D_k^t$, $D_w^1 \leq D_w^2 \leq ... \leq D_w^t$. In every iteration, at least one element of the D_k vector is changed from 1 to 0, and one element of the D_w vector is changed from 0 to 1, so the algorithm converges within M steps: $D_k^T = D_k^{T-1}, D_w^T = D_w^{T-1}, T \leq M$. The convergent vectors (D_k^T, D_w^T) constitute an equilibrium: $D_k^T = K(D_w^T), D_w^T = W(D_k^T)$. Furthermore, this equilibrium gives Kmart the highest profit among the set of all equilibria.

That Kmart prefers the equilibrium (D_k^T, D_w^T) obtained using $D_w^0 = \{0, ..., 0\}$ to all other equilibria follows from two results: first, $D_w^T \leq D_w^*$ for any D_w^* that belongs to an equilibrium; second, $\Pi_k(K(D_w), D_w)$ decreases in D_w, where $K(D_w)$ denotes Kmart's best response function. Together they imply that $\Pi_k(D_k^T, D_w^T) \geq \Pi_k(D_k^*, D_w^*)$, $\forall$ $\{D_k^*, D_w^*\}$ that belongs to the set of Nash equilibria.

[18] See Topkis (1998) for a detailed discussion.

To show the first result, note that $D_w^0 \leq D_w^*$, by the construction of D_w^0. Since $K(D_w)$ decreases in D_w, $D_k^1 = K(D_w^0) \geq K(D_w^*) = D_k^*$. Similarly, $D_w^1 = W(D_k^1) \leq W(D_k^*) = D_w^*$. Repeating this process T times leads to $D_k^T = K(D_w^T) \geq K(D_w^*) = D_k^*$, and $D_w^T = W(D_k^T) \leq W(D_k^*) = D_w^*$. The second result follows from $\Pi_k(K(D_w^*), D_w^*) \leq \Pi_k(K(D_w^*), D_w^T) \leq \Pi_k(K(D_w^T), D_w^T)$. The first inequality holds because Kmart's profit function decreases in its rival's strategy, while the second inequality follows from the definition of the best response function $K(D_w)$.

By the dual argument, starting with $D_k^0 = \inf(\mathbf{D}) = \{0, ..., 0\}$ delivers the equilibrium that is most preferred by Wal-Mart. To search for the equilibrium that favors Wal-Mart in the southern region and Kmart in the rest of the country, one uses the same algorithm to solve the game separately for the south and the other regions.

5.3 Adding small firms

Incorporating small firms into the game is a straightforward application of backward induction, since the number of small firms in the second stage is a well-defined function $N_s(D_k, D_w)$. Chain i's profit function now becomes $\Pi_i(D_i, D_j) = \Sigma_{m=1}^M [D_{i,m} * (X_{im} + \delta_{ii} \Sigma_{l \neq m} \frac{D_{i,l}}{Z_{ml}} + \delta_{ij} D_{j,m} + \delta_{is} \ln(N_s(D_{i,m}, D_{j,m}) + 1)]$, where X_{im} is defined in the previous subsection (5.2). The profit function $\Pi_i(D_i, D_j)$ remains supermodular in D_i with decreasing differences in (D_i, D_j) under a minor assumption, which essentially requires that the net competition effect of rival D_j on chain i's profit is negative.[19]

The main computational burden in solving the full model with both chains and small retailers is the search for the best responses $K(D_w)$ and $W(D_k)$. In appendix 9.6, I discuss a few technical details related with the implementation.

6 Empirical implementation

6.1 Estimation

The model does not yield a closed form solution to firms' location choices conditioning on market size observables and a given vector of parameter values. Hence I turn to simulation methods. The ones most frequently used in the I.O. literature are the method of simulated log-likelihood (MSL) and the method of simulated moments (MSM).

[19] If we ignore the integer problem and approximate $\ln(N_s + 1)$ by $-(X_{sm} + \delta_{sk} D_k + \delta_{sw} D_w)$, then the assumption is: $\delta_{kw} - \frac{\delta_{ks}\delta_{sw}}{\delta_{ss}} < 0, \delta_{wk} - \frac{\delta_{ws}\delta_{sk}}{\delta_{ss}} < 0$. Essentially, these two conditions imply that when there are small stores, the 'net' competition effect of Wal-Mart (its direct impact, together with its indirect impact working through small stores) on Kmart's profit and that of Kmart on Wal-Mart's profit are still negative.

Implementing MSL is difficult because of the complexities in obtaining an estimate of the log-likelihood of the observed sample. The cross-sectional dependence among the observed outcomes in different markets indicates that the log-likelihood of the sample is no longer the sum of the log-likelihood of each market, and one needs an exceptionally large number of simulations to get a reasonable estimate of the sample's likelihood. Thus I adopt the MSM method to estimate the parameters in the profit functions $\theta_0 = \{\beta_i, \delta_{ii}, \delta_{ij}, \rho\}_{i=k,w,s} \in \mathbf{\Theta} \subset \mathbf{R}^P$. The following moment condition is assumed to hold at the true parameter value θ_0:

$$E[g(X_m, \theta_0)] = 0$$

where $g(X_m, \cdot) \in \mathbf{R}^L$ with $L \geq P$ is a vector of moment functions that specifies the differences between the observed equilibrium market structures and those predicted by the model.

A MSM estimator $\hat{\theta}$, minimizes a weighted quadratic form in $\Sigma_{m=1}^{M} \hat{g}(X_m, \theta)$:

$$\theta = \arg\min_{\theta \in \Theta} \frac{1}{M} \left[\sum_{m=1}^{M} \hat{g}(X_m, \theta) \right]' \mathbf{\Omega}_M \left[\sum_{m=1}^{M} \hat{g}(X_m, \theta) \right] \tag{4}$$

where $\hat{g}(\cdot)$ is a simulated estimate of the true moment function, and $\mathbf{\Omega}_M$ is an $L \times L$ positive semidefinite weighting matrix. Assume $\mathbf{\Omega}_M \xrightarrow{p} \mathbf{\Omega}_0$, an $L \times L$ positive definite matrix. Define the $L \times P$ matrix $\mathbf{G}_0 = E[\nabla_\theta g(X_m, \theta_0)]$. Under some mild regularity conditions, Pakes and Pollard (1989) and McFadden (1989) show that:

$$\sqrt{M}(\hat{\theta} - \theta_0) \xrightarrow{d} \text{Normal}(\mathbf{0}, (1 + R^{-1}) * \mathbf{A}_0^{-1} \mathbf{B}_0 \mathbf{A}_0^{-1}) \tag{5}$$

where R is the number of simulations, $\mathbf{A}_0 \equiv \mathbf{G}_0' \mathbf{\Omega}_0 \mathbf{G}_0$, $\mathbf{B}_0 = \mathbf{G}_0' \mathbf{\Omega}_0 \mathbf{\Lambda}_0 \mathbf{\Omega}_0 \mathbf{G}_0$, and $\mathbf{\Lambda}_0 = E[g(X_m, \theta_0) g(X_m, \theta_0)']$ =Var$[g(X_m, \theta_0)]$. If a consistent estimator of $\mathbf{\Lambda}_0^{-1}$ is used as the weighting matrix, the MSM estimator $\hat{\theta}$ is asymptotically efficient, with its asymptotic variance being $\text{Avar}(\hat{\theta}) = (1 + R^{-1}) * (\mathbf{G}_0' \mathbf{\Lambda}_0^{-1} \mathbf{G}_0)^{-1} / M$.

The obstacle in using this standard MSM method is that the moment functions $g(X_m, \cdot)$ are no longer independent across markets when the chain effect induces spatial correlation in the equilibrium outcome. For example, Wal-Mart's entry decision in Benton County, Arkansas directly relates to its entry decision in Carroll County, Arkansas, Benton's neighbor. In fact, any two entry decisions, $D_{i,m}$ and $D_{i,l}$, are correlated because of the chain effect, although the dependence becomes very weak when market m and market l are far apart, since the benefit $\frac{D_{i,l}}{Z_{ml}}$ evaporates with distance.

The MSM estimator remains consistent with such dependent data, but the covariance matrix needs to be corrected. In particular, the asymptotic covariance matrix of the moment functions $\mathbf{\Lambda}_0$ in equation (5) should be replaced by $\mathbf{\Lambda}_0^d = \Sigma_{s \in M} E[g(X_m, \theta_0) g(X_s, \theta_0)']$. Conley

(1999) proposes a nonparametric covariance matrix estimator formed by taking a weighted average of spatial autocovariance terms, with zero weights for observations farther than a certain distance. The method requires the underlying data generating process to satisfy a mixing condition that the dependence among observations dies away quickly as the distance increases. Following Conley (1999) and Conley & Ligon (2002), the estimator of Λ_0^d is:

$$\hat{\Lambda} \equiv \frac{1}{M} \Sigma_m \Sigma_{s \in B_m} \left[\hat{g}(X_m, \theta) \hat{g}(X_s, \theta)' \right] \tag{6}$$

where B_m is the set of markets whose centroid is within fifty miles of market m, including market m. I have also estimated the variance of the moment functions $\hat{\Lambda}$ summing over markets within a hundred miles. All of the parameters that are significant with the smaller set of B_m remain significant, and the changes in the t-statistics are very small.

The estimation procedure is as follows. Start from some initial guess of the parameter values, and draw from the normal distribution four independent vectors: a vector of the market level errors $\{\varepsilon_m\}_{m=1}^M$ and three vectors of firm-specific errors $\{\eta_{k,m}\}_{m=1}^M$, $\{\eta_{w,m}\}_{m=1}^M$, and $\{\eta_{s,m}\}_{m=1}^M$. Obtain the simulated profits $\hat{\Pi}_i, i = k, w, s$ and solve for $\hat{D}_k, \hat{D}_w, \hat{N}_s$. Repeat the simulation R times and formulate $\hat{g}(X_m, \theta)$. Search for parameter values that minimize the objective function (4), while using the same set of simulation draws for all values of θ. To implement the two-step efficient estimator, I use the identity weighting matrix to find a preliminary estimate $\tilde{\theta}$, which is then substituted in equation (6) to compute the optimal weight matrix $\hat{\Lambda}^{-1}$ for the second step.

Instead of the usual machine-generated pseudo-random draws, I use Halton draws, which have better coverage properties and smaller simulation variances.[20] According to Train (2000), 100 Halton draws achieves greater accuracy in his mixed logit estimation than 1000 pseudo-random draws. The parameter estimation exploits 150 Halton simulation draws while the variance is calculated with 300 Halton draws.

There are twenty-six parameters with the following set of moments that match the model-predicted and the observed values of a) numbers of Kmart stores, Wal-Mart stores, and small

[20] A Halton sequence is defined in terms of a given number, usually a prime. As an illustration, consider the prime 3. Divide the unit interval evenly into three segments. The first two terms in the Halton sequence are the two break points: $\frac{1}{3}$ and $\frac{2}{3}$. Then divide each of these three segments into thirds, and add the break points for these segments into the sequence in a particular way: $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}$. Note that the lower break points in all three segments $(\frac{1}{9}, \frac{4}{9}, \frac{7}{9})$ are entered in the sequence before the higher break points $(\frac{2}{9}, \frac{5}{9}, \frac{8}{9})$. Then each of the 9 segments is divided into thirds, and the break points are added to the sequence: $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \frac{10}{27}, \frac{19}{27}, \frac{4}{27}, \frac{13}{27}, \frac{22}{27}$, and so on. This process is continued for as many points as the researcher wants to obtain. See chapter 9 of "Discrete Choice Methods with Simulation (2003)" by Kenneth Train for an excellent discussion of the Halton draws.

firms; b) various kinds of market structures (for example, only a Wal-Mart store but no Kmart stores); c) the number of chain stores in the nearby markets; and d) the interaction between the market size variables and the above items.

6.2 Discussion: a closer look at the assumptions

Now I discuss several assumptions of the model: the game's information structure and issues of multiple equilibria, the symmetry assumption for small firms, and the non-negativity of the chain effect.

6.2.1 Information structure and multiple equilibria

In the empirical entry literature, a common approach is to assume complete information and simultaneous entry. One problem with this approach is the presence of multiple equilibria, which has posed considerable challenges to estimation. Some researchers look for features that are common among different equilibria. For example, Bresnahan and Reiss (1990 and 1991) and Berry (1992) point out that although firm identities differ across different equilibria, the number of entering firms might be unique. Grouping different equilibria by their common features leads to a loss of information and less efficient estimates. Further, common features are increasingly difficult to find when the model becomes more realistic.[21] Others give up point identification of parameters and search for bounds, as in Andrews, Berry and Jia (2004), Chernozhukov, Hong, and Tamer (2004), Pakes, Porter, Ho, and Ishii (2005), and Shaikh (2005). However, a meaningful bound might be difficult to obtain in complicated models as the one employed here, which involves three sets of profit functions with twenty-six parameters.

Given the above considerations, I choose an equilibrium that seems reasonable a priori. In the baseline specification, I estimate the model using the equilibrium that is most profitable for Kmart because Kmart derives from an older entity and historically might have had a first-mover advantage. As a robustness check, I experiment with two other cases. The first one chooses the equilibrium that is most profitable for Wal-Mart. This is the direct opposite of the baseline specification and is inspired by the hindsight of Wal-Mart's success. The second one selects the equilibrium that is most profitable for Wal-Mart in the south and most profitable for Kmart in the rest of the country. This is based on the observation that the northern regions had been Kmart's backyard until recently while Wal-Mart started its business from the south and has expertise in serving the southern population. The estimated parameters for

[21] For example, the number of entering firms is no longer unique in the current example with asymmetric competition effects.

the different cases are very similar to one another, which provides evidence that the results are robust to the equilibrium choice.

6.2.2 The symmetry assumption for small firms

I have assumed that all small firms are symmetric with the same profit function. The assumption is necessitated by data availability, since I do not observe any firm characteristics for small firms. Making this assumption greatly simplifies the complexity of the model with asymmetric competition effects, as it guarantees that in the second stage the equilibrium number of small firms in each market is unique.

6.2.3 The chain effect δ_{ii}

The assumption that $\delta_{ii} \geq 0, i \in \{k, w\}$ is crucial to the solution algorithm, since it implies that the function $V(D)$ defined by the necessary condition (3) is increasing, and that the profit function (1) is supermodular in chain i's own strategy. These results allow me to employ two powerful theorems – Tarski's fixed point theorem and Topkis's theorem – to solve a complicated problem that is otherwise unmanageable. The parameter δ_{ii} does not have to be a constant. It can be region specific, or it can vary with the size of each market (for example, interacting with population), as long as it is weakly positive. However, the algorithm breaks down if either δ_{kk} or δ_{ww} becomes negative, and it excludes scenarios where the chain effect is positive in some regions and negative in others.

The discussion so far has focused on the beneficial aspect of locating stores close to each other. In practice, stores begin to compete for consumers when they get too close. As a result, chains face two opposing forces when making location choices: the chain effect and the business stealing effect. It is conceivable that in some areas stores are so close that the business stealing effect outweighs the gains and δ_{ii} becomes negative.

Holmes (2005) estimates that for places with a population density of 20,000 people per five-mile radius (which is comparable to an average city in my sample counties), 89% of the average consumers visits a Wal-Mart right near by.[22] When the distance increases to 5 miles, 44% of the consumers visits the store. The percentage drops to 7% if the store is 10 miles away. Survey studies also show that few consumers drive further than 10-15 miles for general merchandise shopping. In my sample, the median distance to the nearest store is 21 miles for Wal-Mart stores, and 27 miles for Kmart stores. It seems reasonable to think that the business stealing effect, if it exists, is small.

[22] This is the result from a simulation exercise where the distance is set to 0 mile.

7 Results

7.1 Parameter estimates

The sample includes 2065 small- and medium-sized counties with populations between 5,000 and 64,000. Even though I do not model Kmart's and Wal-Mart's entry decisions in other counties, I incorporate into the profit function the spillover from stores outside the sample. This is especially important for Wal-Mart, as the number of Wal-Mart stores in big counties doubled over the sample period. Table 1 (C) displays the summary statistics of the distance weighted numbers of adjacent Kmart stores $\Sigma_{l \neq m, l \in B_m} \frac{D_{k,l}}{Z_{ml}}$ and Wal-Mart stores $\Sigma_{l \neq m, l \in B_m} \frac{D_{w,l}}{Z_{ml}}$, which measure the spillover from nearby stores (including stores outside the sample). In 1997, the Kmart spillover variable remained roughly the same as in 1988, but the Wal-Mart spillover variable was almost twice as big as in 1988.

The profit functions of all retailers share three common explanatory variables: log of population, log of real retail sales per capita, and the percentage of population that is urban. Many studies have found a pure size effect: there tend to be more stores in a market as the population increases. Retail sales per capita capture the "depth" of a market and explain firm entry behavior better than personal income does. The percentage of urban population measures the degree of urbanization. It is generally believed that urbanized areas have more shopping districts that attract big chain stores.

For Kmart, the profit function includes a dummy variable for the Midwest regions. Kmart's headquarters are located in Troy, Michigan. Until the mid 1980s, this region had always been the "backyard" of Kmart stores. Similarly, Wal-Mart's profit function includes a southern dummy, as well as the log of distance in miles to its headquarters in Benton, Arkansas. This distance variable turns out to be a useful predictor for Wal-Mart stores' location choices. For small firms, everything else equal, there are more small firms in the southern states. It could be that there have always been fewer big retail stores in the southern regions and that people rely on neighborhood small firms for day-to-day shopping.

All coefficients (the market size coefficients β_i, the competitive effects δ_{ij}, and the chain effect δ_{ii}) are allowed to be firm-specific. Before presenting the coefficients from the full model in Table 3, I report the probit coefficients (for Kmart and Wal-Mart) and the ordered probit coefficients (for small retailers) in Table 2. There are six probit and ordered probit regressions, one for each player in each year. The market size coefficients have the same signs as those from the full model, but the estimates for the competition effects are biased toward zero and those for the chain effects have the wrong signs. In subsection 7.3, I show that the ordered probit model underestimates the chain stores' competition effect on small retailers by more than a half.

Table 3 lists the parameter estimates from the full model for 1988 and 1997. In this subsection I focus on the β's; in the next one I discuss δ_{ij} and δ_{ii} in detail. Coefficients for market size variables are highly significant and intuitive, with the exception of the urban variable in the small firms' profit function, which suggests fewer small firms locate in more urbanized areas. ρ is much smaller than 1, indicating the importance of the market level error terms and the necessity of controlling for endogeneity of all firms' entry decisions.

The model is estimated three times, each time with a different equilibrium. Tables 4 (A) and 4 (B) present the three sets of estimates for 1988 and 1997, respectively. In both tables, column one corresponds to the equilibrium most preferred by Kmart; column two uses the equilibrium most preferred by Wal-Mart; column three chooses the one that grants Wal-Mart an advantage in the southern regions and Kmart an advantage in the rest of the country. The estimates are very similar across the different equilibria.

Tables 5(A) and 5(B) display the model's goodness of fit. In Table 5(A), columns one and three display the sample averages, while the other two columns list the model's predicted averages. The model matches exactly the actual average numbers of Kmart and Wal-Mart stores for 1988, and comes very close to them for 1997. The number of small firms is a noisy variable and is much harder to predict. Its sample median was 3, but the maximum was 25 in 1988 and 19 in 1997. The model does a decent job of fitting the data: the sample average was 3.86 per county in 1988 and 3.49 per county in 1997; the model's predictions are 3.79 and 3.43, respectively. Such results might be expected as the parameters are chosen to match these moments. In Table 5(B), I report the correlations between the predicted and observed numbers of Kmart stores, Wal-Mart stores, and small firms in each market. The correlations are between 0.63 and 0.75. These correlations are not included in the set of moment functions, and a high correlation indicates a good fit. Overall, the model explains the data well.

To check whether the estimates are reasonable, Table 6 lists the model's predicted profits and compares them with the accounting profits documented in Kmart's and Wal-Mart's SEC 10-K annual reports. According to the model, the average profit of Wal-Mart stores grew 51% over the sample period, which is consistent with the recorded increase of 41% in Wal-Mart's annual reports. Kmart's accounting profit in 1997 was substantially smaller than that in 1988, due to the financial obligations of divesting several specialized retailing businesses that were overall a financial disappointment. The average sales per Kmart store, in real terms, increased 2.6% over the ten-year period. Considering the various increases in the operating costs documented in Kmart's annual report, the change in its store sales revenue is compatible with the 8% decrease in the store profit predicted by the model.[23,24]

To understand the magnitudes of the market size coefficients, I report in Table 7 the changes in the number of each type of stores when some market size variable changes. For example, to derive the effect of population change on the number of small firms, I fix Kmart's and Wal-Mart's profits, increase small retailers' profit in accordance with a ten percent increase in population, and re-solve for the new equilibrium number of small stores.

The market size variables have a relatively modest impact on the number of small businesses. In 1988, a 10% increase in population attracts 8.7% more firms. The same increase in real retail sales per capita draws 5.4% more firms. The number of small firms declines by about 0.7% when the percentage of urban population goes up by 10%. In comparison, the regional dummy is much more important: everything else equal, changing the southern dummy from 1 for all counties to 0 for all counties leads to 37.2% fewer small firms (6136 small stores vs. 9775 small stores).

Market size variables seem to matter more for big chains. In 1988, A 10% growth in population induces Kmart to enter 12.1% more markets and Wal-Mart 10.3% more markets. A similar increment in retail sales attracts entry of Kmart and Wal-Mart stores in 17.6% and 10.5% more markets, respectively. The results are similar for 1997. These differences indicate that Kmart is much more likely to locate in bigger markets, while Wal-Mart thrives in smaller markets. Perhaps not surprisingly, the regional advantage is substantial for both chains: controlling for the market size, changing the Midwest regional dummy from 1 to 0 for all counties leads to 28% fewer Kmart stores, and changing the Southern regional dummy from 1 to 0 for all counties leads to 52.5% fewer Wal-Mart stores. When distance increases by 10%, the number of Wal-Mart stores drops by 8.4%. Wal-Mart's "home advantage" is much smaller in 1997: everything else the same, changing the south dummy from 1 to 0 for all counties leads to 22% fewer Wal-Mart stores. The regional dummies and the distance variable provide

[23] The model predicted profit for Kmart (Wal-Mart) is the average of the equilibrium profit over all Kmart (Wal-Mart) stores. There are two caveats in the comparison of the profit growth predicted by the model and the reported growth of the accounting profit. First, there is no real unit for the profit derived from the structural profit function – it is scaled by one standard deviation of the unobserved error term in the profit function. The 51% increase in the model predicted profit from 1988 to 1997 assumes that the standard deviation of the error term did not change over this sample period. Second, the 41% growth in the accounting profit is averaged over all stores, including stores in counties outside the sample. The calculation here assumes that profit growth is roughly the same for stores in the sample counties and stores outside the sample counties.

[24] Wal-Mart's 1988 and 1997 annual reports do not separate the profit of Wal-Mart stores from the profit of Sam's clubs. Since the gross markup in Sam's clubs is half of that in the regular Wal-Mart discount stores, two dollars of sales from a Sam's club are assumed to contribute to the total profit the same as one dollar of sales from a Wal-Mart discount store.

a reduced-form way for the static model to capture the path-dependence of the expansion of Wal-Mart stores.

7.2 The competition effect and the chain effect

As shown in Table 3, all of the competition effects and the chain effects, with the exception of the impact of small firms on chain stores, are precisely estimated.

The estimates display several noticeable features. First, the negative impact of Kmart on Wal-Mart's profit δ_{wk} in absolute value is much smaller in 1997 than in 1988, while the opposite is true for Wal-Mart's impact on Kmart's profit δ_{kw}.[25] Both a Cournot model and a Bertrand model with differentiated products predict that reduction in rivals' marginal costs drives down a firm's own profit. I do not observe firms' marginal costs, but these parameter estimates are consistent with evidence that Wal-Mart's marginal cost was declining relative to Kmart's over the sample period. Wal-Mart is famous for its cost-sensitive culture; it is also keen on technology advancement. Holmes (2001) cites evidence that Wal-Mart has been a leading investor in information technology. In contrast, Kmart struggled with its management failures that resulted in stagnant revenue sales, and it either delayed or abandoned store renovation plans throughout the 1990s.

Second, it is somewhat surprising that the negative impact of Kmart on small firms' profit δ_{sk} is comparable to Wal-Mart's impact δ_{sw}, considering the controversies and media reports generated by Wal-Mart. The outcry about Wal-Mart was probably because Wal-Mart had more stores in small- to medium-sized markets where the effect of a big store entry was felt more acutely, and because Wal-Mart kept expanding, while Kmart was consolidating its existing stores with few net openings in these markets over the sample period.

Third, the coefficient for Wal-Mart's chain effect δ_{ww} is smaller in 1997, although the overall effect is bigger, as $\Sigma_{l \neq m, l \in B_m} \frac{D_{w,l}}{Z_{ml}}$ is almost twice as large in 1997 as in 1988.[26] The decline in δ_{ww} suggests that the benefit of scale economies does not grow proportionally. In fact there are good reasons to believe it might not be monotone because, as discussed in section 6.2.3, when chains grow bigger and saturate the area, cannibalization among stores becomes a stronger concern.

To better assess the magnitude of the competition effects, Table 8(A) and Table 8(B) resolve the model for different market structures. The results from Table 8(A) suggest that chains have a substantial competition impact on small firms. In 1988, compared with the scenario where there are neither Kmart nor Wal-Mart stores, adding a Kmart store to each

[25] Due to the high variance of the estimates, the difference is not statistically significant.

[26] The difference between δ_{ww} in 1988 and δ_{ww} in 1997 is not statistically significant.

market reduces the number of small firms by 46%, or 2.69 firms per county; adding a Wal-Mart store reduces it by 43.3%, or 2.53 firms per county. When both a Kmart and a Wal-Mart store enter, the number of small firms plummets by 71.4%, a reduction of 4.17 firms per county. If Wal-Mart takes over Kmart, the number of small firms is 12.3% higher than that observed in the sample when Wal-Mart and Kmart compete against each other.[27] This is due to a business stealing effect between the competing chains – when the two chains merge or one chain takes over the other, the joint-profit-maximizing number of chain stores is smaller, which in turn leads to a larger number of small firms. The patterns are quite similar in 1997: compared with the case of no chain stores, adding a Kmart store to each market decreases the number of small firms by 36.2%, or 1.92 per county; adding a Wal-Mart, 37%, or 1.96 per county; adding both a Kmart and a Wal-Mart store, 61.6%, or 3.27 per county.

Even with the conservative estimate that one Kmart or Wal-Mart store displaces 40% of the small firms, the competition effect of chains on small retailers is sizable, especially since the small discount firms form only a segment of the retailers affected by the entry of chain stores. The combined effect on all small retailers and local communities in general can be much larger.

Table 8 (B) illustrates the competition effect between Kmart and Wal-Mart. Consistent with the changes in δ_{kw} and δ_{wk} from 1988 to 1997, the effect of Kmart's presence on Wal-Mart's profit is much stronger in 1988, while the effect of Wal-Mart's presence on Kmart's profit is stronger in 1997. For example, in 1988, Wal-Mart would only enter 318 markets if there were a Kmart store in every county. When Kmart ceases to exist as a competitor, the number of markets with Wal-Mart stores rises to 846, a net increase of 166%. The same experiment in 1997 leads Wal-Mart to enter 39.3% more markets, from 728 to 1014. The pattern is reversed for Kmart. In 1988, Kmart would enter 42.6% more markets when there is no Wal-Mart stores compared with the case of one Wal-Mart store in every county (479 Kmart stores vs. 336 Kmart stores); in 1997, Kmart would enter 87.1% more markets for the same experiment (610 Kmart stores vs. 326 Kmart stores).[28]

To examine the importance of the chain effect for both chains, consider Table 8 (C). The first row reports the percentage of store profit due to the chain effect; this is the average of $\delta_{ii}\Sigma_{l\neq m,l\in B_m}\frac{D_{i,l}}{Z_{ml}}$ divided by the average store profit (reported in Table 6). For both chains, the chain effect contributes to more than 10% of the store profit. For example, it accounts for

[27] In this counter-factual exercise, Wal-Mart becomes the monopoly chain and competes with small retailers. The total number of chain stores is just the total number of Wal-Mart stores in the new equilibrium, and is smaller than the sum of Wal-Mart stores and Kmart stores before the take-over.

[28] In solving for the number of Wal-Mart (Kmart) stores when Kmart (Wal-Mart) exits, I allow the small firms to compete with the remaining chain.

10.2% of Wal-Mart's profit in 1988, and 12.3% of its profit in 1997. To derive the equilibrium number of stores when there is no chain effect, I set $\delta_{ii} = 0$ for the targeted chain, but keep the rival's δ_{jj} unchanged and re-solve the model. In 1988, without the chain effect, the number of Kmart stores would have decreased by 40, and Wal-Mart would have entered 125 fewer markets. The numbers are comparable for 1997. This result is consistent with Holmes (2005), who also found scale economies to be important. Given the magnitude of these spillover effects, further research that explains their mechanism will help improve our understanding of the retail industry, in particular its productivity gains over the past several decades.[29]

7.3 The impact of Wal-Mart's expansion and related policy issues

Consistent with media reports about Wal-Mart's impact on small retailers, the model predicts that Wal-Mart's expansion contributes to a large percentage of the net decline in the number of small firms over the sample period. The first row in Table 9 (A) records the net decrease of 748 small firms observed over the sample period, or 0.36 per market. To evaluate the impact of Wal-Mart's expansion on small firms separately from other factors (e.g., the change in market sizes or the change in Kmart stores), I re-solve the model using the 1988 coefficients for Kmart's and small firms' profit functions and the 1988 market size variables, but the 1997 coefficients for Wal-Mart's profit function. The experiment corresponds to holding everything the same as in 1988, but allowing Wal-Mart to become more efficient and expand. The predicted number of small firms falls by 558 from the model prediction using the 1988 coefficients for Wal-Mart's profit function. This accounts for 75% of the observed decrease in the number of small firms. Conducting the same experiment but using the 1997 coefficients for Kmart's and small firms' profit functions, the 1997 market size variables, and the 1988 coefficients for Wal-Mart's profit function, I find that Wal-Mart's expansion accounts for 383 stores, or 51% of the observed decrease in the number of small firms.

If we ignore the endogeneity of chains' entry decisions and regress the number of small firms on the number of chains together with the market size variables, we would underestimate the impact of Wal-Mart's expansion on small retailers by a large amount. For example, using the coefficients from an ordered probit model applied to the 1988 data, the difference between the expected number of small firms using Wal-Mart's 1988 store number and the expected number of small firms using Wal-Mart's 1997 store number explains only 33% of the observed decline in the number of small firms. Using the coefficients from the same ordered probit model applied to the 1997 data, Wal-Mart's expansion between 1988 and 1997 accounts for only 20% of the

[29] See Foster *et al* (2002).

observed decline in the number of small firms.[30] Overall, ignoring the endogeneity of chains' entry decisions underestimates the competition effect by fifty to sixty percent.

Using the conservative figure of 383 stores, the absolute impact of Wal-Mart's entry seems modest. However, the exercise here includes only small firms in the discount sector. Both Kmart and Wal-Mart carry a large assortment of products and compete with a variety of stores, like hardware stores, houseware stores, apparel stores, etc., so that their impact on local communities is conceivably much larger. To examine the overall impact of Wal-Mart's expansion, one needs to include a separate profit function for firms in each of these other categories and estimate the system of profit functions jointly.

Government subsidy has long been a policy instrument to encourage firm investment and to create jobs. To evaluate the effectiveness of this policy in the discount retailing sector, I simulate the equilibrium numbers of stores when various firms are subsidized. The results in Table 10 indicate that direct subsidies do not seem to be effective in generating jobs. In 1988, subsidizing Wal-Mart stores 10% of their average profit, which amounts to one million dollars, increases the number of Wal-Mart stores per county only from 0.31 to 0.34.[31] With the average Wal-Mart store hiring fewer than 300 full and part-time employees, the additional number of stores translates to at most nine new jobs.[32] Similarly, subsidizing all small firms by 100% of their average profit increases their number from 3.79 to 4.61, and generates eight jobs if on average a small firm hires ten employees. Together, these exercises suggest that a direct subsidy should be used with caution if it is designed to increase employment in this industry.

8 Conclusion and future work

I have examined the competition effect of chain stores on small firms and the role of the chain effect in firms' entry decisions. The results support the anecdotal evidence that "big drives out small." On average, entry by either a Kmart or a Wal-Mart store displaces forty to fifty percent of the small discount firms. Wal-Mart's expansion from the late 1980s to the late 1990s explains fifty to seventy percent of the net change in the number of small discount firms. Failure to address the endogeneity of firms' entry decisions would result in underestimating

[30] The ordered probit regressions use the same right-hand side variables as the structural model. See Table 2 for the coefficients from the ordered probit regressions.

[31] The average Wal-Mart store's net income in 1988 is about one million in 2004 dollars (see Table 6). Using a discount rate of 10%, the discounted present value of a store's lifetime profit is about ten million. A subsidy of 10% is equivalent to one million dollars.

[32] The equilibrium numbers of Kmart stores and small firms decrease slightly when Wal-Mart is subsidized, but the implied change in employment is tiny.

this impact by fifty to sixty percent. Furthermore, direct subsidies to either chains or small firms are not likely to be effective in creating jobs and should be used with caution.

These results reinforce the concerns raised by many policy observers regarding the subsidies directed to big retail corporations. Perhaps less obvious is the conclusion that subsidies toward small retailers should also be designed carefully.

Like Holmes (2005), I find that scale economies, as captured by the chain effect, generate substantial benefits. Studying these scale economies in more detail is useful for helping firms exploit such advantages and for guiding merger policies or other regulations that affect chains. A better understanding of the mechanism underlying these spillover effects will also help us to gain insight in the productivity gains in the retail industry over the past several decades.

Finally, the algorithm used in this paper can be applied to many industries where scale economies are important. One application is the airline industry, where the network of flight routes exhibits a type of spillover effect similar to the one described here. For example, adding a route from New York to Boston directly affects profits of flights that either originate from or end in Boston and New York. The tools proposed in this paper can be deployed to extend current models of strategic interaction among airlines to incorporate such network effects. Another possible application is to industries with cost complementarity among different products. The algorithm here is particularly suitable for modeling firms' product choices when the product space is large.

9 Appendix A: definitions and proofs

9.1 Verification of the necessary condition (3)

Let $D^* = \arg\max_{D \in \mathbf{D}} \Pi(D)$. The optimality of D^* implies the following set of necessary conditions:

$$\Pi(D_1^*, ..., D_{m-1}^*, D_m^*, D_{m+1}^*, ..., D_M^*) \geq \ \Pi(D_1^*, ..., D_{m-1}^*, D_m, D_{m+1}^*, ..., D_M^*), \forall m, D_m^* \neq D_m$$

Let $\hat{D} = \{D_1^*, ..., D_{m-1}^*, D_m, D_{m+1}^*, ..., D_M^*\}$. $\Pi(D^*)$ differs from $\Pi(\hat{D})$ in two parts: the profit in market m, and the profit in all other markets through the chain effect:

$$\begin{aligned} \Pi(D^*) - \Pi(\hat{D}) &= (D_m^* - D_m)[X_m + \delta \Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}}] + \\ & \quad \delta \Sigma_{l \neq m} D_l^* (\frac{D_m^*}{Z_{lm}}) - \delta \Sigma_{l \neq m} D_l^* (\frac{D_m}{Z_{lm}}) \\ &= (D_m^* - D_m)[X_m + 2\delta \Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}}] \end{aligned}$$

where $Z_{ml} = Z_{lm}$ due to the symmetry. Since $\Pi(D^*) - \Pi(\hat{D}) \geq 0$ for any $D_m^* \neq D_m$, we have $D_m^* = 1, D_m = 0$ if and only if $X_m + 2\delta \Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}} \geq 0$; and $D_m^* = 0, D_m = 1$ if and only if $X_m + 2\delta \Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}} \leq 0$. Together they imply $D_m^* = 1[X_m + 2\delta \Sigma_{l \neq m} \frac{D_l^*}{Z_{ml}} \geq 0]$.

9.2 The set of fixed points of an increasing function that maps a lattice into itself

All of the definitions in this appendix – the definitions of a lattice, a complete lattice, supermodular functions, increasing differences, as well as induced set ordering – are taken from Topkis (1998). The definition of a lattice involves the concepts of a partially ordered set, a join, and a meet. A partially ordered set is a set $\mathbf{X}$ on which there is a binary relation $\preceq$ that is reflexive, antisymmetric, and transitive. If two elements, X' and X'', of a partially ordered set $\mathbf{X}$ have a least upper bound (greatest lower bound) in $\mathbf{X}$, it is their join (meet) and is denoted $X' \vee X''$ ($X' \wedge X''$).

Definition 3 *A partially ordered set that contains the join and the meet of each pair of its elements is a lattice.*

Definition 4 *A lattice in which each nonempty subset has a supremum and an infimum is complete.*

Any finite lattice is complete. A nonempty complete lattice has a greatest element and a least element. Tarski's fixed point theorem, stated in the main body of the paper as Theorem 1, establishes that the set of fixed points of an increasing function that maps from a lattice into itself is a nonempty complete lattice with a greatest element and a least element.

For a counterexample where a decreasing function's set of fixed points is empty, consider the following simplified entry model where three firms compete with each other and decide simultaneously whether to enter the market. Their joint strategy space is $\mathbf{D} = \{0,1\}^3$. The profit functions are as follows:

$$\begin{cases} \Pi_k = D_k(0.5 - D_w - 0.25D_s) \\ \Pi_w = D_w(1 - 0.5D_k - 1.1D_s) \\ \Pi_s = D_s(0.6 - 0.5D_w - 0.7D_s) \end{cases}$$

Let $D = \{D_k, D_w, D_s\} \in \mathbf{D}$, D_{-i} denote rivals' strategies, $V_i(D_{-i})$ denote the best response function for player i, and $V(D) = \{V_k(D_{-k}), V_w(D_{-w}), V_s(D_{-s})\}$ denote the joint best response function. It is easy to show that $V(D)$ is a decreasing function that takes the following values:

$$\begin{cases} V(0,0,0) = \{1,1,1\};\ V(0,0,1) = \{1,0,1\};\ V(0,1,0) = \{0,1,1\};\ V(0,1,1) = \{0,0,1\} \\ V(1,0,0) = \{1,1,0\};\ V(1,0,1) = \{1,0,0\};\ V(1,1,0) = \{0,1,0\};\ V(1,1,1) = \{0,0,0\} \end{cases}$$

The set of fixed points for $V(D)$ is empty, since there does not exist a $D \in \mathbf{D}$ such that $V(D) = D$.

9.3 A tighter lower bound and upper bound for the optimal solution vector D^*

In section 5.1 I have shown that using $\inf(\mathbf{D})$ and $\sup(\mathbf{D})$ as starting points yields, respectively, a lower bound and an upper bound to $D^* = \arg\max_{D\in\mathbf{D}} \Pi(D)$. Here I introduce two bounds that are tighter. The lower bound builds on the solution to a constrained maximization problem:

$$\begin{aligned} \max_{D_1,\ldots,D_M \in \{0,1\}} \Pi &= \sum_{i=1}^{M} \left[D_m * (X_m + \delta \Sigma_{l\neq m} \frac{D_l}{Z_{ml}}) \right] \\ s.t. \text{ if } D_m &= 1, \text{ then } X_m + \delta \Sigma_{l\neq m} \frac{D_l}{Z_{ml}} > 0 \end{aligned}$$

The solution to this constrained maximization problem belongs to the set of fixed points of the vector function $\hat{V}(D) = \{\hat{V}_1(D), ..., \hat{V}_M(D)\}$, where $\hat{V}_m(D) = 1[X_m + \delta \Sigma_{l\neq m} \frac{D_l}{Z_{ml}} > 0]$. The function $\hat{V}(\cdot)$ is increasing and maps from $\mathbf{D}$ into itself: $V : \mathbf{D} \to \mathbf{D}$. Using arguments similar to those in section 5.1, one can show that the convergent vector $\hat{D}$ using $\sup(\mathbf{D})$ as the starting vector is the greatest element of the set of fixed points. Further, $\hat{D}$ achieves a higher profit than any other fixed point of $\hat{V}(\cdot)$, since by construction each non-zero element of the vector $\hat{D}$ adds to the total profit. Changing any non-zero element(s) of $\hat{D}$ to zero reduces the total profit.

To show that $\hat{D} \leq D^*$, the solution to the original unconstrained maximization problem, we construct a contradiction. Since the maximum of an unconstrained problem is always greater than that of a corresponding constrained problem, we have: $\Pi(D^*) \geq \Pi(\hat{D})$. Therefore, D^* can't be strictly smaller than $\hat{D}$, because any vector strictly smaller than $\hat{D}$ delivers a lower profit. Suppose D^* and $\hat{D}$ are unordered. Let $D^{**} = D^* \vee \hat{D}$ (where "$\vee$" defines the element-by-element Max operation). The change from D^* to D^{**} increases total profit, because profits at markets with $D^*_m = 1$ do not decrease after the change, and profits at markets with $D^*_m = 0$ but $\hat{D}_m = 1$ increase from 0 to a positive number after the change. This contradicts the definition of D^*, so $\hat{D} \leq D^*$.

Note that $V(\hat{D}) \geq \hat{D}$, where $V(\cdot)$ is as defined in section (5.1). This follows from $V_m(\hat{D}) = 1[X_m + 2\delta\Sigma_{l \neq m}\frac{\hat{D}_l}{Z_{ml}} \geq 0] \geq 1[X_m + \delta\Sigma_{l \neq m}\frac{\hat{D}_l}{Z_{ml}} \geq 0] = \hat{D}_m, \forall m$, where the last equality holds because $\hat{D}$ is a fixed point of $\hat{V}(\cdot)$. The monotonicity of $V(\cdot)$, together with $\hat{D} \leq D^*$ and $V(\hat{D}) \geq \hat{D}$, implies that the iteration process starting from $\hat{D}$ converges, and that the convergent vector (denoted as $\hat{D}^T$) is a lower bound to D^*.

$\hat{D}^T$ is a tighter lower bound than D^L (discussed in section (5.1)) because $\hat{D} \geq \inf(\mathbf{D})$, so $\hat{D}^T = V^{TT}(\hat{D}) \geq V^{TT}(\inf(\mathbf{D})) = D^L$, with $TT = \max\{T, T'\}$, where T is the number of iterations from $\hat{D}$ to $\hat{D}^T$ and T' is the number of iterations from $\inf(\mathbf{D})$ to D^L.

Since the chain effect is bounded by zero (when there are no other stores anywhere) and $\delta\Sigma_{l \neq m}\frac{1}{Z_{ml}}$ (when there is a store in every market), we can find a tighter upper bound to D^* by starting from the vector $\tilde{D} = \{\tilde{D}_m : \tilde{D}_m = 0 \text{ if } X_m + 2\delta\Sigma_{l \neq m}\frac{1}{Z_{ml}} < 0; \tilde{D}_m = 1 \text{ otherwise}\}$. The markets with $\tilde{D}_m = 0$ contribute a negative element to total profit even with the largest conceivable chain effect, so it is never optimal to enter these markets, i.e., $\tilde{D} \geq D^*$. It is straightforward to show that the iteration converges ($V^T(\tilde{D}) = \tilde{D}^T, T \leq M$) and that the convergent vector $\tilde{D}^T$ is a tighter upper bound to D^* than D^U.[33]

[33] $\tilde{D}^T \leq D^U$ because $\tilde{D}^T = V^{TT}(\tilde{D}) \leq V^{TT}(\sup(\mathbf{D})) = D^U$, with $TT = \max\{T, T'\}$, where T is the number of iterations from $\tilde{D}$ to $\tilde{D}^T$ and T' is the number of iterations from $\sup(\mathbf{D})$ to D^U.

9.4 Verification that the chains' profit functions are supermodular with decreasing differences

Definition 5 *Suppose that $Y(X)$ is a real-valued function on a lattice $\mathbf{X}$. If*

$$Y(X') + Y(X'') \leq Y(X' \vee X'') + Y(X' \wedge X'') \quad (7)$$

for all X' and X'' in $\mathbf{X}$, then $Y(X)$ is supermodular on $\mathbf{X}$. If

$$Y(X') + Y(X'') < Y(X' \vee X'') + Y(X' \wedge X'')$$

for all unordered X' and X'' in $\mathbf{X}$, then $Y(X)$ is strictly supermodular on $\mathbf{X}$. If $-Y(X)$ is (strictly) supermodular, then $Y(X)$ is (strictly) submodular.

Definition 6 *Suppose that $\mathbf{X}$ and K are partially ordered sets and $Y(X,k)$ is a real-valued function on $\mathbf{X} \times K$. If $Y(X,k'') - Y(X,k')$ is increasing, decreasing, strictly increasing, or strictly decreasing in X on $\mathbf{X}$ for all $k' \prec k''$ in K, then $Y(X,k)$ has, respectively, increasing differences, decreasing differences, strictly increasing differences, or strictly decreasing differences in (X,k) on $\mathbf{X}$.*

Now let us verify that chain i's profit function (1) is supermodular in its own strategy $D_i \in \mathbf{D}$. For ease of notation, the firm subscript i is omitted, and $X_m\beta_i + \delta_{ij} D_{j,m} + \delta_{is}\ln(N_{s,m} + 1) + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}$ is absorbed into X_m. Chain i's profit function is simplified to: $\Pi = \sum_{m=1}^{M}\left[D_m * (X_m + \delta\Sigma_{l\neq m}\frac{D_l}{Z_{ml}})\right]$. First it is easy to show that $D' \vee D'' = (D' - \min(D', D'')) + (D'' - \min(D', D'')) + \min(D', D'')$, $D' \wedge D'' = \min(D', D'')$. Let $D' - \min(D', D'')$ be denoted D_1, $D'' - \min(D', D'')$ as D_2, and $\min(D', D'')$ as D_3.The left-hand side of the inequality (7) is:

$$\begin{aligned}
\Pi(D') + \Pi(D'') &= \sum\nolimits_m D'_m(X_m + \delta\Sigma_{l\neq m}\frac{D'_l}{Z_{ml}}) + \sum\nolimits_m D''_m(X_m + \delta\Sigma_{l\neq m}\frac{D''_l}{Z_{ml}}) \\
&= \sum\nolimits_m \left[(D'_m - \min(D'_m, D''_m)) + \min(D'_m, D''_m)\right] * \\
&\quad \left[X_m + \delta\Sigma_{l\neq m}\frac{1}{Z_{ml}}[(D'_l - \min(D'_l, D''_l)) + \min(D'_l, D''_l)]\right] + \\
&\quad \sum\nolimits_m[(D''_m - \min(D'_m, D''_m)) + \min(D'_m, D''_m)] * \\
&\quad \left[X_m + \delta\Sigma_{l\neq m}\frac{1}{Z_{ml}})[(D''_l - \min(D'_l, D''_l)) + \min(D'_l, D''_l)]\right] \\
&= \sum\nolimits_m (D_{1,m} + D_{3,m})(X_m + \delta\Sigma_{l\neq m}\frac{1}{Z_{ml}}(D_{1,l} + D_{3,l})) + \\
&\quad \sum\nolimits_m (D_{2,m} + D_{3,m})(X_m + \delta\Sigma_{l\neq m}\frac{1}{Z_{ml}}(D_{2,l} + D_{3,l}))
\end{aligned}$$

Similarly, the right-hand side of the inequality (7) is:

$$\begin{aligned}
\Pi(D' \vee D'') + \Pi(D' \wedge D'') &= \sum\nolimits_m (D'_m \vee D''_m) \left[X_m + \delta \Sigma_{l \neq m} \frac{1}{Z_{ml}} (D'_l \vee D''_l) \right] + \\
& \sum\nolimits_m (D'_m \wedge D''_m) \left[X_m + \delta \Sigma_{l \neq m} \frac{1}{Z_{ml}} (D'_l \wedge D''_l) \right] \\
&= \sum\nolimits_m (D_{1,m} + D_{2,m} + D_{3,m}) \left[X_m + \delta \Sigma_{l \neq m} \frac{1}{Z_{ml}} (D_{1,l} + D_{2,l} + D_{3,l}) \right] + \\
& \sum\nolimits_m D_{3,m} (X_m + \delta \Sigma_{l \neq m} \frac{1}{Z_{ml}} D_{3,l}) \\
&= \Pi(D') + \Pi(D'') + \delta (\Sigma_m \Sigma_{l \neq m} \frac{D_{2,m} D_{1,l} + D_{1,m} D_{2,l}}{Z_{ml}})
\end{aligned}$$

The profit function is supermodular in its own strategy if the chain effect δ is non-negative. The verification of decreasing differences is also straightforward (here $\delta_{ij} D_{j,m}$ is spelled out, rather than absorbed into X_{im}):

$$\begin{aligned}
\Pi_i(D_i, D''_j) - \Pi_i(D_i, D'_j) &= \sum\nolimits_m \left[D_{i,m} * (X_{im} + \delta_{ii} \Sigma_{l \neq m} \frac{D_{i,l}}{Z_{ml}} + \delta_{ij} D''_{j,m}) \right] - \\
& \sum\nolimits_m \left[D_{i,m} * (X_{im} + \delta_{ii} \Sigma_{l \neq m} \frac{D_{i,l}}{Z_{ml}} + \delta_{ij} D'_{j,m}) \right] \\
&= \delta_{ij} \sum_{m=1}^{M} D_{i,m} (D''_{j,m} - D'_{j,m})
\end{aligned}$$

The difference is decreasing in D_i for all $D'_j < D''_j$ as long as $\delta_{ij} \leq 0$.

9.5 Multiple maximizers to the chains' optimization problem

In the main body of the paper, I have assumed that the optimal solution to the profit maximization problem D^* is unique. To accommodate multiple solutions in the algorithm discussed in subsections 5.1 and 5.2, I need to introduce the definition of the induced set ordering.

Definition 7 *The induced set ordering $\sqsubseteq$ is defined on the collection of nonempty members of the power set $\mathcal{P}(\mathbf{X}) \backslash \{\emptyset\}$ such that $\mathbf{X}' \sqsubseteq \mathbf{X}''$ in $\mathcal{P}(\mathbf{X}) \backslash \{\emptyset\}$ if X' in $\mathbf{X}'$ and X'' in $\mathbf{X}''$ imply that $X' \wedge X''$ is in $\mathbf{X}'$ and $X' \vee X''$ is in $\mathbf{X}''$.*

The power set $\mathcal{P}(\mathbf{X})$ of a set $\mathbf{X}$ is the set of all subsets of $\mathbf{X}$.

Definition 8 *A function whose range is included in the collection of all subsets of some set is a correspondence. A correspondence S_k is increasing (decreasing) in k on K if the domain K is a partially ordered set, the range $\{S_k : k \in K\}$ is in $\mathcal{L}(\mathbf{X})$ where $\mathbf{X}$ is a lattice and $\mathcal{L}(\mathbf{X})$ is a partially ordered set with the ordering relation $\sqsubseteq$, and S_k is an increasing (decreasing) correspondence from K into $\mathcal{L}(\mathbf{X})$ (so $k' \preceq k''$ in K implies $S_{k'} \sqsubseteq S_{k''}$ ($S_{k''} \sqsubseteq S_{k'}$) in $\mathcal{L}(\mathbf{X})$).*

In stating that S_k is increasing (decreasing) in k, it is implicit that each S_k is a nonempty sublattice of $\mathbf{X}$ and that $\sqsubseteq$ is the ordering relation on the sets S_k in $\mathcal{L}(\mathbf{X})$. The following Theorem, discussed in Topkis (1998), states that if S_k is decreasing in k on K and S_k is finite, then S_k has a greatest element and a least element, both of which decrease in k.[34]

Theorem 9 *Suppose that $\mathbf{X}$ is a lattice, K is a partially ordered set, S_k is a subset of $\mathbf{X}$ for each k in K, and S_k is decreasing in k on K. If S_k has a greatest (least) element for each k in K, then the greatest (least) element is a decreasing function of k from K into $\mathbf{X}$. Hence, if S_k is finite for each k in K, or $\mathbf{X}$ is a subset of R^n and S_k is a compact subset of R^n for each k in K, then S_k has a greatest element and a least element for each k in K and the greatest (least) element is a decreasing function of k.*

According to Theorem 2 in the main body of the paper, $\arg\max_{D_i \in \mathbf{D}_i} \Pi_i(D_i, D_j)$ is decreasing in D_j. Since $\arg\max_{D_i \in \mathbf{D}_i} \Pi_i(D_i, D_j) \subset \mathbf{D}_i$ is finite, Theorem 9 implies that the set $\arg\max_{D_i \in \mathbf{D}_i} \Pi_i(D_i, D_j)$ has a greatest element and a least element for each D_j, both of which decease in D_j. The solution algorithm that accommodates multiple solutions to the profit-maximizing problem is as follows: to search for the equilibrium most profitable for Kmart, if there are multiple elements in Kmart's best response correspondence, choose the greatest element; if there are multiple elements in Wal-Mart's best response correspondence, choose the least element. To search for the equilibrium most profitable for Wal-Mart, choose the greatest element in Wal-Mart's best response correspondence and the least element in Kmart's best response.

9.6 Computational issues

The main computational burden of this exercise is the search for the best responses $K(D_w)$ and $W(D_k)$. In section 5.1, I have proposed two bounds D^U and D^L that help to reduce the number of profit evaluations. Appendix 9.3 illustrates a tighter upper and lower bound that work well in the empirical implementation.

When the chain effect δ_{ii} is sufficiently big, it is conceivable that the upper bound and lower bound are far apart from each other. If this happens, computational burden once again becomes an issue, as there will be many vectors between these two bounds.

Two observations work in favor of the algorithm. First, recall that the chain effect is assumed to take place among counties whose centroids are within fifty miles. Markets that are

[34] The original theorem is stated in terms of S_k increasing in k. Replacing k with $-k$ delivers the version of the theorem stated here.

farther away are not directly connected. Conditioning on the entry decisions in other markets, the entry decisions in group A do not depend on the entry decisions in group B if all markets in group A are at least fifty miles away from any market in group B. Therefore, what matters is the size of the largest connected markets different between D^U and D^L, rather than the total number of elements different between D^U and D^L. To illustrate this point, suppose there are ten markets as below:

	1	2	3	
4	5	6	7	8
	9	10		

. Suppose the upper bound D^U and the lower bound D^L are the same for markets 2,6,9, and 10, but differ for the rest six markets: $D^U =$

	1	D_2	1	
1	1	D_6	1	1
	D_9	D_{10}		

, $D^L =$

	0	D_2	0	
0	0	D_6	0	0
	D_9	D_{10}		

. If markets 1, 4, and 5 (group A) are at least fifty miles away from markets 3, 7, and 8 (group B), one only needs to evaluate $2^3 + 2^3 = 16$ vectors, rather than $2^6 = 64$ vectors to find the profit maximizing vector.

The second observation is that even with a sizable chain effect, the event of having D^U and D^L different in a large connected area is extremely unlikely. Let N denote the size of such an area. Both D^U and D^L are the fixed points of $V(\cdot)$, so: $D^U_m = 1[X_m + 2\delta\Sigma_{l\neq m, l\in B_m}\frac{D^U_l}{Z_{ml}} + \xi_m \geq 0]$ and $D^L_m = 1[X_m + 2\delta\Sigma_{l\neq m, l\in B_m}\frac{D^L_l}{Z_{ml}} + \xi_m \geq 0]$, where I have grouped $X_m\beta_i + \delta_{ij}D_{j,m} + \delta_{is}\ln(N_{s,m}+1)$ into X_m, and $\sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}$ into ξ_m. B_m denotes the set of markets that are within fifty miles from market m, including m.) The probability of $D^U_m = 1, D^L_m = 0$ for every market m in the size-N connected area C_N is:

$$\begin{aligned}\Pr(D^U_m &= 1, D^L_m = 0, \forall m \in C_N) \leq \Pr(X_m + \xi_m < 0, X_m + \xi_m + 2\delta\Sigma_{l\neq m, l\in B_m}\frac{1}{Z_{ml}} \geq 0, \forall m \in C_N) \\ &= \Pi^N_{m=1}\Pr(X_m + \xi_m < 0, X_m + \xi_m + 2\delta\Sigma_{l\neq m, l\in B_m}\frac{1}{Z_{ml}} \geq 0)\end{aligned}$$

where $\Pi^N_{m=1}$ denotes the product of the N elements. The equality follows from the i.i.d. assumption of $X_m + \xi_m$. As δ goes to infinity, the probability approaches $\Pi^N_{m=1}\Pr(X_m + \xi_m < 0)$ from below. How fast it decreases when N increases depends on the distribution of ξ_m as well as the distribution of X_m. If ξ_m is i.i.d. normally distributed and X_m is linearly distributed between $[-a, a]$, with a a finite positive number, on average the probability is on the magnitude of $(\frac{1}{2})^N$.

To show this, note that:

$$\begin{aligned} E(\Pi_{m=1}^{N} \Pr(X_m + \xi_m &< 0)) = E(\Pi_{m=1}^{N}(1 - \Phi(X_m)) \\ &= \Pi_{m=1}^{N}[1 - E(\Phi(X_m))] \\ &= (\frac{1}{2})^N \end{aligned}$$

Therefore, even in the worst scenario that the chain effect δ approaches infinity, the probability of having a large connected area that differs between D^U and D^L decreases exponentially with the size of the area. In the current application, the size of the largest connected area that differs between D^L and D^U is seldom bigger than seven or eight markets.

10 Appendix B: data

I went through all the painstaking details to clean the data from the Directory of Discount Stores. After the manually entered data were inspected many times with the hard copy, the stores' cities were matched to belonging counties using a census data.[35] Some city names listed in the directory contained typos, so I first found possible spellings using the census data, then inspected the stores' street addresses and zipcodes using various web sources to confirm the right city name spelling. The final data set appears to be quite accurate. I compared it with Wal-Mart's firm data and found the difference to be quite small.[36] For the sample counties, only thirty to sixty stores were not matched between these two sources for either 1988 or 1997.

[35] Marie Pees kindly provided these data.

[36] I am very grateful to Emek Basker for sharing the Wal-Mart firm data with me.

References

[1] Andrews, Donald W.K., Steven Berry, and Panle Jia (2004), "Confidence Regions for Parameters in Discrete Games with Multiple Equilibria," *Yale University working paper.*

[2] Bajari, Patrick, and Jeremy Fox (2005), "Complementarities and Collusion in an FCC Spectrum Auction," *working paper.*

[3] Bajari, Patrick, Han Hong, and Stephen Ryan (2004), "Identification and Estimation of Discrete Games of Complete Information," *Duke University working paper.*

[4] Basker, Emek (2005a), "Job Creation or Destruction? Labor-Market Effects of Wal-Mart Expansion," *The Review of Economics and Statistics,* vol. 87, No. 1, pp. 174-183.

[5] Basker, Emek (2005b), "Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail Prices," *Journal of Urban Economics*, Vol. 58, No. 2, pp. 203-229.

[6] Berry, Steven (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, vol 60, No. 4, pp. 889-917.

[7] Bresnahan, Timothy and Peter Reiss (1990), "Entry into Monopoly Markets," *Review of Economic Studies*, vol. 57, No. 4, pp. 531-553.

[8] Bresnahan, Timothy and Peter Reiss (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, vol. 95, No. 5, pp. 57-81.

[9] Chernozhukov, Victor, Han Hong, and Elie Tamer (2004), "Inference on Parameter Set in Econometric Models," *Princeton University working paper.*

[10] Committee on Small Business. House (1994), "Impact of Discount Superstores on Small Business and Local Communities," Committee Serial No. 103-99. Congressional Information Services, Inc.

[11] Conley, Timothy (1999), "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, Vol. 92, pp. 1-45.

[12] Conley, Timothy, and Ethan Ligon (2002), "Economic Distance and Cross Country Spillovers", *Journal of Economic Growth*, Vol. 7, pp. 157-187.

[13] Davis, Peter (2005) "Spatial Competition in Retail Markets: Movie Theaters," forthcoming *RAND Journal of Economics.*

[14] Directory of Discount Department Stores (1988-1997), Chain Store Guide, Business Guides, Inc., New York.

[15] Discount Merchandiser (1988-1997), Schwartz Publications, New York.

[16] Foster, Lucia, John Haltiwanger, and C.J. Krizan (2002), "The Link Between Aggregate and Micro Productivity Growth: Evidence from Retail Trade," *NBER working paper*, No. 9120.

[17] Gowrisankaran, Gautam, and Joanna Stavins (2004), "Network Externalities and Technology Adoption: Lessons from Electronic Payments," *Rand Journal of Economics*, Vol. 35, No. 2, pp 260-276.

[18] Hausman, Jerry, and Ephraim Leibtag (2005), "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart," *NBER working paper,* No. 11809.

[19] Holmes, Thomas (2001), "Barcodes Lead to Frequent Deliveries and Superstores," *RAND Journal of Economics,* Vol. 32, No. 4, pp. 708-725.

[20] Holmes, Thomas (2005), "The Diffusion of Wal-Mart and Economies of Density," *University of Minnesota working paper.*

[21] Kmart Inc. (1988-2000), *Annual Report.*

[22] Mazzeo, Michael (2002), "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics,* vol. 33, No. 2, pp. 1-22.

[23] McFadden, Daniel (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, vol. 57, No. 5, pp. 995-1026.

[24] Neumark, David, Junfu Zhang, and Stephen Ciccarella (2005), "The Effects of Wal-Mart on Local Labor Markets," *NBER working paper*, No. 11782.

[25] Pakes, Ariel and David Pollard, "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, vol. 57, No. 5, pp. 1027-1057.

[26] Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii (2005), "Moment Inequalities and Their Application," *Harvard working paper.*

[27] Pinkse, Joris, Margaret Slade, and Craig Brett (2002), "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, Vol. 70, No. 3, pp. 1111-1153.

[28] Seim, Katja (2005), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," forthcoming *RAND Journal of Economics*.

[29] Shaikh, Azeem (2005), "Inference for Partially Identified Econometric Models," *Stanford University working paper.*

[30] Shils, Edward B. (1997), "The Shils Report: Measuring the Economic and Sociological Impact of the Mega-Retail Discount Chains on Small Enterprise in Urban, Suburban and Rural Communities," online: http://www.lawmall.com/rpa/rpashils.htm.

[31] Smith, Howard (2004), "Supermarket Choice and Supermarket Competition in Market Equilibrium," *Review of Economic Studies,* vol. 71, No. 1, pp. 235-263.

[32] Stone, Keneth (1995), "Impact of Wal-Mart Stores On Iowa Communities: 1983-93," *Economic Development Review*, Vol. 13, No. 2, pp. 60-69.

[33] Tarski, Alfred (1955), "A Lattice-Theoretical Fixpoint Theorem and Its Applications," *Pacific Journal of Mathematics*, Vol 5, pp 285-309.

[34] Taylor, Don, and Jeanne Archer (1994), "*Up Against the Wal-Marts (How Your Business Can Survive in the Shadow of the Retail Giants),*" New York: American Management Association.

[35] Topkis, Donald (1978), "Minimizing a submodular function on a lattice," *Operations Research*, Vol 26, pp305-321.

[36] Topkis, Donald (1979), "Equilibrium Points in Nonzero-Sum n-Person Submodular Games," *SIAM Journal of Control and Optimum*, Vol 17, pp 773-787.

[37] Topkis, Donald (1998), "*Supermodularity and Complementarity,*" Princeton University Press, New Jersey.

[38] Train, Kenneth (2000), "Halton Sequences for Mixed Logit," *UC Berkeley working paper.*

[39] Train, Kenneth (2003), "*Discrete Choice Methods with Simulation,*" Cambridge University Press, Cambridge, UK.

[40] Vance, Sandra S., and Roy V. Scott (1994), "*A History of Sam Walton's Retail Phenomenon,*" Twayne Publishers, New York.

[41] Wal-Mart Stores, Inc. (1970-2000), *Annual Report.*

[42] Zhou, Lin (1994), "The Set of Nash Equilibria of a Supermodular Game Is a Complete Lattice," *Games and Economic Behavior*, Vol. 7, pp 295-300.

Table 1 (A): The Discount Industry from 1960 to 1997

Year	Number of Discount Stores	Total Sales (2004 $bill.)	Average Store Size (thou ft²)	Number of Firms
1960	1329	12.8	38.4	1016
1980	8311	119.4	66.8	584
1989	9406	123.4	66.5	427
1997	9741	198.7	79.2	230

Source: various issues of *Discount Merchandiser*. The numbers include only traditional discount stores. Wholesale clubs, super-centers, and special retailing stores are excluded.

Table 1 (B): Summary Statistics for the Data Set

	1988		1997	
Variable	Mean	Std.	Mean	Std.
Population (thou.)	22.47	14.12	24.27	15.67
Per Capita Retail Sales (1984 $thou.)	3.69	1.44	4.05	2.02
Percentage of Urban Population	0.30	0.23	0.33	0.24
Midwest (1 if in the Great Lakes, Plains, or Rocky Mountain Region)	0.41	0.49	0.41	0.49
South (1 if Southwest or Southeast)	0.50	0.50	0.50	0.50
Distance to Benton, AR (100 miles)	6.14	3.88	6.14	3.88
% of Counties with Kmart Stores	0.21	0.42	0.19	0.41
% of Counties with Wal-Mart Stores	0.32	0.53	0.48	0.57
Number of Firms with 1-19 Employees	3.86	2.84	3.49	2.61
Number of Counties	2065			

Source: the 1988 population is from the U.S. Census Bureau's website; the 1997 population is from the website of the Missouri State Census Data Center. Retail sales are from the 1987 and 1997 Economic Census, respectively. The percentage of urban population is from the 1990 and 2000 decennial census, respectively. Region dummies and the distance variable are from the 1990 census. The numbers of Kmart and Wal-Mart stores are from the Directory of Discount Stores, and the number of small stores is from the County Business Patterns.

Table 1 (C): Summary Statistics for the Distance Weighted Number of Adjacent Stores

	1988		1997	
Variable	Mean	Std.	Mean	Std.
Distance Weighted Number of Adjacent Kmart Stores within 50 Miles	0.11	0.08	0.13	0.11
Distance Weighted Number of Adjacent Wal-Mart Stores within 50 Miles	0.10	0.08	0.19	0.19
Number of Counties	2065			

Source: the annual reference "Directory of Discount Department Stores" by Chain Store Guide, Business Guides, Inc., New York.

Table 2: Parameter Estimates from Probit (Kmart and Wal-Mart) and Ordered Probit (Small Firms)

Kmart's Profit			Wal-Mart's Profit			Small Firms' Profit		
Variable	1988	1997	Variable	1988	1997	Variable	1988	1997
Log Population	1.68*	1.75*	Log Population	1.27*	2.23*	Log Population	1.24*	1.18*
	(0.12)	(0.12)		(0.10)	(0.11)		(0.05)	(0.05)
Log Retail Sales	1.69*	1.64*	Log Retail Sales	1.32*	1.58*	Log Retail Sales	0.76*	0.64*
	(0.16)	(0.14)		(0.14)	(0.13)		(0.07)	(0.06)
Urban Ratio	1.56*	1.20*	Urban Ratio	1.27*	0.99*	Urban Ratio	-1.10*	-0.99*
	(0.25)	(0.24)		(0.21)	(0.21)		(0.12)	(0.12)
Midwest	0.43*	0.33*	Log Distance	-1.63*	-1.11*	South	0.62*	0.78*
	(0.09)	(0.10)		(0.09)	(0.08)		(0.05)	(0.05)
Constant	-20.55*	-20.59*	South	1.08*	0.87*			
	(1.36)	(1.28)		(0.09)	(0.09)			
			Constant	-5.75*	-12.88*			
				(1.11)	(1.08)			
δ_{kw}	-0.21*	-0.66*	δ_{wk}	-0.57*	-0.73*	δ_{sk}	-0.34*	-0.11
	(0.09)	(0.11)		(0.11)	(0.11)		(0.07)	(0.07)
δ_{kk}	-1.38*	-1.15†	δ_{ww}	2.88*	-3.41*	δ_{sw}	-0.32*	-0.21*
	(0.67)	(0.64)		(0.79)	(0.63)		(0.06)	(0.06)
δ_{ks}	-0.17†	0.01	δ_{ws}	-0.20*	-0.34*			
	(0.09)	(0.09)		(0.08)	(0.09)			
Observation Number	2065	2065		2065	2065		2065	2065
Log Likelihood	-583.64	-567.77		-680.25	-669.71		-4226	-4035

Note: * denotes significance at the 5% confidence level, and † denotes significance at the 10% confidence level. Standard errors are in parentheses. The estimated cutoff values in the ordered probit regressions for small firms are omitted. Midwest and South are regional dummies, with the Great Lakes region, the Plains region, and the Rocky Mountain region grouped as the "Midwest", the Southwest region and the Southeast region grouped as the "South". $\delta_{ij},\ i, j \in \{k, w, s\}, i \neq j$, denotes the competition effect, while $\delta_{ii}, i \in \{k, w\}$, denotes the chain effect. "*k*" stands for Kmart, "*w*" stands for Wal-Mart, and "*s*" stands for small stores.

Table 3: Parameter Estimates from the Full Model

Kmart's Profit			Wal-Mart's Profit			Small Firms' Profit		
Variable	1988	1997	Variable	1988	1997	Variable	1988	1997
Log Population	1.49*	1.84*	Log Population	1.54*	2.16*	Log Population	1.65*	1.90*
	(0.09)	(0.13)		(0.15)	(0.15)		(0.18)	(0.26)
Log Retail Sales	2.19*	2.01*	Log Retail Sales	1.56*	1.85*	Log Retail Sales	1.04*	1.17*
	(0.25)	(0.23)		(0.35)	(0.25)		(0.12)	(0.16)
Urban Ratio	2.07*	1.55*	Urban Ratio	2.19*	1.73*	Urban Ratio	-0.46*	-0.78*
	(0.42)	(0.29)		(0.35)	(0.40)		(0.21)	(0.24)
Midwest	0.40*	0.32*	Log Distance	-1.31*	-1.01*	South	0.88*	1.03*
	(0.12)	(0.14)		(0.16)	(0.14)		(0.14)	(0.17)
Constant	-24.60*	-24.08*	South	0.94*	0.61*	Constant	-10.22*	-11.89*
	(2.07)	(2.07)		(0.13)	(0.11)		(0.98)	(1.56)
			Constant	-10.90*	-16.37*			
				(2.98)	(2.17)			
δ_{kw}	-0.48*	-0.93*	δ_{wk}	-1.54*	-1.13*	δ_{sk}	-1.20*	-1.00*
	(0.22)	(0.29)		(0.21)	(0.30)		(0.23)	(0.20)
δ_{kk}	0.63*	0.75*	δ_{ww}	1.22*	0.89*	δ_{sw}	-1.11*	-1.03*
	(0.20)	(0.36)		(0.43)	(0.39)		(0.16)	(0.21)
δ_{ks}	-0.07	-0.02	δ_{ws}	-0.03	-0.03	δ_{ss}	-2.14*	-2.41*
	(0.05)	(0.05)		(0.11)	(0.12)		(0.28)	(0.35)
ρ	0.53*	0.65*						
	(0.11)	(0.10)						
Function Value	62.84	30.80				Observation Number	2065	2065

Note: * denotes significance at the 5% confidence level, and † denotes significance at the 10% confidence level. Standard errors are in parentheses. See Table 2 for the explanation of the variables and parameters. $\sqrt{1-\rho^2}$ measures the importance of the market-level profit shocks.

Table 4 (A): Parameter Estimates Using Different Equilibria (1988)

	Favors Kmart	Favors Wal-Mart	Favors Wal-Mart in South
Kmart's Profit			
Log Population	1.49*	1.53*	1.46*
	(0.09)	(0.12)	(0.13)
Log Retail Sales	2.19*	2.16*	2.18*
	(0.25)	(0.23)	(0.30)
Urban Ratio	2.07*	2.15*	2.23*
	(0.42)	(0.36)	(0.36)
Midwest	0.40*	0.32*	0.27*
	(0.12)	(0.10)	(0.09)
Constant	-24.60*	-24.20*	-24.30*
	(2.07)	(1.83)	(2.51)
δ_{kw}	-0.48*	-0.51†	-0.46†
	(0.22)	(0.29)	(0.25)
δ_{kk}	0.63*	0.71†	0.72*
	(0.20)	(0.36)	(0.16)
δ_{ks}	-0.07	-0.13	-0.09
	(0.05)	(0.15)	(0.09)
Wal-Mart's Profit			
Log Population	1.54*	1.52*	1.43*
	(0.15)	(0.21)	(0.13)
Log Retail Sales	1.56*	1.52*	1.55*
	(0.35)	(0.25)	(0.17)
Urban Ratio	2.19*	2.31*	2.45*
	(0.35)	(0.34)	(0.38)
Log Distance	-1.31*	-1.29*	-1.34*
	(0.16)	(0.17)	(0.12)
South	0.94*	1.03*	1.00*
	(0.13)	(0.13)	(0.21)
Constant	-10.90*	-10.88*	-10.58*
	(2.98)	(1.69)	(1.46)
δ_{wk}	-1.54*	-1.58*	-1.51*
	(0.21)	(0.32)	(0.32)
δ_{ww}	1.22*	1.32*	1.25*
	(0.43)	(0.24)	(0.47)
δ_{ws}	-0.03	-0.03	-0.03
	(0.11)	(0.09)	(0.07)
Small Firms' Profit			
Log Population	1.65*	1.64*	1.59*
	(0.18)	(0.26)	(0.13)
Log Retail Sales	1.04*	1.04*	1.04*
	(0.12)	(0.16)	(0.12)
Urban Ratio	-0.46*	-0.53†	-0.46*
	(0.21)	(0.32)	(0.19)
South	0.88*	0.85*	0.93*
	(0.14)	(0.11)	(0.09)
Constant	-10.22*	-10.21*	-10.18*
	(0.98)	(1.40)	(1.05)
δ_{sk}	-1.20*	-1.14*	-1.12*
	(0.23)	(0.19)	(0.18)
δ_{sw}	-1.11*	-1.08*	-1.12*
	(0.16)	(0.15)	(0.13)
δ_{ss}	-2.14*	-2.14*	-2.10*
	(0.28)	(0.37)	(0.16)
ρ	0.53*	0.54*	0.54*
	(0.11)	(0.12)	(0.08)
Function Value	62.84	62.87	71.30
Number of Observations	2065	2065	2065

Note: * denotes significance at the 5% confidence level, and † denotes significance at the 10% confidence level. Standard errors are in parentheses. See Table 2 for the explanation of the variables and parameters.

Table 4 (B): Parameter Estimates Using Different Equilibria (1997)

	Favors Kmart	Favors Wal-Mart	Favors Wal-Mart in South
Kmart's Profit			
Log Population	1.84*	1.66*	1.70*
	(0.13)	(0.13)	(0.15)
Log Retail Sales	2.01*	1.91*	1.83*
	(0.23)	(0.34)	(0.22)
Urban Ratio	1.55*	1.57*	1.74*
	(0.29)	(0.20)	(0.50)
Midwest	0.32*	0.26*	0.22
	(0.14)	(0.10)	(0.14)
Constant	-24.08*	-22.52*	-22.18*
	(2.07)	(3.16)	(1.73)
δ_{kw}	-0.93*	-0.94*	-0.92*
	(0.29)	(0.23)	(0.29)
δ_{kk}	0.75*	0.82*	0.74*
	(0.36)	(0.33)	(0.36)
δ_{ks}	-0.02	-0.02	-0.01
	(0.05)	(0.07)	(0.14)
Wal-Mart's Profit			
Log Population	2.16*	2.05*	2.03*
	(0.15)	(0.14)	(0.17)
Log Retail Sales	1.85*	1.72*	1.74*
	(0.25)	(0.15)	(0.20)
Urban Ratio	1.73*	1.72*	1.90*
	(0.40)	(0.53)	(0.51)
Log Distance	-1.01*	-1.04*	-1.04*
	(0.14)	(0.10)	(0.10)
South	0.61*	0.74*	0.55*
	(0.11)	(0.08)	(0.10)
Constant	-16.37*	-14.80*	-14.87*
	(2.17)	(1.30)	(1.26)
δ_{wk}	-1.13*	-1.26*	-1.15*
	(0.30)	(0.21)	(0.55)
δ_{ww}	0.89*	0.91*	0.87*
	(0.39)	(0.23)	(0.27)
δ_{ws}	-0.03	-0.10	-0.05
	(0.12)	(0.12)	(0.10)
Small Firms' Profit			
Log Population	1.90*	1.92*	1.95*
	(0.26)	(0.15)	(0.17)
Log Retail Sales	1.17*	1.20*	1.15*
	(0.16)	(0.21)	(0.12)
Urban Ratio	-0.78*	-0.80*	-0.73†
	(0.24)	(0.21)	(0.39)
South	1.03*	1.00*	0.97*
	(0.17)	(0.12)	(0.07)
Constant	-11.89*	-12.18*	-11.84*
	(1.56)	(1.82)	(1.21)
δ_{sk}	-1.00*	-1.07*	-1.09*
	(0.20)	(0.18)	(0.30)
δ_{sw}	-1.03*	-1.03*	-1.04*
	(0.21)	(0.15)	(0.23)
δ_{ss}	-2.41*	-2.39*	-2.40*
	(0.35)	(0.20)	(0.17)
ρ	0.65*	0.67*	0.63*
	(0.10)	(0.08)	(0.14)
Function Value	30.80	32.53	37.70
Number of Observations	2065	2065	2065

Note: * denotes significance at the 5% confidence level, and † denotes significance at the 10% confidence level. Standard errors are in parentheses. See Table 2 for the explanation of the variables and parameters.

Table 5 (A): Model's Goodness of Fit

	1988		1997	
Number of:	Sample Mean	Model Mean	Sample Mean	Model Mean
Kmart	0.21	0.21	0.19	0.20
Wal-Mart	0.32	0.32	0.48	0.49
Small Firms	3.86	3.79	3.49	3.43

Table 5 (B): Correlation between Model Prediction and Sample Observation

Number of:	1988	1997
Kmart	0.66	0.64
Wal-Mart	0.73	0.75
Small Firms	0.63	0.64

Table 6: Model Predicted Profit vs. Accounting Profit

	Model Average			Average Accounting Profit		
	1988	1997	1997/ 1988	1988 ($mill.)	1997 ($mill.)	1997/ 1988
Kmart	0.80	0.74	0.92	0.56	0.14*	0.25
Wal-Mart	1.03	1.55	1.51	0.95	1.34	1.41

Source: Kmart’s and Wal-Mart’s SEC 10-K annual report.
*: Kmart’s accounting profit fluctuated dramatically in the 1990s, due to the financial obligations of various divested businesses. A better indicator of its store profit is probably the average store sales, which remained stagnant throughout the 90s.

Table 7 (A): Number of Kmart Stores When the Market Size Changes

	1988		1997	
	Percent	Total	Percent	Total
Base Case	100.0%	431	100.0%	408
Population Up 10%	112.1%	483	115.7%	472
Retail Sales Up 10%	117.6%	507	117.4%	479
Urban Ratio Up 10%	107.0%	461	106.1%	433
Midwest=0 for All Counties	86.1%	371	87.7%	358
Midwest=1 for All Counties	119.3%	514	115.4%	471

Table 7 (B): Number of Wal-Mart Stores When the Market Size Changes

	1988		1997	
	Percent	Total	Percent	Total
Base Case	100.0%	658	100.0%	1014
Population Up 10%	110.3%	726	108.3%	1098
Retail Sales Up 10%	110.5%	727	107.1%	1086
Urban Ratio Up 10%	105.3%	693	102.5%	1039
Distance Up 10%	91.6%	603	96.2%	975
South=0 for All Counties	64.1%	422	88.4%	896
South=1 for All Counties	135.1%	889	113.2%	1148

Table 7 (C): Number of Small Firms When the Market Size Changes

	1988		1997	
	Percent	Total	Percent	Total
Base Case	100.0%	7831	100.0%	7090
Population Up 10%	108.7%	8511	109.0%	7727
Retail Sales Up 10%	105.4%	8253	105.4%	7474
Urban Ratio Up 10%	99.3%	7773	98.7%	7000
South=0 for All Counties	78.4%	6136	76.2%	5404
South=1 for All Counties	124.8%	9775	125.0%	8860

Note: for each of the simulation exercises in all three panels, I fix other firms' profits and change only the profit of the target firm in accordance with the change in the market size. I re-solve the model to obtain the equilibrium numbers of firms. For example, in the second row of Table 7 (A), I increase Kmart's profit according to a ten percent increase in population while holding Wal-Mart and small firms' profits constant. Using this new set of profits, the equilibrium number of Kmart stores is 12.1% higher than in the base case in 1988.

Table 8 (A): Number of Small Firms with Different Market Structure

	1988		1997	
	Percent	Total	Percent	Total
No Kmart or Wal-Mart	100.0%	12070	100.0%	10946
Only Kmart in Each Market	54.0%	6519	63.8%	6985
Only Wal-Mart in Each Market	56.7%	6849	63.0%	6898
Both Kmart and Wal-Mart	28.6%	3457	38.4%	4198
Wal-Mart Competes with Kmart	64.9%	7831	64.8%	7090
Wal-Mart Takes Over Kmart	72.9%	8796	72.3%	7918

Table 8 (B): Competition Effect for Kmart and Wal-Mart

	1988		1997	
Number of Kmart Stores	Percent	Total	Percent	Total
Base Case	100.0%	431	100.0%	408
Wm in Each Market	78.0%	336	79.9%	326
Wm Exits Each Market	111.1%	479	149.5%	610
Not Compete with Small	108.1%	466	102.7%	419

	1988		1997	
Number of Wal-Mart Stores	Percent	Total	Percent	Total
Base Case	100.0%	658	100.0%	1014
Km in Each Market	48.3%	318	71.8%	728
Km Exits Each Market	128.6%	846	108.6%	1101
Not Compete with Small	102.6%	675	101.5%	1029

Table 8 (C) : Chain Effect for Kmart and Wal-Mart

	Kmart		Wal-Mart	
	1988	1997	1988	1997
Percentage of Profit Explained by Chain Effect	14.0%	17.4%	10.2%	12.3%
Reduction in Number of Stores with No Chain Effect	40	46	125	109

Note: for the first four rows in Table 8(A), I fix the number of Kmart and Wal-Mart stores as specified and solve for the equilibrium number of small stores. For the last two rows in Table 8(A) and all rows (except for the rows of 'Base Case') in Table 8(B), I re-solve the full model using the specified assumptions. 'Base Case' in Table 8(B) is what we observe in the data when Kmart competes with Wal-Mart.

Table 9: The Impact of Wal-Mart's Expansion on Small Stores

	1988	1997
Observed Decrease in the Number of Small Stores	748	748
Predicted Decrease from the Full Model	558	383
Percentage Explained	75%	51%
Predicted Decrease from Ordered Probit	247	149
Percentage Explained	33%	20%

Note: for the full model, the predicted 558 store exits in 1988 are obtained by simulating the change in the number of small stores using the 1988 coefficients for Kmart's and the small stores' profit functions, but the 1997 coefficients for Wal-Mart's profit function. The column of 1997 uses the 1997 coefficients for Kmart's and small stores' profit functions, but the 1988 coefficients for Wal-Mart's profit function. For the ordered probit model, the predicted store exits are the difference between the expected number of small stores using Wal-Mart's 1988 store number and the expected number of small stores using Wal-Mart's 1997 store number, both of which calculated using the probit coefficient estimates for the indicated year.

Table 10: The Impact of Government Subsidies

	Average Number of Stores		Changes in the Number of Stores Compared to the Base Case	
	1988	1997	1988	1997
Base Case				
Kmart	0.21	0.20		
Wal-Mart	0.32	0.49		
Small Firms	3.79	3.43		
Subsidize Kmart's Profit by 10%				
Kmart	0.22	0.21	0.01	0.01
Wal-Mart	0.31	0.49	-0.01	0.00
Small Firms	3.77	3.41	-0.03	-0.02
Subsidize Wal-Mart's Profit by 10%				
Kmart	0.21	0.19	0.00	-0.01
Wal-Mart	0.34	0.52	0.02	0.03
Small Firms	3.74	3.39	-0.05	-0.04
Subsidize Small Firms' Profit by 100%				
Kmart	0.21	0.20	0.00	0.00
Wal-Mart	0.32	0.49	0.00	0.00
Small Firms	4.61	4.23	0.81	0.80

Note: for each of these counter-factual exercises, I incorporate the change in the subsidized firm's profit and re-solve the model to obtain the equilibrium numbers of stores.