Econometrica, Vol. 76, No. 6 (November, 2008), 1263–1316

WHAT HAPPENS WHEN WAL-MART COMES TO TOWN: AN EMPIRICAL ANALYSIS OF THE DISCOUNT RETAILING INDUSTRY

BY PANLE JIA[1]

In the past few decades multistore retailers, especially those with 100 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 other retailers and local communities. This paper develops an empirical model to assess the impact of chain stores on other 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 lattice theory is exploited to solve this complicated entry game among chains and other discount retailers in a large number of markets. It is found that the negative impact of Kmart's presence on Wal-Mart's profit was much stronger in 1988 than in 1997, while the opposite is true for the effect of Wal-Mart's presence on Kmart's profit. Having a chain store in a market makes roughly 50% of the discount stores unprofitable. Wal-Mart's expansion from the late 1980s to the late 1990s explains about 40–50% of the net change in the number of small discount stores and 30–40% for all other discount stores. Scale economies were important for Wal-Mart, but less so for Kmart, and the magnitude did not grow proportionately with the chains' sizes.

KEYWORDS: Chain, entry, spatial correlation, Wal-Mart, lattice.

> 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.
>
> *Up Against the Wal-Mart* Archer and Taylor (1994)

> 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]This paper is a revision of Chapter 1 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 also thank the editor, three anonymous referees, Donald Andrews, Pat Bajari, Donald Brown, Judy Chevalier, Tom Holmes, Yuichi Kitamura, Ariel Pakes, Herbert Scarf, and seminar participants at Boston University, Columbia University, Duke University, Harvard, MIT, Northwestern University, NYU, Princeton, Stanford University, UCLA, UCSD, University of Chicago, University of Michigan, University of Minnesota, University of Pennsylvania, and participants of the 2006 *Review of Economic Studies* European Meetings in Oslo, Essex, and Tel Aviv for many helpful comments. Financial support from the Cowles Foundation Prize and a Horowitz Foundation Fellowship is gratefully acknowledged. All errors are my own. Comments are welcome.

© 2008 The Econometric Society DOI: 10.3982/ECTA6649

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* (1988–1997), 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.[2,3] Over the span of several decades, the sector has emerged from the fringe of the retail industry and become one of the major sectors.[4] 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 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.[5] 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 100 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.[6] Since the late 1960s, their share of the retail market more than doubled. In spite of the dominance of

[2]See the annual report, "The True Look of the Discount Industry," in the June 1962 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.

[3]According to *Annual Benchmark Report for Retail Trade and Food Services* (U.S. Census Bureau (1993–1997)), 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.

[4]The other retail sectors are building materials, food stores, automotive dealers, apparel, furniture, eating and drinking places, and miscellaneous retail.

[5]See Shils (1997).

[6]See U.S. Census Bureau (1997).

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.[7] In entry models, for example, the store entry decisions of multiunit 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 2000 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 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 algorithm proposed above exploits the game's supermodularity structure to solve a complicated problem. However, it has a couple of limitations. First, it is not applicable to oligopoly games with three or more chains.[8] The

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

[8]Entry games are not supermodular in general, as the competition effect is usually assumed to be negative. However, with only two chains, we can redefine the strategy space for one player to be the negative of the original space. Then the game associated with the new strategy space is supermodular, provided that each chain's profit function is supermodular in its own strategy. See Section 5.2 for details. This would not work for oligopoly games with three or more chains.

algorithm is also not applicable to situations with negative chain effects, which might happen if a chain's own stores in different markets compete for sales and the negative business stealing effect overwhelms the positive spillover effect. However, the frame is general enough to accommodate business stealing among a chain's own stores within a market. See Section 6.2.5 for further discussions. Nishida (2008) is a nice application that studies retail chains in the Japanese markets.

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.[9] The results indicate that the negative impact of Kmart's presence on Wal-Mart's profit was much stronger in 1988 than in 1997, while the opposite is true for the effect of Wal-Mart's presence on Kmart's profit. Having a chain store in a market makes roughly 50% of the discount stores unprofitable. Wal-Mart's expansion from the late 1980s to the late 1990s explains about 37–55% of the net change in the number of small discount stores and 34–41% for all other discount stores. Scale economies were important to Wal-Mart, but less so for Kmart, and 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.

This paper complements a recent study by Holmes (2005) which analyzes the diffusion process of Wal-Mart stores. Holmes's approach is appealing because he derives the magnitude of the economies of density, a concept similar to the chain effect in this paper, 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 currently 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 three-stage model. In the first stage, or the "pre-chain" period, small retailers make entry decisions without anticipating the future entry of Kmart or Wal-Mart. In the second stage, Kmart or Wal-Mart emerges in the retail industry and optimally locates their stores across the entire set of markets. In the third stage, existing small firms decide whether to continue their business, while potential entrants

[9]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.

decide whether to enter the market and compete with the chains. 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 (1991, 1990) 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 multiproduct 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) considered the quality choices of highway motels and Seim (2006) studied 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 multiproduct 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, Slade, and Brett (2002), Smith (2004), and Davis (2006). 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.

Like many other discrete-choice models with complete information, the entry games generally allow multiple equilibria. There is a very active literature on estimating discrete-choice games that explicitly addresses the issue of multiple equilibria. For example, Tamer (2003) proposed an exclusion condition that leads to point identification in two-by-two games. Andrews, Berry, and Jia (2004), Chernozhukov, Hong, and Tamer (2007), Pakes, Porter, Ho, and Ishii (2005), Romano and Shaikh (2006), and others analyzed bound estimations that exploit inequality constraints derived from necessary conditions. Bajari, Hong, and Ryan (2007) examined the identification and estimation of the equilibrium selection mechanism as well as the payoff function. Ciliberto and Tamer (2006) studied multiple equilibria in the airline markets and used the methodology of Chernozhukov, Hong, and Tamer (2007) to construct the confidence region. Ackerberg and Gowrisankaran (2007) analyzed banks' adoptions of the automated clearinghouse electronic payment system, assuming each network was in one of the two extreme equilibria with a certain probability. In the current application, I estimate the parameters using the equilibrium that is most profitable for Kmart, and also provide the parameter estimates at two other different equilibria as a robustness check.

This paper's algorithm is an application of the lattice theory, in particular Tarski's (1955) fixed point theorem and Topkis's (1978) monotonicity theorem. Milgrom and Shannon (1994) derived a necessary and sufficient condition for the solution set of an optimization problem to be monotonic in the parameters of the problem. Athey (2002) extended the monotone comparative statics to situations with uncertainty.

Finally, this paper is part of the growing literature on Wal-Mart, which includes Stone (1995), Basker (2005b, 2005a), Hausman and Leibtag (2005), Neumark, Zhang, and Ciccarella (2005), and Zhu and Singh (2007).

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. Data and programs are provided as supplemental material (Jia (2008)).

2. INDUSTRY BACKGROUND

Discount retailing is one of the most dynamic sectors in the retail industry. Table I displays some statistics for the industry from 1960 to 1997. The sales revenue for this sector, in 2004 U.S. 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 25 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

TABLE I

THE DISCOUNT INDUSTRY FROM 1960 TO 1997[a]

Year	Number of Discount Stores	Total Sales (2004 \$, billions)	Average Store Size (thousand ft^2)	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

[a]Source: Various issues of *Discount Merchandiser*. The numbers include only traditional discount stores. Wholesale clubs, supercenters, and special retailing stores are excluded.

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, p. 32)). 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, and so on. 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 an 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 continued to grow, other retailers started 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" (House Committee on Small Business (1994)). 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 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 (1988–1997). The directory covers all operating discount stores of more than 10,000 ft^2. For each store, the directory lists its name, size, street address, telephone number, store format, and firm affiliation.[10] 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 industry definition, I restrict the sample period to the 10 years before the classification change. As first documented in Basker (2005a), the directory was not fully updated for some years. Fortunately, it was fairly accurate for the years used in this study. See Appendix A for details.

The second data set, the County Business Patterns (CBP), tabulates at the county level the number of establishments by employment size category by industry sectors.[11] 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. Both Kmart and Wal-Mart are classified as firms in the general merchandise sector. I focus on two groups of retailers that compete with them: (a) small general merchandise stores with 19 or fewer employees; (b) all retailers in the general merchandise sector. I also experimented unsuccessfully with modeling the competition between these chains and retailers in a group of sectors. The model is too stylized to accommodate the vast differences between retailers in different sectors.

The number of retailers in the "pre-chain" period comes from 1978 CBP data. Data prior to 1977 are in tape format and not readily usable. I downloaded the county business pattern data from the Geospatial and Statistical Data Center of University of Virginia.[12]

Data on county level population were downloaded from the websites of the 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.

[10]The directory stopped providing store size information in 1997 and changed the inclusion criterion to 20,000 ft^2 in 1998. The store formats include membership stores, regional offices, and, in later years, distribution centers.

[11]CBP reports data at the establishment level, not the firm level. As it does not include information on firm ownership, I do not know which establishments are owned by the same firm. Given this data limitation, I have assumed that, in contrast to chain stores, all small retailers are single-unit firms. Throughout this paper, the terms "small firms" and "small stores" will be used interchangeably.

[12]The web address (as of January 2008) is http://fisher.lib.virginia.edu/collections/stats/ccdb/.

3.2. *Market Definition and Data Description*

In this paper, a market is defined as a county. Although the Chain Store Guide (1988–1997) 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 5000 and 64,000 from 1988 to 1997. There are 2065 such counties among a total of 3140 in the United States. According to Vance and Scott (1994), the minimum county population for a Wal-Mart store was 5000 in the 1980s, while Kmart concentrated in places with a much larger population. 9% of all U.S. counties were smaller than 5000 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 was very complex with a large number of competitors and many market niches. Defining a county as a market is likely to be problematic for these counties. Given the data limitation, I model entry decisions in those 2065 small- to medium-sized counties, and treat the chain stores in the other counties as exogenously given. The limitation of this approach is that the spillover effect from the chain stores located in large counties is also treated as exogenous. Using a county as a market definition also assumes away the cross-border shopping behavior. In future research, any data on the geographic patterns of consumers' shopping behavior would enable a more reasonable market definition.

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 20 stores in the counties studied here. I do not include Target in the analysis.

In the sample, only 8 counties had two Kmart stores and 49 counties had two Wal-Mart stores in 1988; the figures were 8 and 66 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 of the second store in the same market. In Section 6.2.5, I discuss how to extend the algorithm to allow for multiple stores in any given market.

Table II presents summary statistics of the sample for 1978, 1988, and 1997. The average county population was 21,470 in 1978. It increased by 5% between 1978 and 1988, and by 8% between 1988 and 1997. Retail sales per capita, in 1984 dollars, was $4070 in 1977. It dropped to $3690 in 1988, but recovered to $4050 in 1997. The average percentage of urban population was 30% in 1978. It barely changed between 1978 and 1988, but increased to 33% in 1997. About

TABLE II

SUMMARY STATISTICS FOR THE DATA SET[a]

Variable	1978 Mean	1978 Std.	1988 Mean	1988 Std.	1997 Mean	1997 Std.
Population (thousand)	21.47	13.38	22.47	14.12	24.27	15.67
Per capita retail sales (1984 $, thousands)	4.07	1.42	3.69	1.44	4.05	2.02
Percentage of urban population	0.30	0.23	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	0.41	0.49
South (1 if Southwest or Southeast)	0.50	0.50	0.50	0.50	0.50	0.50
Distance to Benton, AR (100 miles)	6.14	3.88	6.14	3.88	6.14	3.88
% of counties with Kmart stores			0.21	0.41	0.19	0.39
% of counties with Wal-Mart stores			0.32	0.47	0.48	0.50
Number of discount stores with 1–19 employees	4.75	2.86	3.79	2.61	3.46	2.47
Number of all discount stores (excluding Kmart and Wal-Mart)	4.89	3.24	4.54	3.10	4.04	2.85
Number of counties	2065					

[a]Source: The population is from the website of the Missouri State Census Data Center. Retail sales are from the 1977, 1987, and 1997 Economic Census. The percentage of urban population is from the 1980, 1990, and 2000 decennial census. Region dummies are defined according to the 1990 census. The numbers of Kmart and Wal-Mart stores are from the annual reference *Directory of Discount Department Stores* (Chain Store Guide (1988–1997)). The numbers of small discount stores and all other discount stores are from various years of the county business patterns.

one-quarter of the counties were 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 region (which includes 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 in 1988. The number dropped slightly to 19% in 1997. In comparison, Wal-Mart had stores in 32% of the counties in 1988 and 48% in 1997. I do not have data on the number of Kmart and Wal-Mart stores in the sample counties in 1978. Before the 1980s, Kmart was mainly operating in large metropolitan areas, and Wal-Mart was only a small regional firm. I assume that in 1978, retailers in my sample counties did not face competition from these two chains.

In 1978, the average number of discount stores per county was 4.89. The majority of them were quite small, as stores with 1–19 employees accounted for 4.75 of them. In 1988, the number of discount stores (excluding Kmart and Wal-Mart stores) dropped to 4.54, while that of small stores (with 1–19 employees) dropped to 3.79. By 1997, these numbers further declined to 4.04 and 3.46, respectively.

4. MODELING

4.1. *Model Setup*

The model I develop is a three-stage game. Stage 1 corresponds to the pre-chain period when only small firms compete against each other.[13] They enter the market if profit after entry recovers the sunk cost. In stage 2, Kmart and Wal-Mart simultaneously choose store locations to maximize their total profits in all markets. In the last stage, existing small firms decide whether to continue their business, while potential entrants decide whether to enter the market to compete with the chain stores and the existing small stores. Small firms are single-unit stores and only enter one market. In contrast, Kmart and Wal-Mart operate many stores and compete in multiple markets.

This is a complete-information game except for one major difference: in the first stage, small firms make entry decisions without anticipating Kmart and Wal-Mart in the later period. The emergence of Kmart and Wal-Mart in the second stage is an unexpected event for the small firms. Once Kmart and Wal-Mart have appeared on the stage, all firms obtain full knowledge of their rivals' profitability and the payoff structure. Facing a number of existing small firms in each market, Kmart and Wal-Mart make location choices, taking into consideration small retailers' adjustment in the third stage. Finally, unprofitable small firms exit the market and new entrants come in. Once these entry decisions are made, firms compete and profits are realized. Notice that I have implicitly assumed that chains can commit to their entry decisions in the second stage and do not further adjust after the third stage. 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.

The three-stage model is motivated by the fact that small retailers existed long before the era of the discount chains. Accordingly, the first stage should be considered as "historical" and the model uses this stage to fit the number of small retailers before the entry of Kmart and Wal-Mart. Stages 2 and 3 happen roughly concurrently: small stores adjust quickly once they observe big chains' decisions.

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.

[13]In the empirical application, I also estimate the model using all discount stores, not just small stores. See Section 7 for details.

Instead, I follow the convention in the entry literature and assume that firms' profit functions take a linear form and that profits decline in the presence of rivals.

In the first stage, or the pre-chain period, profit for a small store that operates in market m is

$$\Pi^0_{s,m} = X^0_m\beta_s + \delta_{ss}\ln(N^0_{s,m}) + \sqrt{1-\rho^2}\varepsilon^0_m + \rho\eta^0_{s,m} - \text{SC}, \tag{1}$$

where s stands for small stores. Profit from staying outside the market is normalized to 0 for all players.

There are several components in the small store's profit $\Pi^0_{s,m}$: the observed market size $X^0_m\beta_s$ that is parameterized by demand shifters, like population, the extent of urbanization, and so forth; the competition effect $\delta_{ss}\ln(N^0_{s,m})$ that is monotonically increasing (in the absolute value) in the number of competing stores $N_{s,m}$; the unobserved profit shock $\sqrt{1-\rho^2}\varepsilon^0_m + \rho\eta^0_{s,m}$, known to the firms but unknown to the econometrician; and the sunk cost of entry SC. As will become clear below, both the vector of observed market size variables X_m and the coefficients β are allowed to vary across different players. For example, Kmart might have some advantage in the Midwest, Wal-Mart stores might be more profitable in markets close to their headquarters, and small retailers might find it easier to survive in rural areas.

The unobserved profit shock has two elements: ε^0_m, the market-level profit shifter that affects all firms operating in the market, and $\eta^0_{s,m}$, a firm-specific profit shock. ε^0_m is assumed to be independent and identically distributed (i.i.d.) across markets, while $\eta^0_{s,m}$ is assumed to be i.i.d. across both firms and markets. $\sqrt{1-\rho^2}$ (with $0 \le \rho \le 1$) measures the importance of the market common shock. In principle, its impact can differ between chains and 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. $\eta^0_{s,m}$ incorporates the unobserved store level heterogeneity, including the management ability, the display style and shopping environment, and employees' morale or skills. 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. In addition, I assume that both ε^0_m and $\eta^0_{s,m}$ are standard normal random variables for computational convenience. In other applications, a more flexible distribution (like a mixture of normals) might be more appropriate.

As mentioned above, in the pre-chain stage, the small stores make entry decisions without anticipating Kmart's and Wal-Mart's entry. As a result, $N^0_{s,m}$ is only a function of $(X^0_m, \varepsilon^0_m, \eta^0_{s,m})$, and is independent of Kmart's and Wal-Mart's entry decisions in the second stage.

To describe the choices of Kmart and Wal-Mart, let me introduce some vector variables. 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. $D_i = \{D_{i,1}, \dots, D_{i,M}\}$ is a vector indicating chain i's location choices for the entire set of markets. $D_{j,m}$ denotes rival j's strategy in market m. Finally, let Z_{ml} designate the distance from market m to market l in miles and let $Z_m = \{Z_{m1}, \dots, Z_{mM}\}$.

The following equation system describes the payoff structure during the "post-chain" period when small firms compete against the chains.[14] The first equation is the profit for chains, the second equation is the profit for small firms, and the last equation defines how the market unobserved profit shocks evolve over time:

$$
\begin{aligned}
(2)\qquad & \Pi_{i,m}(D_i, D_{j,m}, N_{s,m}; X_m, Z_m, \varepsilon_m, \eta_{i,m}) \\
&\quad = D_{i,m} * \Bigg[X_m\beta_i + \delta_{ij} D_{j,m} + \delta_{is} \ln(N_{s,m} + 1) \\
&\qquad + \delta_{ii} \sum_{l \neq m} \frac{D_{i,l}}{Z_{ml}} + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m} \Bigg], \quad i, j \in \{k, w\}, \\
& \Pi_{s,m}(D_i, D_{j,m}, N_{s,m}; X_m, \varepsilon_m, \eta_{i,m}) \\
&\quad = X_m\beta_s + \sum_{i=k,w} \delta_{si} D_{i,m} + \delta_{ss} \ln(N_{s,m}) \\
&\qquad + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{s,m} - \mathrm{SC} * 1[\text{new entrant}], \quad \rho \in [0, 1], \\
& \varepsilon_m = \tau\varepsilon_m^0 + \sqrt{1-\tau^2}\tilde{\varepsilon}_m, \quad \tau \in [0, 1],
\end{aligned}
$$

where k denotes Kmart, w denotes Wal-Mart, and s denotes small firms. In the following, I discuss each equation in turn.

First, notice the presence of D_i in chain i's profit $\Pi_{i,m}(\cdot)$: profit in market m depends on the number of stores chain i has in other markets. Chains maximize their total profits in all markets $\sum_m \Pi_{i,m}$, internalizing the spillover effect among stores in different locations.

As mentioned above, $X_m\beta_i$ is indexed by i so that the market size can have a differential impact on firms' profitability. The competition effect from the rival chain is captured by $\delta_{ij} D_{j,m}$, where $D_{j,m} = 1$ 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. In equilibrium, the number of small firms in the

[14] I treat stage 2 and stage 3 as happening "concurrently" by assuming that both chains and small firms share the same market-level shock ε_m.

last stage is a function of Kmart's and Wal-Mart's 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.

The chain effect is denoted by $\delta_{ii} \sum_{l \neq m}(D_{i,l}/Z_{ml})$, the benefit that having stores in other markets generates for the profit in market m. δ_{ii} is assumed to be nonnegative. Nearby stores split the costs of operation, delivery, and advertising to achieve scale economies. They also share knowledge of local 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 $\delta_{ii}(D_{i,l}/Z_{ml})$ 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 observations to die away sufficiently fast. I also assume that the chain effect takes place among counties whose centroids are within 50 miles, or roughly an area that expands 75 miles in each direction. Including counties within 100 miles increases 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.

As in small firms' profit function, chain i's profit shock contains two elements: the market shifter common to all firms ε_m and the firm-specific shock $\eta_{i,m}$. Both are assumed to have a standard normal distribution.

Small firms' profit in the last stage $\Pi_{s,m}$ is similar to the pre-chain period, except for two elements. First, $\sum_{i=k,w} \delta_{si} D_{i,m}$ captures the impact of Kmart and Wal-Mart on small firms. Second, only new entrants pay the sunk cost SC.

The last equation describes the evolution of the market-level error term ε_m over time. $\tilde{\varepsilon}_m$ is a pure white noise that is independent across period. τ measures the persistence of the unobserved market features. Notice that both τ and the sunk cost SC can generate history dependence, but they have different implications. Consider two markets A and B that have a similar market size today with the same number of chain stores: $X_A = X_B$, $D_{i,A} = D_{i,B}$, $i \in \{k, w\}$. Market A used to be much bigger, but has gradually decreased in size. The opposite is true for market B: it was smaller before, but has expanded over time. If history does not matter (that is, both τ and SC are zero), these two markets should on average have the same number of small stores. However, if SC is important, then market A should have more small stores that entered in the past and have maintained their business after the entry of chain stores. In other words, big sunk cost implies that everything else equal, markets that were bigger historically carry more small stores in the current period. On the other hand, if τ is important, then some markets have more small stores throughout the time, but there are no systematic patterns between the market size in the previous period and the number of stores in the current period, as the history dependence is driven by the unobserved market shock ε_m that is assumed to be independent of X_m.

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. The chain effect $\delta_{ii}(D_{i,l}/Z_{ml})$ is also endogenous, because a large ε_m is associated with a high value of $D_{i,m}$, which increases the profitability of market l, and hence leads to a high value of $D_{i,l}$. I solve the chains' and small firms' entry decisions simultaneously within the model, and require the model to replicate the location patterns observed in the data.

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

5. SOLUTION ALGORITHM

The unobserved market-level profit shock ε_m, together with the chain effect $\delta_{ii}\sum_{l\neq m}(D_{i,l}/Z_{ml})$, renders all of the discrete variables $N^0_{s,m}$, $D_{i,m}$, $D_{j,m}$, $D_{i,l}$, 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 $D = \{D_1, \dots, D_M\}$. The profit-maximization problem is

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

The choice variable D_m appears in the profit function in two ways. First, it directly determines profit in market m: the firm earns $X_m + \delta\sum_{l\neq m}(D_l/Z_{ml})$ if $D_m = 1$ and earns 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 a 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 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.[15] 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 location choice.

Throughout this paper, the comparison between vectors is coordinatewise. 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,

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

which leads to

$$D_m^* = 1\left[X_m + 2\delta \sum_{l\neq m} \frac{D_l^*}{Z_{ml}} \geq 0\right] \quad \forall m. \tag{3}$$

The derivation of equation (3) is left to Appendix B.1. These conditions have the usual interpretation that $X_m + 2\delta\sum_{l\neq m}(D_l^*/Z_{ml})$ is market m's marginal

[15] A symmetric matrix is positive (negative) semidefinite if and only if all the eigenvalues are nonnegative (nonpositive). The Hessian of the profit function (2) 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 $H = 0$.

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 $V_m(D) = 1[X_m + 2\delta \sum_{l \neq m}(D_l/Z_{ml}) \geq 0]$ and $V(D) = \{V_1(D), \ldots, V_M(D)\}$. $V(\cdot)$ 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 nonnegative. 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 B.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 cannot be replaced by assuming that $Y(X)$ is a monotone function. Appendix B.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 $D^0 = \sup(\mathbf{D}) = \{1, \ldots, 1\}$. 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 construction, 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: $D^0 \geq D^1 \geq \cdots \geq D^t$. 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: $D^T = D^{T+1}$, $T \leq M$. 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 $D^0 = \inf(\mathbf{D}) = \{0, \ldots, 0\}$ is the least element in the set of fixed points. Denote it by D^L. In Appendix B.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 from the rival chain, the profit function for chain i becomes $\Pi_i(D_i, D_j) = \sum_{m=1}^{M} [D_{i,m} * (X_{im} + \delta_{ii} \sum_{l \neq m} (D_{i,l}/Z_{ml}) + \delta_{ij} D_{j,m})]$, 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 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.[16,17]

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 B.4, I verify that the profit function $\Pi_i(D_i, D_j) = \sum_{m=1}^{M} [D_{i,m} * (X_{im} + \delta_{ii} \sum_{l \neq m} (D_{i,l}/Z_{ml}) + \delta_{ij} D_{j,m})]$ 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 function $\arg\max_{D_i \in \mathbf{D}_i} \Pi_i(D_i, D_j)$ decreases in rival j's strategy D_j. Similarly for chain j's best response to i's strategy.

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

[17]See Milgrom and Shannon (1994) for a detailed discussion on the necessary and sufficient conditions for the solution set of an optimization problem to be monotonic in the parameters.

The set of Nash equilibria of a supermodular game is nonempty, and it has a greatest element and a least element.[18,19] The current entry game is not supermodular, as the profit function has decreasing differences in the joint strategy space $\mathbf{D} \times \mathbf{D}$. This leads to a nonincreasing joint best response function, and we know from the discussion after Theorem 1 that a nonincreasing 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 rewritten as

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

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.[20]

To obtain the equilibrium most profitable for Kmart, start with the smallest vector in Wal-Mart's strategy space: $D_w^0 = \inf(\mathbf{D}) = \{0, \dots, 0\}$. 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 this process generates two monotone sequences: $D_k^1 \geq D_k^2 \geq \cdots \geq D_k^t$ and $D_w^1 \leq D_w^2 \leq \cdots \leq D_w^t$. 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: $D_k^T = D_k^{T+1}$ and $D_w^T = D_w^{T+1}$, $T \leq M$. The convergent

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

[19]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.

[20]See page 185 of Topkis (1998) for a detailed discussion.

vectors (D_k^T, D_w^T) constitute an equilibrium: $D_k^T = K(D_w^T)$ and $D_w^T = W(D_k^T)$. 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 $D_w^0 = \{0, \dots, 0\}$ 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.

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 $D_k^0 = \inf(\mathbf{D}) = \{0, \dots, 0\}$ 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*

It is straightforward to solve the pre-chain stage: N_s^0 is the largest integer such that all entering firms can recover their sunk cost,[21]

$$\Pi_{s,m}^0 = X_m^0 \beta_s + \delta_{ss} \ln(N_{s,m}^0) + \sqrt{1-\rho^2}\varepsilon_m^0 + \rho\eta_{s,m}^0 - \mathrm{SC} > 0.$$

After the entry of chain stores, some of the existing small stores will find it unprofitable to compete with chains and exit the market, while other more efficient stores (the ones with larger $\eta_{s,m}$) will enter the market after paying the sunk cost of entry. The number of small stores in the post-chain period is a sum of new entrants N_s^E and the remaining incumbents N_s^I. Except for the idiosyncratic profit shocks, the only difference between these two groups of small firms is the sunk cost:

$$\begin{aligned}\Pi_{s,m} = {} & X_m \beta_s + \sum_{i=k,w} \delta_{si} D_{i,m} + \delta_{ss} \ln(N_{s,m}) + \sqrt{1-\rho^2}\varepsilon_m \\ & + \rho\eta_{s,m} - \mathrm{SC} * 1[\text{new entrant}].\end{aligned}$$

[21]The number of potential small entrants is assumed to be 11, which was within the top two percentile of the distribution of the number of small stores. I also experimented with the maximum number of small stores throughout the sample period. See footnote 28 for details.

Potential entrants will enter the market only if the post-chain profit can recover the sunk cost, while existing small firms will maintain the business as long as the profit is nonnegative.

Both the number of entrants $N_s^E(D_k, D_w)$ and the number of remaining incumbents $N_s^I(D_k, D_w)$ are well defined functions of the number of chain stores. To solve the game between chains and small stores in the post-chain period, I follow the standard backward induction and plug in small stores' reaction functions to the chains' profit function. Specifically, chain i's profit function now becomes

$$\Pi_i(D_i, D_j) = \sum_{m=1}^{M}\Bigg[D_{i,m} * \Bigg(X_{im} + \delta_{ii}\sum_{l\neq m}\frac{D_{i,l}}{Z_{ml}} + \delta_{ij}D_{j,m} + \delta_{is}\ln\big(N_s^E(D_{i,m}, D_{j,m}) + N_s^I(D_{i,m}, D_{j,m}) + 1\big)\Bigg)\Bigg],$$

where X_{im} is defined in Section 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.[22]

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 B.5, I discuss a few technical details related to 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 empirical industrial organization literature are the method of simulated log-likelihood (MSL) and the method of simulated moments (MSM). 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

[22]If we ignore the integer problem and the sunk cost, then $\delta_{ss}\ln(N_s + 1)$ can be approximated by $-(X_{sm} + \delta_{sk}D_k + \delta_{sw}D_w)$, and the assumption is $\delta_{kw} - (\delta_{ks}\delta_{sw}/\delta_{ss}) < 0$ and $\delta_{wk} - (\delta_{ws}\delta_{sk}/\delta ss) < 0$. The expression is slightly more complicated with the integer constraint, and the distinction between existing small stores and new entrants. 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. I have verified in the empirical application that these conditions are indeed satisfied.

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 $\theta_0 = \{\beta_i, \delta_{ii}, \delta_{ij}, \rho, \tau, \mathrm{SC}\}_{i=k,w,s} \in \boldsymbol{\Theta} \subset \mathbf{R}^P$. 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 $\sum_{m=1}^{M} \hat{g}(X_m, \theta)$:

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

where $\hat{g}(\cdot)$ is a simulated estimate of the true moment function and $\boldsymbol{\Omega}_M$ is an $L \times L$ positive semidefinite weighting matrix. Assume $\boldsymbol{\Omega}_M \xrightarrow{p} \boldsymbol{\Omega}_0$, 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) showed 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, $\mathbf{A}_0 \equiv \mathbf{G}_0'\boldsymbol{\Omega}_0\mathbf{G}_0$, $\mathbf{B}_0 = \mathbf{G}_0'\boldsymbol{\Omega}_0\boldsymbol{c}_0\boldsymbol{\Omega}_0\mathbf{G}_0$, and $\boldsymbol{\Lambda}_0 = E[g(X_m, \theta_0)g(X_m, \theta_0)'] = \text{Var}[g(X_m, \theta_0)]$. If a consistent estimator of $\boldsymbol{\Lambda}_0^{-1}$ is used as the weighting matrix, the MSM estimator $\hat{\theta}$ is asymptotically efficient,[23] with its asymptotic variance being $\text{Avar}(\hat{\theta}) = (1 + R^{-1}) * (\mathbf{G}_0'\boldsymbol{\Lambda}_0^{-1}\mathbf{G}_0)^{-1}/M$.

The obstacle in using this standard 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 $D_{i,l}/Z_{ml}$ evaporates with distance. As a result, the covariance matrix in equation (5) is no longer valid.

[23]The MSM estimator $\hat{\theta}$ is asymptotically efficient relative to estimators which minimize a quadratic norm in $g(\cdot)$. Different moments could improve efficiency. I thank the referee for pointing this out.

Conley (1999) discussed method of moments estimators using data that exhibit spatial dependence. That paper provided sufficient conditions for consistency and normality, which require the underlying data generating process to satisfy a strong mixing condition.[24] Essentially, the dependence among observations should die away quickly as the distance increases. In the current application, the consistency condition requires that the covariance between $D_{i,m}$ and $D_{i,l}$ goes to 0 as their geographic distance increases.[25] In other words, the entry decisions in different markets should be nearly independent when the distance between these markets is sufficiently large.

Unlike some other iteration procedures that search for the fixed points, $|\delta_{ii}|$ does not have to be less than 1. To see this, note that by construction,

$$D_{i,m} = 1\left[X_{i,m} + 2\delta_{ii}\sum_{l\neq m}\frac{D_{i,l}}{Z_{ml}} \geq 0\right] \quad \forall m,$$

where $D_{i,m} = 1$ if chain i has a store in market m. The system stays stable as long as δ_{ii} is finite, because $D_{i,m}$ is bounded $\forall m$. The geographic scope of the spillover effect can increase with the sample size, but the sum $\delta_{ii}\sum_{l\neq m}(D_{i,l}/Z_{lm})$ should remain finite to prevent the profit function from exploding. There are many ways to formulate the relationship between the spillover effect and the distance, as long as it guarantees that the (pairwise) covariance between the chain stores' entry decisions in different markets goes to 0 as the geographic distance increases.[26]

[24]The asymptotic arguments require the data to be generated from locations that grow uniformly in spatial dimensions as the sample size increases.

[25]Here I briefly verify that the consistency condition is satisfied. By construction,

$$\begin{aligned} D_{i,m} &= 1\left[X_{i,m} + 2\delta_{ii}\frac{D_{i,l}}{Z_{ml}} + \rho\eta_{i,m} \geq 0\right], \\ D_{i,l} &= 1\left[X_{i,l} + 2\delta_{ii}\frac{D_{i,m}}{Z_{ml}} + \rho\eta_{i,l} \geq 0\right], \end{aligned}$$

where $X_{i,m} = X_m\beta_i + 2\delta_{ii}\sum_{k\neq l,m}(D_{i,k}/Z_{mk}) + \delta_{ij}D_{j,m} + \delta_{is}\ln(N_{s,m}+1) + \sqrt{1-\rho^2}\varepsilon_m$. The covariance between $D_{i,m}$ and $D_{i,l}$ is

$$\begin{aligned} \operatorname{cov}(D_{i,m}, D_{i,l}) &= E(D_{i,m} * D_{i,l}) - E(D_{i,m}) * E(D_{i,l}) \\ &\leq \Pr\left(\rho\eta_{i,m} \geq -X_{i,m} - \frac{2\delta_{ii}}{Z_{ml}},\ \rho\eta_{i,l} \geq -X_{i,l} - \frac{2\delta_{ii}}{Z_{ml}}\right) \\ &\quad - \Pr(\rho\eta_{i,m} \geq -X_{i,m}) * \Pr(\rho\eta_{i,l} \geq -X_{i,l}) \\ &\to 0 \quad \text{as} \quad Z_{ml} \to \infty. \end{aligned}$$

[26]The normality conditions in Conley (1999) require the covariance to decrease at a sufficiently fast rate. See page 9 of that paper for details. These conditions are trivially satisfied here, as the

With the presence of the spatial dependence, the asymptotic covariance matrix of the moment functions $\boldsymbol{\Lambda}_0$ in equation (5) should be replaced by $\boldsymbol{\Lambda}_0^d = \sum_{s\in M} E[g(X_m, \theta_0)g(X_s, \theta_0)']$. Conley (1999) proposed 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. Following Conley (1999) and Conley and Ligon (2002), the estimator of Λ_0^d is

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

where B_m is the set of markets whose centroid is within 50 miles of market m, including market m.[27] The implicit assumption is that the spillover effect is negligible for markets beyond 50 miles. I have also estimated the variance of the moment functions $\hat{\Lambda}$ summing over markets within a 100 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 small.

The estimation procedure is as follows.

Step 1. Start from some initial guess of the parameter values and draw independently from the normal distribution the following vectors: the market-level errors for both the pre-chain period and the post-chain period—$\{\varepsilon_m^0\}_{m=1}^M$ and $\{\tilde{\varepsilon}_m\}_{m=1}^M$; profit shocks for the chains—$\{\eta_{k,m}\}_{m=1}^M$ and $\{\eta_{w,m}\}_{m=1}^M$; and profit shocks for each of the potential small entrants—$\{\eta_{s,m}^0\}_{m=1}^M$ and $\{\eta_{s,m}\}_{m=1}^M$, where $s = 1, \dots, 11$.[28]

Step 2. Obtain the simulated profits $\hat{\Pi}_i$, $i = k, w, s$, and solve for $\hat{N}_s^0$, $\hat{D}_k$, $\hat{D}_w$, and $\hat{N}_s$.

Step 3. Repeat Steps 1 and 2 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 substitute a preliminary estimate $\tilde{\theta}$ into equation (6) to compute the optimal weight matrix $\hat{\Lambda}^{-1}$ for the second step.

spillover effect is assumed to occur only within 50 or 100 miles. In other applications, one needs to verify that these conditions are satisfied.

[27] As mentioned in Conley (1999), this estimator is inefficient and not always positive semidefinite. Newey and West (1987) introduced a weight function $w(l, m)$ as a numerical device to make the estimator positive semidefinite. The weight used in the empirical application is 0.5 for all the neighbors.

[28] The number of potential small entrants is assumed to be 11. During the sample period, only one county had 25 small stores, while the median number was 4 for the 1970s, and 3 for the 1980s and 1990s. As the memory requirement and the computational burden increase with the number of potential entrant, I have capped the maximum number of small stores at 11, which is within the top one percentile of the distribution in the 1990s and the top two percentile in the 1980s. The competition effects of chains on small stores do not change much with a maximum number of 25 small stores.

Instead of the usual machine-generated pseudorandom draws, I use Halton draws, which have better coverage properties and smaller simulation variances.[29,30] According to Train (2000), 100 Halton draws achieved greater accuracy in his mixed logit estimation than 1000 pseudorandom draws. The parameter estimation exploits 150 Halton simulation draws, while the variance is calculated with 300 Halton draws.

There are 29 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 stores in the pre-chain period as well as the post-chain period, (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, (d) the interaction between the market size variables and the above items, and (e) the difference in the number of small stores between the pre-chain and post-chain periods, interacted with the changes in the market size variables between these two periods.

6.2. *Discussion: A Closer Look at the Assumptions and Possible Extensions*

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 nonnegativity of the chain effect. In the last subsection, I consider possible extensions.

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, 1991) and Berry

[29]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 Train (2003) for an excellent discussion of the Halton draws.

[30]In situations of high-dimensional simulations (as is the case here), the standard Halton draws display high correlations. The estimation here uses shuffled Halton draws, as proposed in Hess and Polak (2003), which has documented that the high correlation can be easily removed by shuffling the Halton draws.

(1992) pointed 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.[31] Other researchers give up point identification of parameters and search for bounds. These papers typically involve bootstraps or subsampling, and are too computationally intensive to be applicable here.[32]

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 the different cases are very similar to one another, which provides evidence that the results are robust to the equilibrium choice. In Section 7.1, I also investigate the differences between these equilibria at the estimated parameter values.[33] On average, they differ only in a small portion of the sample, and results from the counterfactual exercises do not vary much across different equilibria.

6.2.2. *The Symmetry Assumption for Small Firms*

I have assumed that all small firms have the same profit function and only differ in the unobserved profit shocks. The assumption is necessitated by data

[31]For example, the number of entering firms in a given market is no longer unique in the current application with the chain effect. See footnote 40 for an illustration.

[32]For example, the methods proposed in Andrews, Berry, and Jia (2004), Chernozhukov, Hong, and Tamer (2007), and Romano and Shaikh (2006) all involve estimating the parameters for each bootstrap sample or subsample. It takes more than a day to estimate the model once; it will take about a year if 300 bootstrap samples or subsamples are used for inference. The method proposed by Pakes, Porter, Ho, and Ishii (2005) is less computationally demanding, but as the authors pointed out, the precision of their inference is still an open question.

[33]If we were to formally test whether the data prefer one equilibrium to the other, we need to derive the asymptotic distribution of the difference between the two minimized objective function values, each associated with a different equilibrium. It is a nonnested test that is somewhat involved, as one objective function contains moments with a nonzero mean at the true parameter values. In the current application, the objective function values are very similar in 1997 (108.68 for the objective function that uses the equilibrium most profitable for Kmart, 105.02 for the equilibrium most profitable for Wal-Mart, and 103.9 for the equilibrium that grants a regional advantage to each player), but differ somewhat in 1988 (the objective function values are 120.26, 120.77, and 136.74, respectively).

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 first and the third 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 (2) 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 monotonicity 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 the distance becomes sufficiently small. 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) estimated that for places with a population density of 20,000 people per 5-mile radius (which is comparable to an average city in my sample counties), 89% of the average consumers visit a Wal-Mart nearby.[34] When the distance increases to 5 miles, 44% of the consumers visit the store. The percentage drops to 7% if the store is 10 miles away. Survey studies also show that few consumers drive farther 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.

6.2.4. *Independent Error Terms:* ε_m

In the model, I have assumed that the market-level profit shocks ε_m are independent across markets. Under this assumption, the chain effect is identified from the geographic clustering pattern of the store locations. Theoretically, one can use the number of small stores across markets to identify the correlation in the error term, because small stores are assumed to be single-unit firms. Conditioning on the covariates, the number of small stores across markets is independent if there is no cross-sectional dependence in the error terms. Once

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

we control for the cross-sectional dependence, the extra clustering exhibited by the chain stores' location choice should be attributed to the chain effect. However, implementing this idea is difficult, as there is no easy way to simulate a large number of error terms that exhibit dependence with irregular spatial patterns. Therefore, cross-sectional dependence of the error term is potentially another explanation for the spatial clustering pattern that I currently attribute to the chain effect.

6.2.5. *Extensions*

Extending the model to allow for multiple stores in any given market involves only a slight modification. In solving the best response given the rival's strategy, instead of starting from $D = \{1, \dots, 1\}$, we use $D = \{N_1, \dots, N_M\}$, where N_m is the maximum number of stores a chain can potentially open in a given market m. The iteration will converge within $\sum_m N_m$ steps. Notice that even though the size of the strategy space has increased from 2^M to $\prod_{m=1}^M N_m$, the number of iterations only increases linearly, rather than exponentially, as there are at most $\sum_m N_m$ steps for $D = \{N_1, \dots, N_M\}$ to *monotonically* decrease to $\{0, \dots, 0\}$. In general, the computational complexity increases linearly with the number of stores in each market. There is one caveat: when there are more stores in an area that are owned by the same firm, the negative business stealing effect across markets can potentially outweigh the positive spillover effect. As a result, the assumption that $\delta_{ii} \geq 0$ might not be supported by data in some applications.

7. RESULTS

7.1. *Parameter Estimates*

The sample includes 2065 small- and medium-sized counties with populations between 5000 and 64,000 in the 1980s. 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 III displays the summary statistics of the distance weighted numbers of adjacent Kmart stores $\sum_{l \neq m, l \in B_m} (D_{k,l}/Z_{ml})$ and Wal-Mart stores $\sum_{l \neq m, l \in B_m} (D_{w,l}/Z_{ml})$, which measure the spillover from nearby stores (including stores outside the sample). In 1997, the Kmart spillover variable was slightly higher than in 1988 (0.13 vs. 0.11), but the Wal-Mart spillover variable was almost twice as big as in 1988 (0.19 vs. 0.10).

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

TABLE III

SUMMARY STATISTICS FOR THE DISTANCE WEIGHTED NUMBER OF ADJACENT STORES[a]

Variable	1988 Mean	1988 Std.	1997 Mean	1997 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			

[a]Source: *Directory of Discount Department Stores* (Chain Store Guide (1988–1997)).

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 Bentonville, 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. The constant in the small firms' profit function is allowed to differ between the pre-chain period and the post-chain period to capture some general trend in the number of small stores that is unrelated with chain stores' entry.

Tables IV and V list the parameter estimates for the full model in six different specifications. Table IV uses the 1988 data for the post-chain period, while Table V uses the 1997 data for this period. The first five columns focus on the competition between chains and small discount stores. The last column estimates the model using Kmart, Wal-Mart, and all other discount stores, including the small ones. The first column is the baseline specification, where the estimates are obtained using the equilibrium most profitable for Kmart. The second column estimates the model using the equilibrium most profitable for Wal-Mart, while the third column repeats the exercise using the equilibrium that grants Wal-Mart an advantage in the south and Kmart an advantage in the rest of the country. The estimates are quite similar across the different equilibria.

One might be concerned that retail sales is endogenous: conditioning on demographics, counties with a Kmart or Wal-Mart store will generate more retail

TABLE IV

PARAMETER ESTIMATES FROM DIFFERENT SPECIFICATIONS—1988[a]

	Baseline	Favors Wal-Mart	Regional Advantage	Personal Income	Rival Stores in Neighborhood	All Other Discount Stores
Kmart's profit						
Log population	1.40*	1.43*	1.44*	2.09*	1.38*	1.55*
	(0.11)	(0.09)	(0.09)	(0.11)	(0.10)	(0.08)
Log retail sales/log personal income	2.20*	2.27*	2.18*	1.78*	2.20*	2.25*
	(0.08)	(0.07)	(0.07)	(0.10)	(0.08)	(0.07)
Urban ratio	2.29*	2.37*	2.31*	2.98*	2.20*	2.24*
	(0.35)	(0.32)	(0.25)	(0.45)	(0.37)	(0.22)
Midwest	0.52*	0.54*	0.52*	0.27*	0.55*	0.47*
	(0.14)	(0.11)	(0.12)	(0.12)	(0.20)	(0.14)
Constant	−24.59*	−25.28*	−24.49*	−25.47*	−24.54*	−25.17*
	(0.73)	(0.51)	(0.50)	(0.67)	(0.69)	(0.58)
delta_kw	−0.33*	−0.28*	−0.31	−0.31*	−0.31	−0.25†
	(0.15)	(0.12)	(0.20)	(0.15)	(0.25)	(0.15)
delta_kk	0.59	0.64*	0.63	0.53*	0.57*	0.56*
	(0.68)	(0.16)	(0.50)	(0.27)	(0.28)	(0.22)
delta_ks	−0.01	−0.02	−0.01	−0.04	−0.001	−0.11
	(0.07)	(0.09)	(0.08)	(0.09)	(0.13)	(0.10)
delta_kw2					0.19	
					(4.76)	
Wal-Mart's profit						
Log population	1.39*	1.43*	1.40*	2.05*	1.37*	1.86*
	(0.08)	(0.09)	(0.09)	(0.16)	(0.15)	(0.12)
Log retail sales/log personal income	1.68*	1.73*	1.62*	1.22*	1.68*	1.62*
	(0.07)	(0.06)	(0.05)	(0.08)	(0.08)	(0.07)
Urban ratio	2.40*	2.43*	2.43*	3.37*	2.24*	2.15*
	(0.38)	(0.27)	(0.33)	(0.38)	(0.39)	(0.26)
Log distance	−1.49*	−1.54*	−1.42*	−1.49*	−1.48*	−1.57*
	(0.12)	(0.10)	(0.10)	(0.11)	(0.16)	(0.12)
South	1.06*	1.11*	1.05*	1.62*	1.08*	1.24*
	(0.16)	(0.13)	(0.15)	(0.19)	(0.14)	(0.14)
Constant	−10.70*	−11.04*	−10.66*	−11.14*	−10.73*	−10.72*
	(1.03)	(0.87)	(0.75)	(0.80)	(1.08)	(0.66)
delta_wk	−1.10*	−1.18*	−1.13*	−1.10*	−0.93*	−0.85*
	(0.28)	(0.29)	(0.18)	(0.24)	(0.28)	(0.28)
delta_ww	1.31*	1.36*	1.36*	1.34*	1.36*	1.30*
	(0.64)	(0.53)	(0.33)	(0.37)	(0.56)	(0.51)
delta_ws	−0.02	−0.02	−0.02	−0.01	−0.02	−0.37*
	(0.07)	(0.05)	(0.11)	(0.09)	(0.07)	(0.10)
rho	0.68*	0.71*	0.69*	0.90*	0.71*	0.87*
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
delta_wk2					0.18	
					(2.75)	

(*Continues*)

TABLE IV—*Continued*

	Baseline	Favors Wal-Mart	Regional Advantage	Personal Income	Rival Stores in Neighborhood	All Other Discount Stores
Small stores' profit/all other discount stores' profit						
Log population	1.53*	1.57*	1.50*	1.45*	1.52*	1.75*
	(0.06)	(0.07)	(0.06)	(0.07)	(0.06)	(0.06)
Log retail sales	1.15*	1.19*	1.14*	1.12*	1.17*	1.34*
	(0.06)	(0.07)	(0.05)	(0.05)	(0.05)	(0.04)
Urban ratio	−1.42*	−1.46*	−1.38*	−1.55*	−1.44*	−0.73*
	(0.13)	(0.14)	(0.14)	(0.12)	(0.14)	(0.10)
South	0.92*	0.96*	0.91*	0.87*	0.92*	0.77*
	(0.06)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)
Constant_88	−9.71*	−10.01*	−9.57*	−9.32*	−9.75*	−11.73*
	(0.46)	(0.63)	(0.48)	(0.42)	(0.37)	(0.36)
delta_sk	−0.99*	−0.98*	−0.97*	−0.63*	−0.98*	−0.76*
	(0.15)	(0.13)	(0.16)	(0.12)	(0.13)	(0.12)
delta_sw	−0.93*	−0.94*	−0.93*	−0.63*	−0.96*	−0.95*
	(0.13)	(0.14)	(0.15)	(0.13)	(0.18)	(0.12)
delta_ss	−2.31*	−2.39*	−2.26*	−2.26*	−2.32*	−2.24*
	(0.09)	(0.10)	(0.09)	(0.11)	(0.09)	(0.10)
tao	0.58*	0.68*	0.54*	0.67*	0.61*	0.26*
	(0.12)	(0.11)	(0.10)	(0.15)	(0.10)	(0.10)
Constant_78	−8.62*	−8.86*	−8.50*	−7.80*	−8.60*	−10.14*
	(0.50)	(0.60)	(0.63)	(0.60)	(0.47)	(0.42)
Sunk cost	−1.80*	−1.86*	−1.80*	−2.07*	−1.90*	−2.32*
	(0.33)	(0.25)	(0.34)	(0.35)	(0.42)	(0.26)
Function value	120.26	120.77	136.74	155.65	119.62	96.05
Observations	2065	2065	2065	2065	2065	2065

[a]Asterisks (*) denote significance at the 5% confidence level daggers and (†) denote significance at the 10% confidence level. Standard errors are in parentheses. Midwest and South are regional dummies, with the Great Lakes region, the Plains region, and the Rocky Mountain region grouped as the Midwest, and the Southwest region and the Southeast region grouped as the South. delta_kw, delta_ks, delta_wk, delta_ws, delta_sk, delta_sw, and delta_ss denote the competition effect, while delta_kk and delta_ww denote the chain effect. "k" stands for Kmart, "w" stands for Wal-Mart, and "s" stands for small stores in the first five columns, and all discount stores (except Kmart and Wal-Mart stores) in the last column. $\sqrt{1-\rho^2}$ measures the importance of the market-level profit shocks. In the first three columns, the parameters are estimated using the equilibrium most profitable for Kmart, the equilibrium most profitable for Wal-Mart, and the equilibrium that grants Kmart advantage in the Midwest region and Wal-Mart advantage in the South, respectively. In the last three columns, the parameters are estimated using the equilibrium that is most favorable for Kmart. In the fourth column, log of personal income per capita is used in Kmart's and Wal-Mart's profit function. In the fifth column, the existence of rival stores in neighboring markets matters. The sixth column estimates the model using Kmart, Wal-Mart, and all other discount stores, not just small stores.

sales.[35] In the fourth column, I estimated the model using personal income per

[35]According to annual reports for Kmart (1988, 1997) and Wal-Mart (1988, 1997), the combined sales of Kmart and Wal-Mart accounted for about 2% of U.S. retail sales in 1988 and 4% in 1997. As I do not observe sales at the store level, I cannot directly measure how much a Kmart or a Wal-Mart store contributes to the total retail sales in the counties where it is located. However, given that there are on average 400–500 retailers per county and that these two firms only

TABLE V

PARAMETER ESTIMATES FROM DIFFERENT SPECIFICATIONS–1997

	Baseline	Favors Wal-Mart	Regional Advantage	Personal Income	Rival Stores in Neighborhood	All Other Discount Stores
Kmart's profit						
Log population	1.50*	1.45*	1.42*	1.34*	1.50*	1.65*
	(0.11)	(0.21)	(0.14)	(0.10)	(0.10)	(0.09)
Log retail sales/log personal income	2.16*	2.08*	2.17*	2.06*	2.16*	2.14*
	(0.16)	(0.13)	(0.13)	(0.09)	(0.09)	(0.08)
Urban ratio	1.36*	1.43*	1.41*	1.79*	1.25*	1.47*
	(0.23)	(0.41)	(0.24)	(0.28)	(0.20)	(0.42)
Midwest	0.38*	0.42*	0.33†	0.37*	0.35†	0.36*
	(0.13)	(0.20)	(0.18)	(0.15)	(0.18)	(0.12)
Constant	−24.26*	−23.47*	−24.20*	−25.04*	−24.26*	−24.70*
	(1.59)	(0.69)	(0.87)	(0.73)	(0.59)	(0.61)
delta_kw	−0.74*	−0.77*	−0.59*	−0.96*	−0.67*	−0.64*
	(0.19)	(0.25)	(0.14)	(0.18)	(0.31)	(0.23)
delta_kk	0.63	0.69	0.85*	0.56*	0.64	0.51
	(0.54)	(0.53)	(0.32)	(0.27)	(0.55)	(0.33)
delta_ks	−0.03	−0.002	−0.003	−0.02	−0.01	−0.07
	(0.20)	(0.18)	(0.08)	(0.09)	(0.12)	(0.08)
delta_kw2					0.27	
					(1.99)	
Wal-Mart's profit						
Log population	2.02*	1.97*	2.00*	2.31*	2.01*	2.01*
	(0.08)	(0.11)	(0.14)	(0.16)	(0.15)	(0.12)
Log retail sales/log personal income	1.99*	1.93*	1.99*	1.82*	2.00*	1.94*
	(0.06)	(0.08)	(0.12)	(0.08)	(0.12)	(0.08)
Urban ratio	1.63*	1.71*	1.57*	1.74*	1.48*	1.64*
	(0.29)	(0.20)	(0.63)	(0.34)	(0.36)	(0.24)
Log distance	−1.06*	−1.03*	−1.07*	−1.10*	−1.05*	−1.00*
	(0.10)	(0.15)	(0.16)	(0.09)	(0.11)	(0.04)
South	0.88*	0.94*	0.81*	0.99*	0.88*	0.93*
	(0.20)	(0.21)	(0.21)	(0.11)	(0.13)	(0.13)
Constant	−16.95*	−16.53*	−16.68*	−18.38*	−16.95*	−16.58*
	(0.76)	(0.87)	(1.08)	(0.95)	(1.20)	(0.51)
delta_wk	−0.68*	−0.74*	−0.59*	−0.68*	−0.53†	−0.87*
	(0.26)	(0.34)	(0.16)	(0.21)	(0.27)	(0.18)
delta_ww	0.79*	0.76	0.86*	0.77*	0.73†	0.76*
	(0.36)	(0.50)	(0.33)	(0.29)	(0.41)	(0.23)
delta_ws	−0.10	−0.10	−0.12†	−0.06	−0.10	−0.28*
	(0.13)	(0.07)	(0.07)	(0.08)	(0.17)	(0.08)
rho	0.86*	0.86*	0.90*	0.85*	0.88*	0.90*
	(0.06)	(0.08)	(0.05)	(0.04)	(0.06)	(0.05)
delta_wk2					0.10	
					(3.46)	

(*Continues*)

TABLE V—*Continued*

	Baseline	Favors Wal-Mart	Regional Advantage	Personal Income	Rival Stores in Neighborhood	All Other Discount Stores
Small stores' profit/all other discount stores' profit						
Log population	1.64*	1.62*	1.67*	1.66*	1.65*	1.92*
	(0.10)	(0.08)	(0.10)	(0.09)	(0.11)	(0.07)
Log retail sales	1.37*	1.33*	1.38*	1.37*	1.37*	1.37*
	(0.07)	(0.07)	(0.06)	(0.06)	(0.08)	(0.06)
Urban ratio	−1.87*	−1.76*	−1.91*	−1.95*	−1.88*	−0.80*
	(0.18)	(0.17)	(0.19)	(0.13)	(0.17)	(0.11)
South	1.14*	1.11*	1.13*	1.19*	1.13*	0.89*
	(0.09)	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)
Constant_97	−11.75*	−11.46*	−11.84*	−11.75*	−11.76*	−12.35*
	(0.61)	(0.52)	(0.43)	(0.77)	(0.68)	(0.42)
delta_sk	−0.45*	−0.44*	−0.41†	−0.43*	−0.39†	−0.38*
	(0.15)	(0.15)	(0.22)	(0.15)	(0.21)	(0.12)
delta_sw	−0.79*	−0.71*	−0.64*	−0.78*	−0.72*	−0.96*
	(0.17)	(0.14)	(0.15)	(0.15)	(0.16)	(0.12)
delta_ss	−2.68*	−2.64*	−2.75*	−2.73*	−2.69*	−2.69*
	(0.19)	(0.11)	(0.14)	(0.21)	(0.21)	(0.10)
tao	0.57*	0.53*	0.63*	0.61*	0.60*	0.11
	(0.21)	(0.19)	(0.24)	(0.17)	(0.16)	(0.13)
Constant_78	−9.62*	−9.33*	−9.48*	−9.98*	−9.56*	−9.77*
	(0.65)	(0.63)	(0.73)	(1.25)	(0.93)	(0.54)
Sunk cost	−2.36*	−2.31*	−2.50*	−1.90*	−2.40*	−2.69*
	(0.40)	(0.44)	(0.62)	(0.78)	(0.60)	(0.30)
Function value	108.68	105.02	103.90	216.24	104.64	91.24
Observations	2065	2065	2065	2065	2065	2065

[a] Asterisks (*) denote significance at the 5% confidence level and daggers (†) denote significance at the 10% confidence level. Standard errors are in parentheses. See Table IV for the explanation of the variables and the different specifications for each column.

capita in place of the retail sales variable.[36] Neither the competition effects nor the chain effects change much. The objective function value is higher, indicating a worse fit of the data.

The model assumes that stores in different markets do not compete with each other. However, it is possible that a chain store becomes a stronger competitor when it is surrounded by a large number of stores owned by the same firm in nearby markets. As a result, rival stores in neighboring markets can in-

accounted for 2–4% of total retail sales, the endogeneity of the retail sales is not likely to be a severe problem.

[36] I did not use personal income per capita in small stores' profit function, because it does not explain variations in the number of small stores very well. In the ordinary least squares regression of the number of small stores on market size variables, personal income per capita is not significant once population is included.

directly affect competition between stores in a given market. The fifth column estimates the following profit function for chain stores:

$$\Pi_{i,m}(D_i, D_{j,m}, N_{s,m}; X_m, Z_m, \varepsilon_m, \eta_{i,m})$$
$$= D_{i,m} * \left[X_m\beta_i + \delta_{ij}D_{j,m} * \left(1 + \delta_{ij,2}\sum_{l\neq m}\frac{D_{j,l}}{Z_{ml}}\right) + \delta_{is}\ln(N_{s,m}+1) \right.$$
$$\left. + \delta_{ii}\sum_{l\neq m}\frac{D_{i,l}}{Z_{ml}} + \sqrt{1-\rho^2}\varepsilon_m + \rho\eta_{i,m}\right], \quad i, j \in \{k, w\},$$

where the competition effect $\delta_{ij}D_{j,m}$ is augmented by $\delta_{ij,2}\sum_{l\neq m}(D_{j,l}/Z_{ml})$, which is the distance weighted number of rival stores in the nearby markets. Neither $\delta_{kw,2}$ nor $\delta_{wk,2}$ is significant. The magnitude is also small: on average, the competition effect is only raised by 2–3% due to the presence of surrounding stores.

In the rest of this section, I focus on the coefficients of the market size variables β. I discuss the competition effects and the chain effects in the next section.

The β coefficients are highly significant and intuitive, with the exception of the Midwest dummy, which is marginally significant in two specifications in 1997. ρ is smaller than 1, indicating the importance of the market-level error terms and the necessity of controlling for endogeneity of all firms' entry decisions.

Tables VI and VII display the model's goodness of fit for the baseline specification.[37] In Table VI, the first and third columns display the sample averages, while the other two columns list the model's predicted averages. The model matches exactly the observed average numbers of Kmart and Wal-Mart stores. The number of small firms is a noisy variable and is much harder to predict. Its sample median is around 3 or 4, but the maximum is 20 in 1978, 25 in 1988,

TABLE VI

MODEL'S GOODNESS OF FIT FOR THE BASELINE SPECIFICATION

	1988		1997	
Number of	Sample Mean	Model Mean	Sample Mean	Model Mean
Kmart	0.21	0.21	0.19	0.19
Wal-Mart	0.32	0.32	0.48	0.48
Small stores in 1978	4.75	4.80	4.75	4.74
Small stores	3.79	3.78	3.46	3.39

[37]The results for the rest of the specifications are available upon request.

TABLE VII

CORRELATION BETWEEN MODEL PREDICTION AND SAMPLE OBSERVATION

Number of	1988	1997
Kmart	0.66	0.63
Wal-Mart	0.72	0.75
Small stores in 1978	0.61	0.61
Small stores	0.65	0.67

and 19 in 1997. The model does a decent job of fitting the data. The sample average is 4.75, 3.79, and 3.46 per county in 1978, 1988, and 1997, respectively. The model's prediction is 4.80, 3.78, and 3.39, respectively.[38] Such results might be expected, as the parameters are chosen to match these moments. In Table VII, I report the correlations between the predicted and observed numbers of Kmart stores, Wal-Mart stores, and small firms in each market. The numbers vary between 0.61 and 0.75. These correlations are not included in the set of moment functions, and a high value indicates a good fit. Overall, the model explains the data well.

To investigate the differences across equilibria, Table VIII reports the percentage of markets where the two extreme equilibria differ. It turns out that these equilibria are not very different from each other. For example, in 1988, using the baseline parameter estimates, the equilibrium most profitable for Kmart and the equilibrium most profitable for Wal-Mart differ in only 1.41%

TABLE VIII

PERCENTAGE OF MARKETS WHERE THE TWO EXTREME EQUILIBRIA DIFFER[a]

	1988	1997
Using parameters associated with the equilibrium most profitable for Kmart	1.41%	1.58%
Using parameters associated with the equilibrium most profitable for Wal-Mart	1.20%	2.03%
Using parameters associated with the equilibrium that favors Wal-Mart in the South	1.45%	1.11%

[a]For each of these exercises, I solve the two extreme equilibria (the one most profitable for Kmart and the one most profitable for Wal-Mart) evaluated at the same set of parameter values, compute their difference, and average over 300 simulations.

[38]I have estimated the three-stage model twice. The first time, I used data in 1978 for the pre-chain period and data in 1988 for the post-chain period. The second time, I used data in 1978 and data in 1997 for the pre- and post-chain periods, respectively. Therefore, the model has two predictions for the number of small stores in 1978, one from each estimation. In both cases, the model's prediction comes very close to the sample mean.

of the markets.[39] As all equilibria are bounded between these two extreme equilibria, the difference between any pair of equilibria can only be (weakly) smaller.

In the absence of the chain effect, the only scenario that accommodates multiple equilibria is when both a Kmart store and a Wal-Mart store are profitable as the only chain store in the market, but neither is profitable when both stores are in the market.[40] Accordingly, the two possible equilibrium outcomes for a given market are Kmart in and Wal-Mart out or Kmart out and Wal-Mart in.[41] Using the baseline parameter estimates, on average, this situation arises in 1.1% of the sample in 1988 and 1.4% of the sample in 1997. These findings suggest that while multiple equilibria are potentially an issue, they do not represent a prevalent phenomenon in the data.[42,43] It also suggests that using different profit functions for different firms helps to reduce the occurrence of multiple equilibria in this entry model, because the more asymmetric firms are in any given market, the less likely the event occurs where both firms are profitable as the only chain store, but neither is profitable when both operate in the market.

To understand the magnitudes of the market size coefficients, I report in Tables IX, X, and XI the changes in the number of each type of store when some market size variable changes using the estimates from the baseline specifications.[44] For example, to derive the effect of population change on the number of Kmart stores, I fix Wal-Mart's and small stores' profits, increase Kmart's profit in accordance with a 10% increase in population, and resolve the full model to obtain the new equilibrium number for 300 simulations. For each of these counterfactual exercises, the columns labeled Favors Kmart use the equilibrium that is most favorable for Kmart, while the columns labeled Favors

[39]The numbers reported here are the average over 300 simulations.

[40]With the chain effect, all four cases—(Kmart out, Wal-Mart out), (Kmart out, Wal-Mart in), (Kmart in, Wal-Mart out), and (Kmart in, Wal-Mart in)—can be the equilibrium outcome for a given market. Consider an example with two markets. In market A, $\Pi_k^A = -0.2 - 0.6D_w^A$ and $\Pi_w^A = -0.2 + 0.3D_w^B - 0.7D_k^A$; in market B, $\Pi_k^B = 0.1 - 0.6D_w^B$ and $\Pi_w^B = 0.1 + 0.3D_w^B - 0.7D_k^B$. One can verify that there are two equilibria in this game. The first is ($D_k^A = 0, D_w^A = 0; D_k^B = 1, D_w^B = 0$) and the second is ($D_k^A = 0, D_w^A = 1; D_k^B = 0, D_w^B = 1$). In this simple example, both (Kmart out, Wal-Mart out) and (Kmart out, Wal-Mart in) can be the equilibrium outcome for market A.

[41]The discussion on multiple equilibria has ignored the small stores, as the number of small stores is a well defined function of a given pair of (Kmart, Wal-Mart).

[42]In their study of banks' adoption of the automated clearinghouse electronic payment system, Ackerberg and Gowrisankaran (2007) also found that the issue of multiple equilibria is not economically significant.

[43]As one referee pointed out, multiple equilibria could potentially be more important if the sample is not restricted to small- and medium-sized counties. The exercise here has taken entry decisions and the benefit derived from stores located in the metropolitan areas as given. It is possible that multiple equilibria will occur more frequently once these entry decisions are endogenized.

[44]To save space, results from other specifications are not reported here. They are not very different from those of the baseline specification.

TABLE IX

NUMBER OF KMART STORES WHEN THE MARKET SIZE CHANGES[a]

	1988				1997			
	Favors Kmart		Favors Wal-Mart		Favors Kmart		Favors Wal-Mart	
	Percent	Total	Percent	Total	Percent	Total	Percent	Total
Base case	100.0	437	100.0	413	100.0	393	100.0	362
Population increases 10%	110.5	482	110.9	458	113.1	445	113.5	411
Retail sales increases 10%	116.8	510	117.4	485	118.8	467	119.4	432
Urban ratio increases 10%	107.2	468	107.6	445	105.4	415	105.6	382
Midwest = 0 for all counties	82.7	361	81.8	338	84.6	333	84.5	306
Midwest = 1 for all counties	123.7	540	124.0	512	118.7	467	119.2	432

[a] For each of the simulation exercises in all Tables IX–XI, I fix other firms' profits and change only the profit of the target firm in accordance with the change in the market size. I resolve the entire game to obtain the new equilibrium numbers of firms. Columns labeled Favors Kmart use the equilibrium most profitable for Kmart, and columns labeled Favors Wal-Mart use the equilibrium most profitable for Wal-Mart. For example, in the second row of Table IX, I increase Kmart's profit according to a 10% increase in population while holding Wal-Mart's and small firms' profit the same as before. Using this new set of profits and the equilibrium that favors Kmart most, the number of Kmart stores is 10.5% higher than in the base case in 1988.

Wal-Mart uses the other extreme equilibrium. They provide an upper (lower) and lower (upper) bound for the number of Kmart (Wal-Mart) stores. It should not come as a surprise that results of these two scenarios are quite similar, since

TABLE X

NUMBER OF WAL-MART STORES WHEN THE MARKET SIZE CHANGES[a]

	1988				1997			
	Favors Kmart		Favors Wal-Mart		Favors Kmart		Favors Wal-Mart	
	Percent	Total	Percent	Total	Percent	Total	Percent	Total
Base case	100.0	651	100.0	680	100.0	985	100.0	1016
Population increases 10%	108.6	707	108.2	736	107.4	1058	106.9	1086
Retail sales increases 10%	110.3	718	109.9	747	107.3	1057	106.8	1085
Urban ratio increases 10%	105.4	686	105.2	715	102.2	1007	102.1	1037
Distance increases 10%	91.2	594	91.5	622	96.0	946	96.3	978
South = 0 for all counties	63.6	414	65.5	445	83.8	825	85.0	863
South = 1 for all counties	135.7	884	134.9	917	117.8	1160	116.3	1182

[a] See the footnote to Table IX for comments.

TABLE XI

NUMBER OF SMALL FIRMS WHEN THE MARKET SIZE CHANGES[a]

	1988				1997			
	Favors Kmart		Favors Wal-Mart		Favors Kmart		Favors Wal-Mart	
	Percent	Total	Percent	Total	Percent	Total	Percent	Total
Base case	100.0	7808	100.0	7803	100.0	6995	100.0	6986
Population increases 10%	106.6	8319	106.6	8314	106.3	7437	106.3	7427
Retail sales increases 10%	104.9	8191	104.9	8186	105.3	7365	105.3	7355
Urban ratio increases 10%	98.2	7665	98.2	7660	97.6	6827	97.6	6817
South = 0 for all counties	80.6	6290	80.6	6285	78.3	5476	78.3	5467
South = 1 for all counties	120.8	9431	120.8	9425	123.3	8625	123.3	8612
Sunk cost increases 10%	95.9	7485	95.9	7481	95.6	6689	95.6	6680

[a]See the footnote to Table IX for comments.

the two equilibria are not very different. In the following discussion, I focus on the equilibrium most profitable for Kmart.

Market size variables are important for big chains. In 1988, a 10% growth in population induces Kmart to enter 10.5% more markets and Wal-Mart to enter 8.6% more markets. A similar increment in retail sales attracts the entry of Kmart and Wal-Mart stores in 16.8% and 10.3% 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 33.1% fewer Kmart stores, and changing the southern regional dummy from 1 to 0 for all counties leads to 53.2% fewer Wal-Mart stores. When distance increases by 10%, the number of Wal-Mart stores drops by 8.8%. Wal-Mart's "home advantage" is much smaller in 1997: everything else the same, changing the southern dummy from 1 to 0 for all counties leads to 29% fewer Wal-Mart stores, and a 10% increase in distance reduces the number of Wal-Mart stores by only 4%. As the model is static in nature (all Kmart and Wal-Mart stores are opened in one period), the regional dummies and the distance variable provide a reduced-form way to capture the path dependence of the expansion of chain stores.

The market size variables have a relatively modest impact on the number of small businesses. In 1988, a 10% increase in population attracted 6.6% more stores. The same increase in real retail sales per capita draws 4.9% more stores. The number of small stores declines by about 1.8% 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 to 0 for all counties leads to 33.3% fewer small stores (6290 stores vs. 9431 stores). When the sunk cost increases by 10%, the number of small stores reduces by 4.1%.

7.2. *The Competition Effect and the Chain Effect*

As shown in Tables IV and V, all competition effects in the profit function of the small stores and that of all other discount stores are precisely estimated. The chain effect and the competition effect in Wal-Mart's profit function are also reasonably well estimated. The results for Kmart's profit function appear to be the weakest: although the size of the coefficients is similar, the standard errors are large for some columns. For example, the chain effect is significant in 4 out of 6 specifications in 1988 and in only two specifications in 1997. The competition effect of Wal-Mart on Kmart is big and significant in all cases in 1997, but insignificant in two specifications in 1988. The impact of small stores on the chain stores is never very significant. With one exception in 1997, both τ and the sunk cost are significant and sizeable, indicating the importance of history dependence.

To assess the magnitude of the competition effects for the chains, Table XII resolves the equilibrium number of Kmart and Wal-Mart stores under different assumptions of the market structure. The negative impact of Kmart's presence

TABLE XII

COMPETITION EFFECT AND CHAIN EFFECT FOR KMART (KM) AND WAL-MART (WM)[a]

	1988		1997	
Number of	Percent	Total	Percent	Total
Kmart stores				
Base case	100.0	437	100.0	393
Wm in each market	85.1	371	82.2	323
Wm exits each market	108.6	474	141.9	558
Not compete with small stores	101.3	442	104.3	410
No chain effect	94.7	414	93.5	368
Wal-Mart stores				
Base case	100.0	651	100.0	985
Km in each market	61.4	400	82.2	809
Km exits each market	119.5	778	105.7	1042
Not compete with small stores	101.7	662	105.1	1035
No chain effect	84.4	550	92.9	915

[a]Base case is the number of stores observed in the data. For each exercise, I resolve the full model under the specified assumptions. For the last two rows of both panels where the counterfactual exercise involves multiple equilibria, I solve the model using the equilibrium that is most profitable for Kmart.

on Wal-Mart's profit is much stronger in 1988 than in 1997, while the opposite is true for the effect of Wal-Mart's presence on Kmart's profit. For example, in 1988, Wal-Mart would only enter 400 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 778, a net increase of 94.5%. The same experiment in 1997 leads Wal-Mart to enter 28.8% more markets, from 809 to 1042. The pattern is reversed for Kmart. In 1988, Kmart would enter 27.8% more markets when there were no Wal-Mart stores compared with the case of one Wal-Mart store in every county (474 Kmart stores vs. 371 Kmart stores); in 1997, Kmart would enter 72.8% more markets for the same experiment (558 Kmart stores vs. 323 Kmart stores).[45] These estimates are consistent with the observation that Wal-Mart grew stronger during the sample period and replaced Kmart as the largest discounter in 1991.

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 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) cited 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.

To investigate the importance of the chain effect for both chains, the last row of both panels in Table XII reports 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 resolve the model. The difference in the number of stores with or without δ_{ii} captures the advantage of chains over single-unit retailers. In 1988, without the chain effect, the number of Kmart stores would have decreased by 5.3% and Wal-Mart would have entered 15.6% fewer markets. In 1997, Kmart would have entered 6.5% fewer markets, while Wal-Mart would have entered 7.1%. The decline in Wal-Mart's chain effect 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.

As I do not observe the stores' sales or profit, I cannot estimate the dollar value of these spillover benefits. However, given the low markup of these discount stores (the average gross markup was 20.9% from 1993 to 1997, see footnote 3), these estimates appear to be large. The results are consistent with Holmes (2005), who also found scale economies to be important. Given the

[45] 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.

TABLE XIII

NUMBER OF SMALL STORES WITH DIFFERENT MARKET STRUCTURE[a]

	Profit Positive		Profit Recovers Sunk Cost	
	Percent	Total	Percent	Total
		1988		
No Kmart or Wal-Mart	100.0	9261		
Only Kmart in each Market	76.2	7057	47.9	4440
Only Wal-Mart in each Market	77.5	7173	49.1	4542
Both Kmart and Wal-Mart	56.1	5195	31.6	2925
		1997		
No Kmart or Wal-Mart	100.0	8053		
Only Kmart in each Market	89.8	7228	54.1	4357
Only Wal-Mart in each Market	82.4	6634	47.9	3854
Both Kmart and Wal-Mart	72.9	5868	40.3	3244

[a] I fix the number of Kmart and Wal-Mart stores as specified and solve for the equilibrium number of small stores. If stores have perfect foresight, the columns labeled Profit Recovers Sunk Cost would have been the number of stores that we observe, as they would not have entered in the pre-chain period if their profit after entry could not recover the sunk cost.

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.[46]

Table XIII studies the competition effects of chains on small discount stores. Here I distinguish between two cases. The first two columns report the number of small stores predicted by the model, where small stores continue their business after the entry of Kmart and Wal-Mart as long as their profit is positive, even if they cannot recover the sunk cost paid in the first stage. The second two columns report the number of small stores whose post-chain profit is higher than the sunk cost. If small stores had perfect foresight and could predict the entry of Kmart and Wal-Mart, these two columns would be the number of stores that we observe. The results suggest that chains have a substantial competition impact on small firms. In 1988, compared with the scenario with no chain stores, adding a Kmart store to each market reduces the number of small firms by 23.8% or 1.07 stores per county. Of the remaining stores, more than one-third could not recover their sunk cost of entry. Had they learned of the entry of the chains stores in the first stage, they would not have entered the market. Thus, adding a Kmart store makes 52.1% of the small stores or 2.33 stores per county either unprofitable or unable to recover their sunk cost. The story is similar for the entry of Wal-Mart stores. When both a Kmart and a Wal-Mart store enter, 68.4% of the small stores or 3.07 stores per county cannot recoup their sunk cost of entry.

[46] See Foster, Haltiwanger, and Krizan (2002) for a detailed study of the productivity growth in the retail industry.

TABLE XIV

NUMBER OF ALL DISCOUNT STORES (EXCEPT FOR KMART AND WAL-MART STORES) WITH DIFFERENT MARKET STRUCTURE[a]

	Profit Positive		Profit Recovers Sunk Cost	
	Percent	Total	Percent	Total
		1988		
No Kmart or Wal-Mart	100.0	10,752		
Only Kmart in each Market	82.7	8890	47.1	5064
Only Wal-Mart in each Market	78.5	8443	43.6	4692
Both Kmart and Wal-Mart	62.7	6741	31.5	3383
		1997		
No Kmart or Wal-Mart	100.0	9623		
Only Kmart in each Market	91.9	8842	51.7	4976
Only Wal-Mart in each Market	79.8	7683	42.0	4043
Both Kmart and Wal-Mart	72.4	6964	36.5	3508

[a] I fix the number of Kmart and Wal-Mart stores as specified and solve for the number of all other discount stores. See the additional comments in the footnote to Table XIII.

Looking at the discount industry as a whole, the impact of Kmart and Wal-Mart remains significant, although Kmart's impact is slightly diminished in 1997. Table XIV shows that when a Wal-Mart store enters a market in 1988, 21.5% of the discount firms will exit the market and 56.4% of the firms cannot recover their sunk cost. These numbers translate to 1.1 stores and 2.9 stores per county, respectively.

It is somewhat surprising that the negative impact of Kmart on other firms' profit is comparable to Wal-Mart's impact, 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.

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 XV records the net decrease of 693 small firms observed over the sample period or 0.34 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 resolve the model using the 1988 coefficients and the 1988 market size variables for Kmart's and small firms' profit functions, but the 1997 coefficients and 1997 market size variables for Wal-Mart's profit function. The experiment corresponds to holding small stores and

TABLE XV

THE IMPACT OF WAL-MART'S EXPANSION[a]

	1988	1997
Observed decrease in the number of small stores between 1988 and 1997	693	693
Predicted decrease from the full model	380	259
Percentage explained	55%	37%
Observed decrease in the number of all discount stores (except for Kmart and Wal-Mart stores) between 1988 and 1997	1021	1021
Predicted decrease from the full model	416	351
Percentage explained	41%	34%

[a] In the top panel, the predicted 380 store exits in 1988 are obtained by simulating the change in the number of small stores using Kmart's and the small stores' profit in 1988, but Wal-Mart's profit in 1997. The column of 1997 uses Kmart's and small stores' profit in 1997, but Wal-Mart's profit in 1988. Similarly for the second panel.

Kmart the same as in 1988, but allowing Wal-Mart to become more efficient and expand. The predicted number of small firms falls by 380. This accounts for 55% of the observed decrease in the number of small firms. Conducting the same experiment but using the 1997 coefficients and the 1997 market size variables for Kmart's and small firms' profit functions, and the 1988 coefficients and 1988 market size variables for Wal-Mart's profit function, I find that Wal-Mart's expansion accounts for 259 stores or 37% of the observed decrease in the number of small firms.

Repeating the same exercise using all discount stores, the prediction is similar: roughly 30–40% of store exits can be attributed to the expansion of Wal-Mart stores. Overall, the absolute impact of Wal-Mart's entry seems modest. However, the exercise here only looks at 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, and apparel stores, so their impact on local communities is conceivably much larger.

I tried various specifications that group retailers in different sectors, for example, all retailers in the discount sector, the building materials sector, and the home-furnishing sector. None of these experiments was successful, as the retailers in different sectors differ substantially and the simple model cannot match the data very well. Perhaps a better approach is to use a separate profit function for firms in each sector and estimate the system of profit functions jointly. This is beyond the scope of this paper and is left for future research.

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 XVI indicate that direct subsidies do not seem to be effective in generating jobs. In 1988, subsidizing Wal-Mart stores 10% of their average profit increases the number of Wal-Mart

TABLE XVI

THE IMPACT OF GOVERNMENT SUBSIDIES: CHANGES IN THE NUMBER OF JOBS IN THE DISCOUNT SECTOR[a]

	1988	1997
Subsidize Kmart's profit by 10%		
Increase in Kmart's employees	4	4
Decrease in other stores' employees	−1	−1
Subsidize Wal-Mart's profit by 10%		
Increase in Wal-Mart's employees	7	8
Decrease in other stores' employees	−1	−1
Subsidize small stores' profit by 100%		
Increase in small stores' employees	13	12
Decrease in other stores' employees	0	−2
Subsidize all other discount stores' profit by 100%		
Increase in other discount stores' employees	40	34
Decrease in Kmart and Wal-Mart stores' employees	−6	−4

[a]For each of these counterfactual exercises, I incorporate the change in the subsidized firm's profit as specified, solve for the equilibrium numbers of stores, and obtain the estimated change in employment assuming that (a) a Kmart or a Wal-Mart store employs 300 employees, (b) a small discount store employs 10 employees, and (c) an average discount store employs 25 employees.

stores per county from 0.32 to 0.34.[47,48] With the average Wal-Mart store hiring fewer than 300 full- and part-time employees, the additional number of stores translates to roughly 7 new jobs. Wal-Mart's expansion crowds out other stores, which brings the net increase down to 6 jobs. Similarly, subsidizing all small firms by 100% of their average profit increases their number from 3.78 to 5.07, and generates 13 jobs if, on average, a small firm hires 10 employees. Repeating the exercise with subsidizing all discount stores (except for Kmart and Wal-Mart stores) by 100% of their average profit leads to a net increase of 34 jobs. Together, these exercises suggest that a direct subsidy does not seem to be very effective in generating employment in this industry. 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.

[47]The average Wal-Mart store's net income in 1988 is about 1 million in 2004 dollars according to its Securities and Exchange Commission annual report. Using a discount rate of 10%, the discounted present value of a store's lifetime profit is about 10 million. A subsidy of 10% is roughly 1 million dollars.

[48]In this exercise, I first simulate the model 300 times, obtain the mean profit for all Wal-Mart stores for each simulation, and average it across simulations. Then I increase Wal-Mart's profit by 10% of this average (that is, I add this number to the constant of Wal-Mart's profit function) and simulate the model 300 times to obtain the number of Wal-Mart stores after the subsidy.

8. CONCLUSION AND FUTURE WORK

In this paper, I have examined the competition effect between Kmart stores, Wal-Mart stores, and other discount stores, as well as the role of the chain effect in firms' entry decisions. The negative impact of Kmart's presence on Wal-Mart's profit is much stronger in 1988 than in 1997, while the opposite is true for the effect of Wal-Mart's presence on Kmart's profit. On average, entry by either a Kmart or a Wal-Mart store makes 48–58% of the discount stores (2–3 stores) either unprofitable or unable to recover their sunk cost. Wal-Mart's expansion from the late 1980s to the late 1990s explains 37–55% of the net change in the number of small discount stores and 34–41% of the net change in the number of all discount stores.

Like Holmes (2005), I find that scale economies, as captured by the chain effect, generate substantial benefits. Without the spillover effect, the number of Kmart stores would have decreased by 5.3% in 1988 and 6.5% in 1997, while Wal-Mart would have entered 15.6% fewer markets in 1988 and 7.1% fewer markets in 1997. Studying these scale economies in more detail is useful 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 insights to the productivity gains in the retail industry over the past several decades.

Finally, the algorithm used in this paper can be applied to industries where scale economies are important. One 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.

APPENDIX A: 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 census data.[49] 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 zip codes 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.[50] For the sample counties, only 30–60 stores were not matched between these two sources for either 1988 or 1997.

[49] Marie Pees from the Census Bureau kindly provided these data.

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

APPENDIX B: DEFINITIONS AND PROOFS

B.1. *Verification of the Necessary Condition (3)*

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

$$\begin{aligned}&\Pi(D_1^*, \ldots, D_{m-1}^*, D_m^*, D_{m+1}^*, \ldots, D_M^*)\\&\quad \geq \Pi(D_1^*, \ldots, D_{m-1}^*, D_m, D_{m+1}^*, \ldots, D_M^*) \quad \forall m, D_m^* \neq D_m.\end{aligned}$$

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

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

where $Z_{ml} = Z_{lm}$ due to symmetry. Since $\Pi(D^*) - \Pi(\hat{D}) \geq 0$, $D_m^* \neq D_m$, it must be that $D_m^* = 1$ and $D_m = 0$ if and only if $X_m + 2\delta \sum_{l\neq m}(D_l^*/Z_{ml}) \geq 0$, and $D_m^* = 0$ and $D_m = 1$ if and only if $X_m + 2\delta \sum_{l\neq m}(D_l^*/Z_{ml}) < 0$. Together we have $D_m^* = 1[X_m + 2\delta \sum_{l\neq m}(D_l^*/Z_{ml}) \geq 0]$.[51]

B.2. *The Set of Fixed Points of an Increasing Function That Maps a Lattice Into Itself*

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. The profit functions are

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

[51] I have implicitly assumed that when $X_m + 2\delta \sum_{l\neq m}(D_l^*/Z_{ml}) = 0$, $D_m^* = 1$.

Let $D = \{D_k, D_w, D_s\} \in \mathbf{D} = \{\mathbf{0}, \mathbf{1}\}^3$, let D_{-i} denote rivals' strategies, let $V_i(D_{-i})$ denote the best response function for player i, and let $V(D) = \{V_k(D_{-k}), V_w(D_{-w}), V_s(D_{-s})\}$ denote the joint best response function. It is easy to show that $V(D)$ is a decreasing function that takes the values

$$\begin{aligned} &V(0,0,0) = \{1,1,1\}; \quad &&V(0,0,1) = \{1,0,1\};\\ &V(0,1,0) = \{0,1,1\}; \quad &&V(0,1,1) = \{0,0,1\},\\ &V(1,0,0) = \{1,1,0\}; \quad &&V(1,0,1) = \{1,0,0\};\\ &V(1,1,0) = \{0,1,0\}; \quad &&V(1,1,1) = \{0,0,0\}.\end{aligned}$$

The set of fixed points of $V(D)$ is empty.

B.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:

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

$$\text{s.t. if } D_m = 1, \text{ then } X_m + \delta \sum_{l\neq m} \frac{D_l}{Z_{ml}} > 0.$$

The solution to this constrained maximization problem belongs to the set of fixed points of the vector function $\hat{V}(D) = \{\hat{V}_1(D), \dots, \hat{V}_M(D)\}$, where $\hat{V}_m(D) = 1[X_m + \delta \sum_{l\neq m}(D_l/Z_{ml}) > 0]$. When $\delta > 0$, the function $\hat{V}(\cdot)$ is increasing and maps from $\mathbf{D}$ into itself: $\hat{V}: \mathbf{D} \to \mathbf{D}$. Let $\hat{D}$ denote the convergent vector using $\sup(\mathbf{D})$ as the starting point for the iteration on $\hat{V}$: $\hat{V}(\hat{D}) = \hat{D}$. Using arguments similar to those in Section 5.1, one can show that $\hat{D}$ is the greatest element among the set of $\hat{V}$'s fixed points. Further, $\hat{D}$ achieves a higher profit than any other fixed point of $\hat{V}(\cdot)$, since by construction each nonzero element of the vector $\hat{D}$ adds to the total profit. Changing any nonzero 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^* cannot 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 profit at markets with $D^*_m = 1$ does not decrease after the change, and profit at markets with $D^*_m = 0$ but $\hat{D}_m = 1$ is positive by construction. This contradicts the definition of D^*, so $\hat{D} \leq D^*$.

Note that $V(\hat{D}) \geq \hat{V}(\hat{D}) = \hat{D}$, where $V(\cdot)$ is defined in Section 5.1. As in Section 5.1, iterating V on both sides of the inequality $V(\hat{D}) \geq \hat{D}$ generates an increasing sequence. Denote the convergent vector as $\hat{D}^T$. This is a tighter lower bound of D^* than D^L (discussed in Section 5.1) because $\hat{D}^T = V^{TT}(\hat{D}) \geq V^{TT}(\inf(\mathbf{D})) = D^L$, 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 and $\delta \sum_{l \neq m} \frac{1}{Z_{ml}}$, it is never optimal to enter markets that contribute a negative element to the total profit even with the largest conceivable chain effect. Let $\tilde{D} = \{\tilde{D}_m : \tilde{D}_m = 0 \text{ if } X_m + 2\delta \sum_{l \neq m}(1/Z_{ml}) < 0; \tilde{D}_m = 1 \text{ otherwise}\}$. We know that $\tilde{D} \geq D^*$. Using the argument above, the convergent vector $\tilde{D}^T$ from iterating V on $\tilde{D}$ is a tighter upper bound to D^* than D^U.

B.4. *Verification That the Chains' Profit Functions Are Supermodular With Decreasing Differences*

DEFINITION 3: Suppose that $Y(X)$ is a real-valued function on a lattice $\mathbf{X}$. If

$$Y(X') + Y(X'') \leq Y(X' \vee X'') + Y(X' \wedge X'') \tag{7}$$

for all X' and X'' in $\mathbf{X}$, then $Y(X)$ is supermodular on $\mathbf{X}$.[52]

DEFINITION 4: 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 in the equation system (2) 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. The profit function is simplified to $\Pi = \sum_{m=1}^{M}[D_m * (X_m + \delta\sum_{l \neq m}(D_l/Z_{ml}))]$. First it is easy to show that $D' \vee D'' = (D' - \min(D', D'')) +$

[52]Both definitions are taken from Chapter 2 of Topkis (1998).

$(D'' - \min(D', D'')) + \min(D', D'')$ and $D' \wedge D'' = \min(D', D'')$. Let $D' - \min(D', D'')$ be denoted as D_1, denote $D'' - \min(D', D'')$ as D_2, and denote $\min(D', D'')$ as D_3. The left-hand side of the inequality (7) is

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

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

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

The profit function is supermodular in its own strategy if the chain effect δ is nonnegative. To verify that the profit function Π_i has decreasing differences in (D_i, D_j), write

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

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

B.5. *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 B.3 illustrates a tighter upper bound 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 50 miles. Markets that are 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 50 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 10 markets:

	1	2	3	
4	5	6	7	8
	9	10		

,

where

$$D^U = \begin{array}{|c|c|c|c|c|} \cline{2-4} \multicolumn{1}{c|}{} & 1 & D_2 & 1 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & D_6 & 1 & 1 \\ \hline \multicolumn{1}{c|}{} & D_9 & D_{10} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\ \cline{2-3} \end{array}$$

and

$$D^L = \begin{array}{|c|c|c|c|c|} \cline{2-4} \multicolumn{1}{c|}{} & 0 & D_2 & 0 & \multicolumn{1}{c}{} \\ \hline 0 & 0 & D_6 & 0 & 0 \\ \hline \multicolumn{1}{c|}{} & D_9 & D_{10} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\ \cline{2-3} \end{array}.$$

D^U and D^L are the same in markets 2, 6, 9, and 10, but differ for the rest. If markets 1, 4, and 5 (group A) are at least 50 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 C_N. Let ξ_m denote the random shocks in the profit function. By construction, $D_m^U = 1[X_m + 2\delta \sum_{l\neq m, l\in B_m} (D_l^U / Z_{ml}) + \xi_m \geq 0]$ and $D_m^L = 1[X_m + 2\delta \sum_{l\neq m, l\in B_m} (D_l^L / Z_{ml}) + \xi_m \geq 0]$. The probability of $D_m^U = 1, D_m^L = 0$ for every market in the size-N connected area C_N is

$$\begin{aligned} &\Pr(D_m^U = 1, D_m^L = 0, \forall m \in C_N) \\ &\quad \leq \prod_{m=1}^{N} \Pr\left(X_m + \xi_m < 0, X_m + \xi_m + 2\delta \sum_{l\neq m, l\in B_m} \frac{1}{Z_{ml}} \geq 0 \right), \end{aligned}$$

where ξ_m is assumed to be i.i.d. across markets. As δ goes to infinity, the probability approaches $\prod_{m=1}^{N} \Pr(X_m + \xi_m < 0)$ from below. How fast it decreases when N increases depends on the distribution assumption. If ξ_m is i.i.d. normal and X_m is i.i.d. uniformly distributed between $[-a, a]$, with a a finite positive number, on average, the probability is on the magnitude of $(\frac{1}{2})^N$:

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

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.

REFERENCES

ACKERBERG, D., AND G. GOWRISANKARAN (2007): "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *RAND Journal of Economics*, 37, 738–761. [1267,1298]

ANDREWS, D., S. BERRY, AND P. JIA (2004): "Confidence Regions for Parameters in Discrete Games With Multiple Equilibria," Unpublished Manuscript, Yale University. [1267,1288]

ARCHER, J., AND D. TAYLOR (1994): *Up Against the Wal-Marts (How Your Business Can Survive in the Shadow of the Retail Giants)*. New York: American Management Association. [1263]

ATHEY, S. (2002): "Monotone Comparative Statics Under Uncertainty," *Quarterly Journal of Economics*, 117, 187–223. [1267]

BAJARI, P., AND J. FOX (2005): "Complementarities and Collusion in an FCC Spectrum Auction," Unpublished Manuscript, University of Chicago. [1276]

BAJARI, P., H. HONG, AND S. RYAN (2007): "Identification and Estimation of Discrete Games of Complete Information," Unpublished Manuscript, Duke University. [1267]

BASKER, E. (2005a): "Job Creation or Destruction? Labor-Market Effects of Wal-Mart Expansion," *Review of Economics and Statistics*, 87, 174–183. [1268,1270]

——— (2005b): "Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail Prices," *Journal of Urban Economics*, 58, 203–229. [1268]

BERRY, S. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60, 889–917. [1267,1287,1288]

BRESNAHAN, T., AND P. REISS (1990): "Entry Into Monopoly Markets," *Review of Economic Studies*, 57, 531–553. [1267,1287]

——— (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 95, 57–81. [1267,1287]

CHAIN STORE GUIDE (1988–1997): *Directory of Discount Department Stores*. New York: Business Guides Inc. [1270-1272,1291]

CHERNOZHUKOV, V., H. HONG, AND E. TAMER (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284. [1267,1288]

CILIBERTO, F., AND E. TAMER (2006): "Market Structure and Multiple Equilibria in Airline Markets," Unpublished Manuscript, Northwestern University. [1267]

CONLEY, T. (1999): "GMM Estimation With Cross Sectional Dependence," *Journal of Econometrics*, 92, 1–45. [1265,1284-1286]

CONLEY, T., AND E. LIGON (2002): "Economic Distance and Cross Country Spillovers," *Journal of Economic Growth*, 7, 157–187. [1286]

DAVIS, P. (2006): "Spatial Competition in Retail Markets: Movie Theaters," *RAND Journal of Economics*, 37, 964–994. [1267]

Discount Merchandiser (1988–1997). New York: Schwartz Publications. [1264]

FOSTER, L., J. HALTIWANGER, AND C. J. KRIZAN (2002): "The Link Between Aggregate and Micro Productivity Growth: Evidence From Retail Trade," Working Paper 9120, NBER. [1303]

HAUSMAN, J., AND E. LEIBTAG (2005): "Consumer Benefits From Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart," Working Paper 11809, NBER. [1268]

HESS, S., AND J. POLAK (2003): "A Comparison of Scrambled and Shuffled Halton Sequences for Simulation Based Estimation," Unpublished Manuscript, Imperial College London. [1287]

HOLMES, T. (2001): "Barcodes Lead to Frequent Deliveries and Superstores," *RAND Journal of Economics*, 32, 708–725. [1302]

——— (2005): "The Diffusion of Wal-Mart and Economies of Density," Unpublished Manuscript, University of Minnesota. [1265,1266,1289,1302,1307]

HOUSE COMMITTEE ON SMALL BUSINESS (1994): "Impact of Discount Superstores on Small Business and Local Communities," Committee Serial No. 103-99, Congressional Information Services, Inc. [1269]

JIA, P. (2008): "Supplement to 'What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry'," *Econometrica Supplemental Material*, 76, http://www.econometricsociety.org/ecta/Supmat/6649_data and programs.zip. [1268]

KMART INC. (1988–2000): Annual Report. [1293]

MAZZEO, M. (2002): "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33, 1–22. [1267]

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026. [1284]

MILGROM, P., AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180. [1267,1280]

NEUMARK, D., J. ZHANG, AND S. CICCARELLA (2005): "The Effects of Wal-Mart on Local Labor Markets," Working Paper 11782, NBER. [1268]

NEWEY, W., AND K. WEST (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708. [1286]

NISHIDA, M. (2008): "Estimating a Model of Strategic Store Network Choice With Policy Simulation," Unpublished Manuscript, University of Chicago. [1266]

PAKES, A., AND D. POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057. [1284]

PAKES, A., J. PORTER, K. HO, AND J. ISHII (2005): "Moment Inequalities and Their Application," Unpublished Manuscript, Harvard University. [1267,1288]

PINKSE, J., M. SLADE, AND C. BRETT (2002): "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70, 1111–1153. [1267]

ROMANO, J., AND A. SHAIKH (2006): "Inference for the Identified Set in Partially Identified Econometric Models," Unpublished Manuscript, University of Chicago. [1267,1288]

SEIM, K. (2006): "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," *RAND Journal of Economics*, 37, 619–640. [1267]

SHILS, E. B. (1997): *The Shils Report: Measuring the Econometric and Sociological Impact of the Mega-Retailer. Discount Chains on Small Enterprises in Urban, Suburban and Rural Communities*. Philadelphia, PA: Wharton School, University of Pennsylvania. [1264]

SMITH, H. (2004): "Supermarket Choice and Supermarket Competition in Market Equilibrium," *Review of Economic Studies*, 71, 235–263. [1265,1267]

STONE, K. (1995): "Impact of Wal-Mart Stores on Iowa Communities: 1983–93," *Economic Development Review*, 13, 60–69. [1268]

TAMER, E. (2003): "Incomplete Simultaneous Discrete Response Model With Multiple Equilibria," *Review of Economic Studies*, 70, 147–165. [1267]

TARSKI, A. (1955): "A Lattice-Theoretical Fixpoint and Its Applications," *Pacific Journal of Mathematics*, 5, 285–309. [1267,1279]

TOPKIS, D. (1978): "Minimizing a Submodular Function on a Lattice," *Operations Research*, 26, 305–321. [1267,1280,1281]

——— (1998): *Supermodularity and Complementarity*. Princeton, NJ: Princeton University Press. [1281,1310]

TRAIN, K. (2000): "Halton Sequences for Mixed Logit," Unpublished Manuscript, UC Berkeley. [1287]

——— (2003): *Discrete Choice Methods With Simulation*. Cambridge, U.K.: Cambridge University Press. [1287]

U.S. CENSUS BUREAU (1993–1997): *Annual Benchmark Report for Retail Trade and Food Services*. Washington, DC: U.S. Government Printing Office. [1264]

——— (1997): *Establishment and Firm Size (Including Legal Form of Organization)*, Economics Census Retail Trade Series. Washington, DC: U.S. Government Printing Office. [1264]

VANCE, S., AND R. SCOTT (1994): *A History of Sam Walton's Retail Phenomenon*. New York: Twayne Publishers. [1269,1271]

WAL-MART STORES, INC. (1970–2000): Annual Report. [1293]

ZHOU, L. (1994): "The Set of Nash Equilibria of a Supermodular Game Is a Complete Lattice," *Games and Economic Behavior*, 7, 295–300. [1281]

ZHU, T., AND V. SINGH (2007): "Spatial Competition With Endogenous Location Choices: An Application to Discount Retailing," Unpublished Manuscript, University of Chicago. [1268]

Dept. of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142, U.S.A. and NBER; pjia@mit.edu.

Manuscript received August, 2006; final revision received January, 2008.