MARKETING SCIENCE
Vol. 27, No. 5, September–October 2008, pp. 811–828
ISSN 0732-2399 | EISSN 1526-548X | 08 | 2705 | 0811

DOI 10.1287/mksc.1080.0398

Supermarket Pricing Strategies

Paul B. Ellickson
Department of Economics, Duke University, Durham, North Carolina 27708,
paul.ellickson@duke.edu

Sanjog Misra
William E. Simon School of Business Administration, University of Rochester,
Rochester, New York 14627, misra@simon.rochester.edu

Most supermarket firms choose to position themselves by offering either everyday low prices (EDLP) across several items or offering temporary price reductions (promotions) on a limited range of items. While this choice has been addressed from a theoretical perspective in both the marketing and economic literature, relatively little is known about how these decisions are made in practice, especially within a competitive environment. This paper exploits a unique store level data set consisting of every supermarket operating in the United States in 1998. For each of these stores, we observe the pricing strategy the firm has chosen to follow, as reported by the firm itself. Using a system of simultaneous discrete choice models, we estimate each store's choice of pricing strategy as a static discrete game of incomplete information. In contrast to the predictions of the theoretical literature, we find strong evidence that firms cluster by strategy by choosing actions that agree with those of its rivals. We also find a significant impact of various demographic and store/chain characteristics, providing some qualified support for several specific predictions from marketing theory.

Key words: EDLP; promotional pricing; positioning strategies; supermarkets; discrete games
History: Received: March 22, 2006; accepted: February 27, 2008; processed by David Bell.

1. Introduction

While firms compete along many dimensions, pricing strategy is clearly one of the most important. In many retail industries, pricing strategy can be characterized as a choice between offering relatively stable prices across a wide range of products (often called everyday low pricing) or emphasizing deep and frequent discounts on a smaller set of goods (referred to as promotional or PROMO pricing). Although Wal-Mart did not invent the concept of everyday low pricing, the successful use of everyday low pricing (EDLP) was a primary factor in their rapid rise to the top of the Fortune 500, spawning a legion of followers selling everything from toys (Toys Я Us) to building supplies (Home Depot). In the 1980s, it appeared that the success and rapid diffusion of the EDLP strategy could spell the end of promotions throughout much of retail. However, by the late 1990s, the penetration of EDLP had slowed, leaving a healthy mix of firms following both strategies, and several others employing a mixture of the two.

Not surprisingly, pricing strategy has proven to be a fruitful area of research for marketers. Marketing scientists have provided both theoretical predictions and empirical evidence concerning the types of consumers that different pricing policies are likely to attract (e.g. Lal and Rao 1997, Bell and Lattin 1998). While we now know quite a bit about where a person is likely to shop, we know relatively little about how pricing strategies are chosen by retailers. There are two primary reasons for this. First, these decisions are quite complex: managers must balance the preferences of their customers and their firm's own capabilities against the expected actions of their rivals. Empirically modeling these actions (and reactions) requires formulating and then estimating a complex discrete game, an exercise which has only recently become computationally feasible. The second is the lack of appropriate data. While scanner data sets have proven useful for analyzing consumer behavior, they typically lack the breadth necessary for tackling the complex mechanics of inter-store competition.[1] The goal of this paper is to combine newly developed methods for estimating static games with a rich, national data set on store level pricing policies to identify the primary factors that drive pricing behavior in the supermarket industry.

Exploiting the game theoretic structure of our approach, we aim to answer three questions that have not been fully addressed in the existing literature. First, to what extent do supermarket chains tailor their pricing strategies to local market conditions? Second, do certain types of chains or stores

[1] Typical scanner data usually reflect decisions made by only a few stores in a limited number of markets.

have advantages when it comes to particular pricing strategies? Finally, how do firms react to the expected actions of their rivals? We address each of these questions in detail.

The first question naturally invites a market pull driven explanation in which consumer demographics play a key role in determining which pricing strategy firms choose. In answering this question, we also aim to provide additional empirical evidence that will inform the growing theoretical literature on pricing related games. Since we are able to assess the impact of local demographics at a much broader level than previous studies, our results provide more conclusive evidence regarding their empirical relevance.

The second question concerns the match between a firm's strategy and its chain-specific capabilities. In particular, we examine whether particular pricing strategies (e.g., EDLP) are more profitable when firms make complementary investments (e.g. larger stores and more sophisticated distribution systems). The empirical evidence on this matter is scant—this is the first paper to address this issue on a broad scale. Furthermore, because our data set includes all existing supermarkets, we are able to exploit variation both within and across chains to assess the impact of store and chain level differences on the choice of pricing strategy.

Finally, we address the role of competition posed in our third question by analyzing firms' reactions to the expected choices of their rivals. In particular, we ask whether firms face incentives to distinguish themselves from their competitors (as in most models of product differentiation) or instead face pressures to conform (as in network or switching cost models)? This question is the primary focus of our paper and the feature that most distinguishes it from earlier work.

Our results shed light on all three questions. First, we find that consumer demographics play a significant role in the choice of local pricing strategies: firms choose the policy that their consumers demand. Furthermore, the impact of these demographic factors is consistent with both the existing marketing literature and conventional wisdom. For example, EDLP is favored in low income, racially diverse markets, while PROMO clearly targets the rich. However, a key implication of our analysis is that these demographic factors act as a coordinating device for rival firms, helping shape the pricing landscape by defining an equilibrium correspondence. Second, we find that complementary investments are key: larger stores and vertically integrated chains are significantly more likely to adopt EDLP. Finally, and most surprisingly, we find that stores competing in a given market have incentives to coordinate their actions. Rather than choosing a pricing strategy that distinguishes them from their rivals, stores choose strategies that match. This finding is in direct contrast to existing theoretical models that view pricing strategy as a form of differentiation, providing a clear comparative static that future pricing models must address.

Our paper makes both substantive and methodological contributions to the marketing literature. On the substantive front, our results offer an in-depth look at the supermarket industry's pricing practices, delineating the role of three key factors (demand, supply, and competition) on the choice of pricing strategy. We provide novel, producer-side empirical evidence that complements various consumer-side models of pricing strategy. In particular, we find qualified support for several claims from the literature on pricing demographics, including Bell and Lattin's (1998) model of basket size and Lal and Rao's (1997) positioning framework, while at the same time highlighting the advantages of chain level investment. Our focus on competition also provides a structural complement to Shankar and Bolton's (2004) descriptive study of price variation in supermarket scanner data, which emphasized the role of rival actions. Our most significant contribution, however, is demonstrating that stores in a particular market do not use pricing strategy as a differentiation device but instead coordinate their actions. This result provides a direct challenge to the conventional view of retail competition, opening up new and intriguing avenues for future theoretical research. Our econometric implementation also contributes to the growing literature in marketing and economics on the estimation of static discrete games, as well as the growing literature on social interactions.[2] In particular, our incorporation of multiple sources of private information and our construction of competitive beliefs are novel additions to these emerging literatures.

The rest of the paper is organized as follows. Section 2 provides an overview of the pricing landscape, explicitly defining each strategy and illustrating the importance of local factors in determining store level decisions. Section 3 introduces our formal model of pricing strategy and briefly outlines our estimation approach. Section 4 describes the data set. Section 5 provides the details of how we implement the model, including the construction of distinct geographic markets, the selection of covariates, our two-step estimation method, and our identification strategy. Section 6

[2] Recent applications of static games include technology adoption by internet service providers (Augereau et al. 2006), product variety in retail eyewear (Watson 2005), location of ATM branches (Gowrisankaran and Krainer 2004), and spatial differentiation among supermarkets (Orhun 2005), discount stores (Zhu et al. 2005), and video stores (Seim 2006). Structural estimation of social interactions is the focus of papers by Brock and Durlauf (2002), Bayer and Timmins (2006), and Bajari et al. (2005).

provides our main empirical results and discusses their implications. Section 7 concludes with directions for future research.

2. The Supermarket Pricing Landscape

2.1. Pricing Strategy Choices

Competition in the supermarket industry is a complex phenomenon. Firms compete across the entire retail and marketing mix, enticing customers with an attractive set of products, competitive prices, convenient locations, and a host of other services, features, and promotional activities. In equilibrium, firms choose the bundle of services and features that maximize profits, conditional on the types of consumers they expect to serve and their beliefs about the actions of their rivals. A supermarket's pricing strategy is a key element in this multidimensional bundle.

The majority of both marketers and practitioners frame a store's pricing decision as a choice between offering everyday low prices or deep but temporary discounts, labeling the first strategy EDLP and the second PROMO (Table 1).[3,4] Not surprisingly, the simple EDLP-PROMO dichotomy is too narrow to adequately capture the full range of firm behavior. In practice, firms can choose a mixture of EDLP and PROMO, varying either the *number* of categories they put on sale or changing the *frequency* of sales across some or all categories of products. Practitioners have coined a term for these practices—hybrid pricing. What constitutes HYBRID pricing is necessarily subjective, depending on an individual's own beliefs regarding how much price variation constitutes a departure from pure EDLP. Both the data and definitions used in this paper are based on a specific store level survey conducted by Trade Dimensions in 1998, which asked individual store managers to choose which of the following categories best described their store's pricing policy:

- **Everyday Low Price (EDLP):** Little reliance on promotional pricing strategies such as temporary price cuts. Prices are consistently low across the board, throughout all packaged food departments.
- **Promotional (Hi-Lo) Pricing:** Heavy use of specials, usually through manufacturer price breaks or special deals.
- **Hybrid EDLP/Hi-Lo:** Combination of EDLP and Hi-Lo pricing strategies.

According to Trade Dimensions, the survey was designed to allow for a broad interpretation of the HYBRID strategy, as they wanted it to capture deviations along either the temporal (i.e., number of sales per year) or category based dimensions (i.e., number of categories on deal). We believe that pricing strategy is best viewed as a continuum, with pure EDLP (i.e., constant margins across all categories) on one end and pure PROMO (i.e. frequent sales on all categories) at the other. This data set represents a coarse discretization of that continuum.

Table 1 Descriptive Statistics

Variable	Obs	Mean	Std. dev.	Min.	Max
Strategy					
EDLP	17,388	0.28	0.45	0	1
HYBRID	17,388	0.38	0.48	0	1
PROMO	17,388	0.34	0.47	0	1
MSA characteristics					
Size (sq. miles)	333	1,868.31	1,943.99	46.40	11,229.6
Density (pop '000 per sq. mile)	333	10.42	9.62	0.91	49.06
Avg. food expenditure ($ '000)	333	663.64	1,201.37	16.04	9,582.09
Market variables					
Median household size	8,000	2.66	0.35	1.32	5.69
Median HH income	8,000	35,255.59	9,753.95	18,109.60	81,954.60
Proportion Black	8,000	0.08	0.14	0.00	0.97
Proportion Hispanic	8,000	0.06	0.13	0.00	0.98
Median vehicles in HH	8,000	2.12	0.33	0.56	3.37
Chain/store characteristics					
Vertically integrated	17,388	0.51	0.50	0.00	1.00
Store size (sqft '000)	17,388	28.99	16.34	2.00	250.00
Independent store	17,388	0.23	0.42	0.00	1.00
Number of stores in chain	804	390.15	478.45	1.00	1,399.00

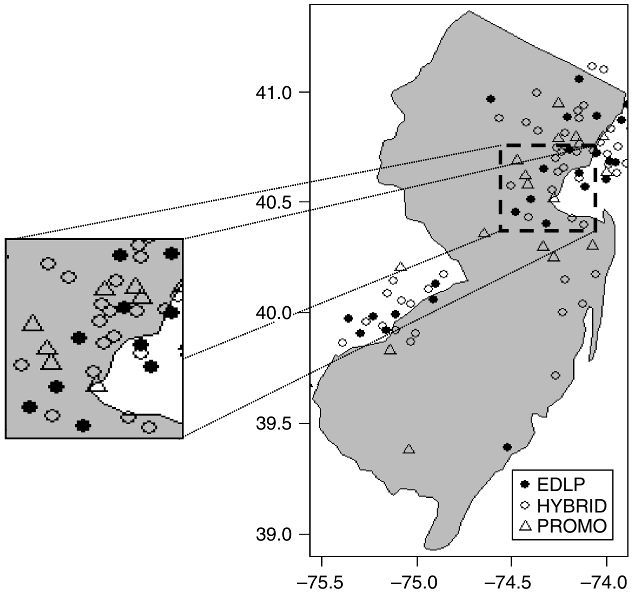
2.2. Supermarket Pricing: A Closer Look

Without observing data on individual stores, it might be tempting to conclude that all pricing strategies are determined at the level of the chain. While there are certainly incentives to choose a consistent policy, the data reveals a remarkable degree of local heterogeneity. To examine the issue more closely, we focus in on a single chain in a single market: the Pathmark chain in New Jersey. Figure 1 shows the spatial locations of

[3] This is clearly a simplification—a supermarket's pricing policy is closely tied to its overall positioning strategy. Pricing strategies are typically chosen to leverage particular operational advantages and often have implications for other aspects of the retail mix. For example, successful implementation of EDLP may involve offering a deeper and narrower product line than PROMO, allowing firms to exploit scale economies (in particular categories), reduce their inventory carrying costs, and lower their advertising expenses. On the other hand, PROMO pricing gives firms greater flexibility in clearing overstock, allows them to quickly capitalize on deep manufacturer discounts, and facilitates the use of consumer loyalty programs (e.g. frequent shopper cards). In other words, the choice of pricing strategy is more than just how prices are set: it reflects the overall positioning of the store. This paper focuses on the pricing dimension alone, taking the other aspects of the retail mix as given. While this is limiting, modeling the entire retail mix is beyond the scope of this paper.

[4] Note that we focus on the choice of pricing *strategy* and abstract away from issues related to more *tactical* decisions about how prices are (or should be) set (see e.g., Kumar and Rao 2006).

Figure 1 **Pathmark Stores in New Jersey**



every Pathmark store in New Jersey, along with its pricing strategy. Two features of the data are worth emphasizing. We address them in sequence.

First, Pathmark does not follow a single strategy across its stores: 42% of the stores use PROMO pricing, 33% follow EDLP, and the remaining 25% use HYBRID. The heterogeneity in pricing strategy observed in the Pathmark case is not specific to this particular chain. Table 2 shows the store level strategies chosen by the top 15 U.S. supermarkets (by total volume) along with their total store counts. As with Pathmark, the major chains are also surprisingly heterogeneous. While some firms do have a clear focus (e.g. Wal-Mart, H.E. Butt, Stop & Shop), others are more evenly split (e.g. Lucky, Cub Foods). This pattern extends to the full set of firms. Table 3 shows the pricing strategies chosen by large and small chains, using four alternative definitions of "large" and small.[5] While large chains seem evenly distributed across the strategies and "small" chains seem to favor PROMO, firm size is not the primary determinant of pricing strategy.

The second noteworthy feature of the Pathmark data is that even geographically proximate stores adopt quite different pricing strategies. While there is some clustering at the broader spatial level (e.g. north versus south New Jersey), the extent to which these strategies are interlaced is striking. Again, looking beyond Pathmark and New Jersey confirms that this within-chain spatial heterogeneity is not unique to this particular example: while some chains clearly favor a consistent strategy, others appear quite responsive to local factors. Broadly speaking, the data reveal only a weak relationship between geography and pricing strategy. While southern chains such as Food Lion are widely perceived to favor EDLP and Northeastern chains like Stop & Shop are thought to prefer PROMO, regional variation does not capture the full story. Table 4 shows the percent of stores that choose either EDLP, HYBRID, or PROMO pricing in eight geographic regions of the United States. While PROMO pricing is most popular in the Northeast, Great Lakes, and central Southern regions, it is far from dominant, as both the EDLP and HYBRID strategies enjoy healthy shares there as well. EDLP is certainly favored in the South and Southeast, but PROMO still draws double digit shares in both regions. This heterogeneity in pricing strategy can be illustrated using the spatial structure of our data set. Figure 2 plots the geographic location of every store in the United States, along with their pricing

Table 2 **Pricing Strategies of the Top 15 Supermarkets**

Firm	Stores	% PROMO	% HYBRID	% EDLP
Kroger	1,399	47	40	13
Safeway	1,165	52	43	5
Albertson's	922	11	41	48
Winn-Dixie	1,174	3	30	67
Lucky	813	35	38	27
Giant	711	29	60	11
Fred Meyer	821	22	60	18
Wal-Mart	487	1	26	73
Publix	581	13	71	16
Food Lion	1,186	2	12	86
A&P	698	55	30	15
H.E. Butt	250	1	3	96
Stop & Shop	189	50	43	7
Cub foods	375	26	34	40
Pathmark	135	42	25	33

Table 3 **Pricing Strategy by Firm Type**

	% EDLP	% HYBRID	% PROMO
"Large" firms:			
Chain	33	37	30
Vertically integrated	35	36	29
Large store size	32	38	30
Many checkouts	31	39	30
"Small" firms:			
Independent	22	28	50
Not vertically integrated	21	32	47
Small store size	23	26	52
Few checkouts	22	26	52

[5] The four definitions of firm size are: chain/independent, vertically integrated and not, large/small store, and many/few checkouts. A chain is defined as having 11 or more stores, while an independent has 10 of fewer. Vertically integrated means the firm operates its own distribution centers. Large versus small store size and many versus few checkouts are defined by the upper and lower quartiles of the full store level census.

Table 4 Pricing Strategies by Region

Region	% PROMO	% HYBRID	% EDLP
West Coast	39	39	22
Northwest	32	51	17
South West	20	48	32
South	32	25	43
Southern Central	45	27	28
Great Lakes	54	29	17
North East	40	37	23
South East	23	37	40

strategy. As is clear from the three panels corresponding to each pricing strategy, there is no obvious pattern: all three strategies exhibit quite uniform coverage. Taken together, these observations suggest looking elsewhere for the primary determinants of pricing

Figure 2 Spatial Distribution of Store Pricing Strategy

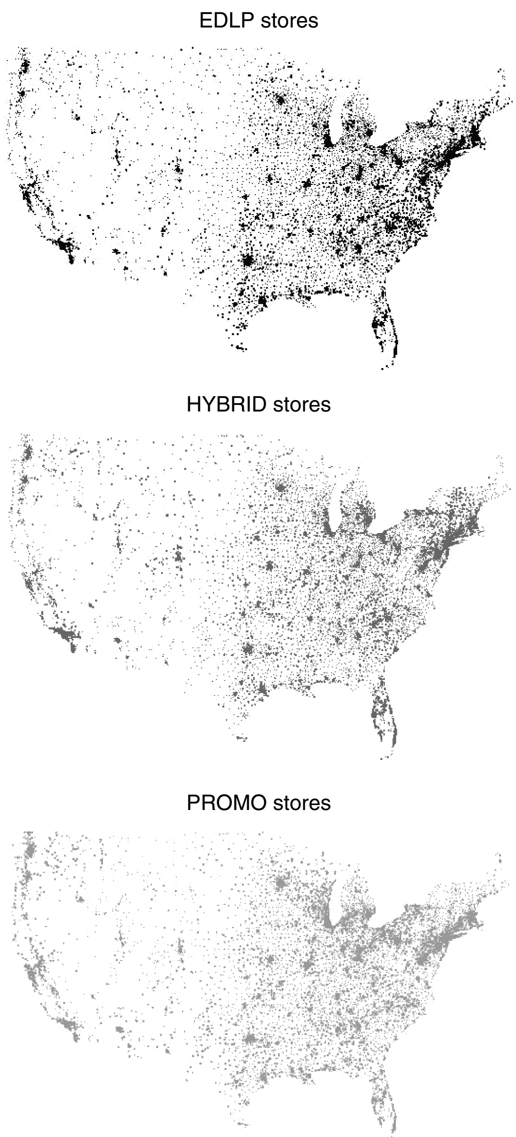


Table 5 Local Factors

	EDLP	HYBRID	PROMO
Local demographics			
Median household size	2.84 (0.331)	2.81 (0.337)	2.80 (0.329)
Median household income	34,247 (14,121)	36,194 (15,121)	36,560 (16,401)
Median vehicles in HH	2.12 (0.302)	2.13 (0.303)	2.09 (0.373)
Median age	35.4 (4.59)	35.8 (4.98)	35.7 (4.25)
Proportion Black	0.128 (0.182)	0.092 (0.158)	0.110 (0.185)
Proportion Hispanic	0.078 (0.159)	0.073 (0.137)	0.070 (0.135)
Strategies of rivals			
Percent of rivals using same strategy	49 (31)	49 (25)	52 (23)

Note. The main numbers in each cell are means, standard deviations are in parentheses.

strategy. We turn next to the role of market demographics and then to the nature and degree of competition.

Table 5 contains the average demographic characteristics of the local market served by stores of each type.[6] PROMO pricing is associated with smaller households, higher income, fewer automobiles per capita, and less racial diversity, providing some initial support for Bell and Lattin's (1998) influential model of basket size.[7] However, the differences in demography, while intuitive, are not especially strong. This does not mean that demographics are irrelevant, but rather that the *aggregate* level patterns linking pricing strategy and demographics are not overwhelming. Isolating the pure impact of demographic factors will require a formal model, which we provide below.

The final row of Table 5 contains the share of rival stores in the competing market that employ the same strategy as the store being analyzed. Here we find a striking result: 50% of a store's rivals in a given location employ the same pricing strategy as the focal store. Competitor factors also played a lead role in the work of Shankar and Bolton (2004), which analyzed pricing variability in supermarket scanner data. In particular, they note that "what is most striking, however, is that the competitor factors are the *most dominant* determinants of retailer pricing in a broad framework that included several other factors" (p. 43). Even at this rather coarse level of analysis, the data

[6] Roughly corresponding to areas the size of a ZipCode, these "local markets" are defined explicitly in §5.2.

[7] Bell and Lattin (1998) find that the most important features of shopping behavior can be captured by two interrelated choices: basket size (how much you buy) and shopping frequency (how often you go). They suggest that large or fixed basket shoppers (i.e. those who buy more and shop less) will more sensitive to the overall basket price than those who shop frequently and will therefore prefer EDLP pricing to PROMO. They present empirical evidence that is consistent with this prediction.

reveal that most stores choose similar pricing strategies to their rivals. This pattern clearly warrants a more detailed investigation and is the focus of our structural model.

Stepping back, three key findings emerge. First, supermarket chains often adopt heterogeneous pricing strategies, suggesting that demand related forces can sometimes outweigh the advantages of chain level specialization. Second, *local* market factors play a key role in shaping demand characteristics. Finally, any empirical analysis of pricing strategy must address the role of competition. While investigating the role of market demographics and firm characteristics is not conceptually difficult, quantifying the structural impact of rival pricing strategies on firm behavior requires a formal game theoretic model of pricing behavior that accounts for the simultaneity of choices. In the following section, we embed pricing strategy in a discrete game that accommodates both local demographics and the strategies of rival firms. We then estimate this model using the two-step approach developed by Bajari et al. (2005).

3. A Strategic Model of Supermarket Pricing

A supermarket's choice of pricing strategy is naturally framed as a discrete game between a finite set of players. Each firm's optimal choice is determined by the underlying market conditions, its own characteristics and relative strengths, as well as its expectations regarding the actions of its rivals. Ignoring strategic expectations, pricing strategy could be modeled as a straightforward discrete choice problem. However, since firms condition their strategies on their beliefs regarding rivals' actions, this discrete choice must be modeled as a system of simultaneous equations. In our framework, firms (i.e., supermarket chains[8]) make a discrete choice of pricing strategy, selecting among three alternatives: everyday low pricing, promotional pricing, and a hybrid strategy. While there is clearly a role for dynamics in determining an optimal pricing policy, we assume that firms act simultaneously in a static environment, taking entry decisions as given. This static treatment of competition is not altogether unrealistic since these pricing strategies involve substantial store level investments in communication and positioning related costs that are not easily reversed.[9]

We assume that competition takes place in "local" markets, each contained in a global market (here, an MSA). Before proceeding further, we must introduce some additional notation. Stores belonging to a given chain $c = 1, \ldots, C$, that are located in a local market $l_m = 1, \ldots, L_m$, in an MSA $m = 1, \ldots, M$, will be indexed using $i_c^{l_m} = 1, \ldots, N_c^{l_m}$. The total number of stores in a particular chain in a given MSA is $N_c^m = \sum_{l_m=1}^{L_m} N_c^{l_m}$, while the total number of stores in that chain across all MSAs is $N_c = \sum_{m=1}^{M} N_c^m$. In each local market, chains select a pricing strategy (action) a from the three element set $K = \{E, H, P\}$, where $E \equiv$ EDLP, $H \equiv$ HYBRID, and $P \equiv$ PROMO. If we observe a market l_m containing $N^{l_m} = \sum_{c=1}^{C} N_c^{l_m}$ players for example, the set of possible action profiles is then $A_{l_m} = \{E, H, P\}^{N_c^{l_m}}$ with generic element $a_{l_m} = (a_1, a_2, \ldots, a_{i_c^{l_m}}, \ldots, a_{N_c^{l_m}})$. The vector of actions of store $i_c^{l_m}$'s competitors is denoted $a_{-i_c^{l_m}} = (a_1, \ldots, a_{i_c^{l_m}-1}, a_{i_c^{l_m}+1}, \ldots, a_{N_c^{l_m}})$.

In a given market, a particular chain's state vector is denoted $s_c^m \in S_c^m$, while the state vector for the market as a whole is $s^m = (s_1^m, \ldots, s_{N_c}^m) \in \prod_{c=1}^{N_c^m} S_c^m$. The state vector s^m is known to all firms and observed by the econometrician. It describes features of the market and characteristics of the firms that we assume are determined exogenously. For each firm, there are also three *unobserved* state variables (corresponding to the three pricing strategies) that are treated as private information of the firm. These unobserved state variables are denoted $\epsilon_{i_c^{l_m}}(a_{i_c^{l_m}})$, or more compactly $\epsilon_{i_c^{l_m}}$, and represent firm specific shocks to the profitability of each strategy. The private information assumption makes this a game of incomplete or asymmetric information (e.g. Harsanyi 1973) and the appropriate equilibrium concept one of Bayesian Nash Equilibrium (BNE). For any given market, the $\epsilon_{i_c^{l_m}}$'s are assumed to be i.i.d. across firms and actions, and drawn from a distribution $f(\epsilon_{i_c^{l_m}})$ that is known to everyone, including the econometrician.

Firms maximize store-level profits, choose pricing strategies independently across stores. In market l_m, the profit earned by store i_c is given by

$$\pi_{i_c^{l_m}} = \Pi_{i_c^{l_m}}\left(s^m, a_{i_c^{l_m}}, a_{-i_c^{l_m}}\right) + \epsilon_{i_c^{l_m}}, \tag{1}$$

where $\Pi_{i_c^{l_m}}$ is a known and deterministic function of states and actions (both own and rival's). Since the ϵ's are private information, each firm's decision rule $a_{i_c^{l_m}} = d_{i_c^{l_m}}(s^m, \epsilon_{i_c^{l_m}})$ is a function of the common state vector and its own ϵ, but *not* the private information of its rivals. From the perspective of *both* its rivals and the econometrician, the probability that a given firm chooses action k conditional on the common state vector is then given by

$$P_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k\right) = \int 1\left\{d_{i_c^{l_m}}\left(s^m, \epsilon_{i_c^{l_m}}\right) = k\right\} f\left(\epsilon_{i_c^{l_m}}\right) d\epsilon_{i_c^{l_m}}, \tag{2}$$

where $1\{d_{i_c^{l_m}}(s, \epsilon_{i_c^{l_m}}) = k\}$ is an indicator function equal to 1 if store $i_c^{l_m}$ chooses action k and 0 otherwise.

[8] Henceforth, we will use chains and firms interchangeably.

[9] As discussed above, pricing decisions are relatively sunk, due to the positioning costs associated with conveying a consistent store-level message to a group of repeat customers. Furthermore, since this is not an entry game, we are not particularly concerned about the possibility of ex post regret that can sometimes arise in static games (Einav 2003).

We let $\mathbf{P}_{l_m}$ denote the set of these probabilities for a given local market. Since the firm does not observe its competitors actions prior to choosing its own action, it makes decisions based on its expectations. The expected profit for firm $i_c^{l_m}$ from choosing action $a_{i_c^{l_m}}$ is then

$$\begin{aligned}\tilde{\pi}_{i_c^{l_m}}\left(a_{i_c^{l_m}}, s^m, \epsilon_i, \mathbf{P}_{l_m}\right)&\\ &= \tilde{\pi}_{i_c^{l_m}}\left(a_{i_c^{l_m}}, s^m\right) + \epsilon_{i_c^{l_m}} &(3)\\ &= \sum_{a_{-i_c^{l_m}}} \Pi_{i_c^{l_m}}\left(s^m, a_{i_c^{l_m}}, a_{-i_c^{l_m}}\right) P_{-i_c^{l_m}} + \epsilon_{i_c^{l_m}}, &(4)\end{aligned}$$

where $P_{-i_c^{l_m}} = \prod_{j \neq i_c^{l_m}} P_j(a_j \mid s^m)$. Given these expected profits, the optimal action for a store is then

$$\begin{aligned}\Psi_{i_c^{l_m}} = \Pr\Big\{&\tilde{\pi}_{i_c^{l_m}}\left(a_{i_c^{l_m}}, s^m\right) + \epsilon_{i_c^{l_m}}\left(a_{i_c^{l_m}}\right)\\ &> \tilde{\pi}_{i_c^{l_m}}\left(a'_{i_c^{l_m}}, s^m\right) + \epsilon_{i_c^{l_m}}\left(a'_{i_c^{l_m}}\right)\ \forall a'_{i_c^{l_m}} \neq a_{i_c^{l_m}}\Big\}, \quad (5)\end{aligned}$$

which is the system of equations that define the (pure strategy) BNE of the game. Because a firm's optimal action is unique by construction, there is no need to consider mixed strategies.

If the ϵ's are drawn from a Type I Extreme Value distribution, this BNE must satisfy a system of logit equations (i.e. best response probability functions). The general framework described above has been applied in several economic settings and its properties are well understood. Existence of equilibrium follows directly from Brouwer's fixed point theorem.

To proceed further, we need to choose a particular specification for the expected profit functions. We will assume that the profit that accrues to store $i_c^{l_m}$ from choosing strategy k in location l_m is given by

$$\begin{aligned}\tilde{\pi}_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k, s^m, \epsilon_i, \mathbf{P}_{l_m}\right) &= s^{m\prime}\beta_k + \rho^E_{-i_c^{l_m}}\alpha_{k1} + \rho^P_{-i_c^{l_m}}\alpha_{k2}\\ &\quad + \xi_c^m(k) + \zeta_c(k) + \varepsilon_{i_c^{l_m}}(k) \quad (6)\end{aligned}$$

where s^m is the common state vector of both market (local and MSA) and firm characteristics (chain and store level). The $\rho^E_{-i_c^{l_m}}$ and $\rho^P_{-i_c^{l_m}}$ terms represent the expected proportion of a store's competitors in market l_m that choose EDLP and PROMO strategies, respectively

$$\rho^{(k)}_{-i_c^{l_m}} = \frac{1}{N^{l_m}} \sum_{j \neq i_c^{l_m}} P_j(a_j = k).$$

Note that we have assumed that payoffs are a linear function of the share of stores that choose EDLP and PROMO, which simplifies the estimation problem and eliminates the need to consider the share who choose HYBRID (H). We further normalize the average profit from the PROMO strategy to zero, one of three assumptions required for identification (we discuss our identification strategy in detail in §5.7). In addition, we have assumed that the private information available to store $i_c^{l_m}$ (i.e. $\epsilon_{i_c^{l_m}}$) can be decomposed into three additive stochastic components

$$\epsilon_{i_c^{l_m}}(k) = \xi_c^m(k) + \zeta_c(k) + \varepsilon_{i_c^{l_m}}(k), \quad (7)$$

where $\varepsilon_{i_c^{l_m}}(k)$ represents local market level private information, $\xi_c^m(k)$ is the private information that a chain possesses about a particular global market (MSA), and $\zeta_c(k)$ is a nonspatial component of private information that is chain specific. Following our earlier discussion, we assume that $\varepsilon_{i_c^{l_m}}(k)$ is an i.i.d. Gumbel error. We further assume that the two remaining components are jointly distributed with distribution function $F(\xi_c^m(k), \zeta_c(k); \Omega)$, where Ω is a set of parameters associated with F. Denoting the parameter vector $\Theta = \{\beta, \alpha, \Omega\}$ and letting $\delta_{i_c^{l_m}}(k)$ be an indicator function such that

$$\delta_{i_c^{l_m}}(k) = \begin{cases} 1 & \text{if } a_{i_c^{l_m}} = k, \\ 0 & \text{if } a_{i_c^{l_m}} \neq k, \end{cases} \quad (8)$$

the optimal choice probabilities (conditional on $\xi_c^m(k), \zeta_c(k)$) for a given store can be written as

$$\begin{aligned}&\Psi_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k \mid \Theta, \mathbf{P}_{l_m}, \mathbf{X}, \xi_c^m(k), \zeta_c(k)\right)\\ &= \frac{\exp\left(s^{m\prime}\beta_k + \rho^E_{-i_c^{l_m}}\alpha_{k1} + \rho^P_{-i_c^{l_m}}\alpha_{k2} + \xi_c^m(k) + \zeta_c(k)\right)}{\sum_{k' \in \{E,H,P\}} \exp\left(s^{m\prime}\beta_{k'} + \rho^E_{-i_c^{l_m}}\alpha_{k'1} + \rho^P_{-i_c^{l_m}}\alpha_{k'2} + \xi_c^m(k') + \zeta_c(k')\right)}\end{aligned} \quad (9)$$

while the likelihood can be constructed as

$$\begin{aligned}&\prod_{c \in C} \int_{\zeta_c(k)} \prod_{m \in M} \int_{\xi_c^m(k)} \Bigg\{ \prod_{l_m \in L_m} \prod_{i_c^{l_m} \in N_c^{l_m}} \Big[\Psi_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k \mid \Theta, \mathbf{P}_{l_m}, \mathbf{s},\right.\\ &\qquad\qquad \left.\xi_c^m(k), \zeta_c(k)\right)\Big]^{\delta_{i_c^{l_m}}(k)} \Bigg\} dF(\xi_c^m(k), \zeta_c(k); \Omega)\\ &\text{s.t.}\quad \mathbf{P}_{l_m} = \mathscr{E}_{\{\xi,\zeta\}}\left[\boldsymbol{\Psi}_{l_m}\left(\Theta, \mathbf{P}_{l_m}, \mathbf{s}, \xi_c^m(k), \zeta_c(k)\right)\right]. \quad (10)\end{aligned}$$

Note that the construction of the likelihood involves a system of discrete choice equations that must satisfy a fixed point constraint ($\mathbf{P}_{l_m} = \boldsymbol{\Psi}_{l_m}$). There are two main approaches for dealing with the recursive structure of this system, both based on methods originally applied to dynamic discrete choice problems. The first, based on Rust's (1987) Nested Fixed Point (NFXP) algorithm, involves solving for the fixed point of the system at every candidate parameter vector and then using these fixed point probabilities to evaluate the likelihood. However, the NFXP approach is both computationally demanding and straightforward to apply only when the equilibrium of the system is unique.[10] An alternate method, based on Hotz and Miller's (1993) Conditional Choice Probability (CCP) estimator, involves using a two-step approach that is both computationally light and more robust to multiplicity.[11] The first step of this procedure involves obtaining consistent estimates of each firm's beliefs regarding the strategic actions of its rivals. These "expectations" are then used in a second stage optimization procedure to obtain the structural parameters of interest. Given the complexity of our problem, we chose to adopt a two-step approach based on Bajari et al. (2005), who were the first to apply these methods to static games.

4. Data Set

The data for the supermarket industry are drawn from Trade Dimension's 1998 *Supermarkets Plus Database*, while corresponding consumer demographics are taken from the decennial Census of the United States. Descriptive statistics are presented in Table 1. Trade Dimensions collects store level data from every supermarket operating in the United States for use in their *Marketing Guidebook* and *Market Scope* publications, as well as selected issues of *Progressive Grocer* magazine. The data are also sold to marketing firms and food manufacturers for marketing purposes. The (establishment level) definition of a supermarket used by Trade Dimensions is the government and industry standard: a store selling a full line of food products and generating at least $2 million in yearly revenues. Foodstores with less than $2 million in revenues are classified as convenience stores and are not included in the data set.[12]

Information on pricing strategy, average weekly volume, store size, number of checkouts, and additional store and chain level characteristics was gathered using a survey of each store manager, conducted by their principal food broker. With regard to pricing strategy, managers are asked to choose the strategy that is closest to what their store practices on a general basis: either EDLP, PROMO or HYBRID. The HYBRID strategy is included to account for the fact that many practitioners and marketing theorists view the spectrum of pricing strategies as more a continuum than a simple EDLP-PROMO dichotomy (Shankar and Bolton 2004). The fact that just over a third of the respondents chose the HYBRID option is consistent with this perception.

5. Empirical Implementation

The empirical implementation of our framework requires three primary inputs. First, we need to choose an appropriate set of state variables. These will be the market, store and chain characteristics that are most relevant to pricing strategy. To determine which specific variables to include, we draw heavily on the existing marketing literature. Second, we will need to define what we mean by a "market." Finally, we need to estimate beliefs and construct the empirical likelihood. We outline each of these steps in the following subsections, concluding with a discussion of unobserved heterogeneity and our strategy for identification.

5.1. Determinants of Pricing Strategy

The focus of this paper is the impact of rival pricing policies on a firm's own pricing strategy. However, there are clearly many other factors that influence pricing behavior. Researchers in both marketing and economics have identified several, including consumer demographics, rival pricing behavior, and market, chain, and store characteristics (Shankar and Bolton 2004). Since we have already discussed the role of rival firms, we now focus on the additional determinants of pricing strategy.

Several marketing papers highlight the impact of demographics on pricing strategy (Ortmeyer et al. 1991, Hoch et al. 1994, Lal and Rao 1997, Bell and Lattin 1998). Of particular importance are consumer factors such as income, family size, age, and access to automobiles. In most strategic pricing models, the PROMO strategy is motivated by some form of spatial or temporal price discrimination. In the spatial models (e.g. Lal and Rao 1997, Varian 1980), PROMO pricing is aimed at consumers who are either willing or able to visit more than one store (i.e. those with low travel costs) or, more generally, those who are more informed about prices. The EDLP strategy instead targets consumers who have higher travel costs or are less informed (perhaps due to heterogeneity in the cost of acquiring price information). In the case of temporal discrimination (Bell and Lattin 1998, Bliss 1988), PROMO pricing targets customers who are willing to either delay purchase or stockpile products, while EDLP targets customers that prefer to purchase their entire basket in a single trip or at a

[10] It is relatively simple to construct the likelihood function when there is a unique equilibrium, although solving for the fixed point at each iteration can be computationally taxing. However, constructing a *proper* likelihood (for the NFXP) is generally intractable in the event of multiplicity, since it involves *both* solving for all the equilibria and specifying an appropriate selection mechanism. Simply using the first equilibrium you find will result in mispecification. A version of the NFXP that is robust to multiplicity has yet to be developed.

[11] Instead of requiring a unique equilibrium to the whole game, two-step estimators simply require a unique equilibrium be played in the data. Futhermore, if the data can be partioned into distinct markets with sufficient observations (as is the case in our application), this requirement can be weakened further.

[12] Firms in this segment operate very small stores and compete only with the smallest supermarkets (Ellickson 2006, Smith 2006).

single store. Clearly, the ability to substitute over time or across stores will depend on consumer characteristics. To account for these factors, we include measures of family size, household income, median vehicle ownership, and racial composition in our empirical analysis.

Since alternative pricing strategies will require differing levels of fixed investment (Lattin and Ortmeyer 1991), it is important to control for both store and chain level characteristics. For example, large and small chains may differ in their ability to efficiently implement particular pricing strategies (Dhar and Hoch 1997). Store level factors also play a role (Messinger and Narasimhan 1997). For example, EDLP stores may need to carry a larger inventory (to satisfy large basket shoppers), while PROMO stores might need to advertise more heavily. Therefore, we include a measure of store size and an indicator variable for whether the store is part of a vertically integrated chain. Finally, since the effectiveness of pricing strategies might vary by market size (e.g. urban versus rural), we include measures of geographic size, population density, and average expenditures on food.

5.2. Market Definition

The supermarket industry is composed of a large number of firms operating anywhere from 1 to 1,200 outlets. We focus on the choice of pricing strategy at an individual store, abstracting away from the more complex issue of how decisions are made at the level of the chain. This requires identifying the primary trading area from which each store draws potential customers. Without disaggregate, consumer-level information, the task of defining local markets requires some simplifying assumptions. In particular, we assume markets can be defined by spatial proximity alone, a strong assumption in some circumstances (Bell et al. 1998). However, absent detailed consumer level purchase information, we cannot relax this assumption further. Therefore, we will try to be as flexible as possible in defining spatial markets.

Although there are many ways to group firms using existing geographic boundaries (e.g. ZipCodes or Counties), these pre-specified regions all share the same drawback: they increase dramatically in size from east to west, reflecting established patterns of population density.[13] Rather than imposing this structure exogenously, we allow the data to sort itself by using cluster analysis. In particular, we assume that a market is a contiguous geographic area, measurable by geodesic distance and containing a set of competing stores. Intuitively, markets are groups of stores that are located close to one another. To construct these markets, we used a statistical clustering method (K-means) based on latitude, longitude, and ZipCode information.[14] Our clustering approach produced a large set of distinct clusters that we believe to be a good approximation of the actual markets in which supermarkets compete. These store clusters are somewhat larger than a typical ZipCode, but significantly smaller than the average county. As robustness checks, we experimented with the number of clusters, broader and narrower definitions of the market (e.g. ZipCodes and MSAs), as well as nearest neighbor methods and found qualitatively similar results (see Appendix B.1).

[13] One exception is Census block groups, which are about half the size of a typical ZipCode. However, we feel that these areas are too small to constitute reasonably distinct supermarket trading areas.

5.3. Estimation Strategy

As noted above, the system of discrete choice equations presents a challenge for estimation. We adopt a two stage approach based on Bajari et al. (2005). The first step is to obtain a consistent estimate of $\mathbf{P}_{l_m}$, the probabilities that appear (implicitly) on the right hand side of Equation (9).[15] These estimates ($\widehat{\mathbf{P}}_{l_m}$) are used to construct the $\rho_{-i_c^{l_m}}$'s, which are then plugged into the likelihood function. Maximization of this (pseudo) likelihood constitutes the second stage of the procedure. Consistency and asymptotic normality has been established for a broad class of two-step estimators by Newey and McFadden (1994), while Bajari et al. (2005) provide formal results for the model estimated here. We note in passing that consistency of the estimator is maintained even with the inclusion of the two random effect terms (ζ and ξ), since these variables are treated as private information of each store. A final comment relates to the construction of standard errors. Because the two-step approach precludes using the inverse information matrix, we use a bootstrap approach instead.[16]

5.4. The Likelihood

In our econometric implementation, we will assume that ζ and ξ are independent, mean zero normal errors, so that

$$\begin{aligned} &F(\xi_c^m(k), \zeta_c(k); \Omega) \\ &\quad = F_\xi(\xi_c^m(k); \Omega_\xi(k)) \times F_\zeta(\zeta_c(k); \Omega_\zeta(k)), \end{aligned} \tag{11}$$

[14] ZipCodes are required to ensure contiguity: without ZipCode information, stores in Manhattan would be included in the same market as stores in New Jersey.

[15] The $\rho_{-i_c^{l_m}}$'s are functions of $\mathbf{P}_{l_m}$.

[16] In particular, we bootstrapped across markets (not individual stores) and held the pseudorandom draws in the simulated likelihood fixed across bootstrap iterations. To save time we used the full data estimates as starting values in each bootstrap iteration.

where both F_ξ and F_ζ are mean zero normal distribution functions with finite covariance matrices. For simplicity, we also assume that the covariance matrices are diagonal with elements $\tau_\zeta^2(k)$ and $\tau_\xi^2(k)$. For identification, consistent with our earlier independence and normalization assumptions, we assume that $\xi_c^m(P) = \zeta_c(P) = 0\ \forall c \in C, m \in M$. We can then use a simulated maximum likelihood procedure that replaces (10) with its sample analog

$$\widetilde{\mathscr{L}}(\Theta) = \prod_{c\in C} R_\zeta^{-1} \sum_{r_\zeta=1}^{R_\zeta} \prod_{m\in M} \left[R_\xi^{-1} \sum_{r_\xi=1}^{R_\xi} \left\{ \prod_{l_m \in L_m} \prod_{i_c^{l_m} \in N_c^{l_m}} \left[\Psi_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k \mid \mathbf{\Theta}, \widehat{\mathbf{P}}_{l_m}, \mathbf{s}, \xi_c^m(k), \zeta_c(k)\right)\right]^{\delta_{i_c^{l_m}}(k)} \right\} \right]. \quad (12)$$

In the simulation procedure, $[\xi_c^m(k)]_{r_\xi}$ and $[\zeta_c(k)]_{r_\zeta}$ are drawn from mean zero normal densities with variances $\tau_\xi^2(k)$ and $\tau_\zeta^2(k)$ respectively. We use $R_\xi = R_\zeta = 500$ and maximize (12) to obtain estimates of the structural parameters. Note that the fixed point restriction, $\mathbf{P}_{l_m} = \mathbf{\Psi}_{l_m}$, no longer appears since we have replaced $\mathbf{P}_{l_m}$ with $\widehat{\mathbf{P}}_{l_m}$ in the formulae for $\rho^E_{-i_c^{l_m}}$ and $\rho^P_{-i_c^{l_m}}$, which are used in constructing $\Psi_{i_c^{l_m}}$ (see (9)). We now turn to estimating beliefs.

5.5. Estimating Beliefs

In an ideal setting, we could recover estimates of each store's beliefs regarding the conditional choice probabilities of its competitors using fully flexible nonparametric methods (e.g. kernel regressions or sieves). Unfortunately, our large state vector makes this infeasible. Instead, we employ a parametric approach for estimating $\hat{\rho}_{-i_c^{l_m}}$, using a mixed multinomial logit (MNL) specification to recover these first stage choice probabilities (Appendix B.4 provides a semi-parametric robustness analysis). This is essentially the same specification employed in the second stage procedure (outlined above), only the store's beliefs regarding rivals' actions are excluded from this reduced form. Note that we do not require an explicit exclusion restriction, since our specification already contains natural exclusion restrictions due to the presence of state variables that vary across stores and chains.

We implement an estimator similar to (12), but with the coefficients on the $\rho_{-i_c^{l_m}}$'s (i.e. α's) set to zero. Let the parameters in the first stage be denoted by $\Lambda_1 = \{\beta_1, \Omega_1\}$[17] and the first stage likelihood for a given store be denoted by $\mathscr{L}_{i_c^{l_m}}(\Lambda, \xi_c^m(k), \zeta_c(k))$. Using a simulated maximum likelihood (SML) approach, we obtain $\hat{\Lambda}_1$, the SML estimate of Λ_1. Given these estimates, and applying Bayes' rule, the posterior expectation of $P(a_{i_c^{l_m}} = k \mid s, \xi_c^m(k), \zeta_c(k))$ can be obtained via the following computation

$$\left(\int_\zeta \int_\xi \Psi_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k \mid \hat{\Lambda}, \xi_c^m(k), \zeta_c(k)\right) \cdot \mathscr{L}_{i_c^{l_m}}(\hat{\Lambda}, \xi_c^m(k), \zeta_c(k))\, dF(\xi_c^m(k), \zeta_c(k); \widehat{\Omega}_1(k)) \right) \cdot \left(\int_\zeta \int_\xi \mathscr{L}_{i_c^{l_m}}(\hat{\Lambda}, \xi_c^m(k), \zeta_c(k))\, dF(\xi_c^m(k), \zeta_c(k); \widehat{\Omega}_1(k)) \right)^{-1}. \quad (13)$$

While this expression is difficult to evaluate analytically, the vector of beliefs defined by

$$\hat{P}_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k\right) = \mathscr{E}_{\{\zeta,\xi\}}\left[\Psi_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k \mid \hat{\Lambda}, \xi_c^m(k), \zeta_c(k)\right)\right] \quad (14)$$

can be approximated by its simulation analog

$$\hat{P}_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k\right) \simeq \frac{\sum_{r=1}^R \Psi_{i_c^{l_m}}\left(a_{i_c^{l_m}} = k \mid \hat{\Lambda}, [\xi_c^m(k), \zeta_c(k)]_r\right) \mathscr{L}_{i_c^{l_m}}(\hat{\Lambda}, [\xi_c^m(k), \zeta_c(k)]_r)}{\sum_{r=1}^R \mathscr{L}_{i_c^{l_m}}(\hat{\Lambda}, [\xi_c^m(k), \zeta_c(k)]_r)} \quad (15)$$

in which $[\xi_c^m(k), \zeta_c(k)]_r$ are draws from a distribution $F(\xi_c^m(k), \zeta_c(k); \widehat{\Omega})$ with similar properties to those described in §5.4. Again, we use $R = 500$ simulation draws. Recalling that $k \in K = \{E, H, P\}$, we can now define a consistent estimator of $\rho^{(k)}_{-i_c^{l_m}}$ as

$$\hat{\rho}^{(k)}_{-i_c^{l_m}} = \left(\sum_{v \neq c} N_v^{l_m} \right)^{-1} \sum_{j \neq i_c^{l_m}} \hat{P}_j\left(a_{i_c^{l_m}} = k\right). \quad (16)$$

5.6. Common Unobservables

While our data set is rich enough to include a large number of covariates upon which firms may condition their actions, the strong emphasis we have placed on strategic interaction may raise concerns regarding the role of unobserved heterogeneity. In particular, how can we be sure that firms are actually reacting to the actions of their rivals, rather than simply optimizing over some common features of the local market that we do not observe? Manski (1993) frames this as the problem of distinguishing between endogenous and correlated effects. Although the presence of both effects yields collinearity in the linear in means model that Manski analyzes (i.e. the reflection problem), the nonlinearity of the discrete choice framework eliminates this stark nonidentification result in our setting. However, the presence of correlated unobservables remains a concern. In what follows, we outline two strategies for handling this problem. The first incorporates a fixed effect at the MSA level, while the second

[17] The subscript 1 indicates that these are first stage estimates.

incorporates a random effect at the level of the cluster. Our main results are robust to either alternative.

The most direct solution is to add a common unobservable, denoted η_{l_m}, to the strategy specific profit function of each store. Using the notation defined earlier, this can be written

$$\tilde{\pi}_{i_c^{l_m}}\left(a_{i_c^{l_m}}=k, s^m, \epsilon_i, \mathbf{P}_{l_m}\right) = s^{m\prime}\beta_k + \rho^E_{-i_c^{l_m}}\alpha_{k1} + \rho^P_{-i_c^{l_m}}\alpha_{k2} + \eta_{l_m} + \epsilon_{i_c^{l_m}}(k). \quad (17)$$

Ideally, one would estimate each η_{l_m} as a cluster specific fixed effect. However, this would require estimating 8,000 additional parameters with less than 18,000 observations, which is clearly infeasible. A feasible alternative is to model the common unobservable at the level of the MSA (i.e. include η_m instead of η_{l_m}). In practice, this simply requires running the first stage separately for each MSA and then adding an MSA level fixed effect to the second stage procedure. This has the added benefit of relaxing the equilibrium restriction: we need now only assume that a unique equilibrium is played in every MSA, instead of across all MSAs. We implement this strategy below. However, given the local nature of the strategic interaction documented here, an MSA level common unobservable may not be sufficient to account for the relevant correlated effects.

A second alternative is to use a cluster level *random effect* (i.e. assume the unobservables come from a pre-specified density $g(\eta_{l_m})$) and simply integrate out over η_{l_m} in the second stage estimation procedure, maximizing the resulting marginalized sample likelihood. However, there is an additional impediment to implementing this strategy: the fact that η_{l_m} is a *common* unobservable prevents the econometrician from obtaining a consistent first stage estimate of $\mathbf{P}_{l_m}$, a requirement of the two-stage procedure employed above. (Note that this is *not* a problem if the first stage can be estimated separately for each market, as was the case with the MSA level unobservable.) To accommodate cluster level random effects, we adopt an approach based on Aguirregabiria and Mira (2007) that is tailored to our particular setting (the details of our algorithm are provided in Appendix C).

5.7. Identification

Bajari et al. (2005) establish identification of the structural parameters of a broad class of discrete games of incomplete information, of which ours is a subcase. Their identification argument rests on three assumptions. The first two have already been (implicitly) stated, but will be repeated here more formally. The first assumption is that the error terms ϵ are distributed i.i.d. across players and actions in any given local market (i.e., cluster)[18] and are drawn from a distribution of known parametric form. This is clearly satisfied by the assumptions imposed above. The second assumption normalizes the expected profit associated with one strategy to zero. This is a standard identification condition of any multinomial choice model. We normalize the mean profit of the PROMO strategy to zero. The final assumption is an exclusion restriction.

The need for an exclusion restriction can be illustrated using Equation (9). Our two-step approach involves estimating the shares ($\rho_{-i_c^{l_m}}$'s) on the right hand side of (9) in a first stage. These shares, which are simple functions of each firm's beliefs regarding the conditional choice probabilities of its rival's, depend on the same state vector (s^m) as the first term of the profit function ($s^{m\prime}\beta_k$), creating a potential collinearity problem. Of course, identification can be trivially preserved by the inherent nonlinearity of the discrete choice problem, but this follows directly from functional form. An alternative strategy (suggested by Bajari et al. 2005) involves identifying one or more continuous covariates that enter firm i's payoffs, but not the payoffs of any of its rivals. Note that each firm's private shock ($\epsilon_{i_c^{l_m}}$) has already been assumed to satisfy this restriction, creating at least one set of "natural" exclusion restrictions. The characteristics of rival firms constitute an additional exclusion. However, a more subtle identification issue concerns the source of exogenous variation in the data that can pin down the form of strategic interaction. For this, we exploit the specific structure of the private information term and the presence of large multi-market chains. The two random effect terms in (7) capture each firm's tendency to employ a consistent strategy within an MSA ($\xi_c^m(k)$) and/or across all stores ($\zeta_c(k)$) in the chain. These firm level tendencies vary across chains and markets, providing a source of variation for the local interactions that take place in any given cluster. The key assumption is that we sometimes see firms that follow a consistent strategy (EDLP, for example) at the market level deviate in a local cluster by playing either PROMO or HYBRID when the demographics of the local market or its beliefs regarding rival strategies outweigh its desire to follow a consistent (chain or MSA-wide) strategy. This has the flavor of an instrumental variable approach, where the instruments are measures of the overall strategy a chain adopts outside the local market or MSA. In order to maintain the static, local, simultaneous move structure of the game, we have restricted these

[18] Note that the i.i.d. requirement need only hold at the cluster level. In particular, it's fine to include random effects in the error term, so long as they are treated as private information. This is the approach we adopt in our main specification.

firm level tendencies to be privately observed random effects. However, an alternative specification in which we conditioned directly on the average strategies that firms follow outside a given MSA yielded similar results.

6. Results and Discussion

As noted earlier, choosing an optimal pricing strategy is a complex task, forcing firms to balance the preferences of their customers against the strategic actions of their rivals. A major advantage of our two-step estimation approach is that, by estimating best response probability functions rather than equilibrium correspondences, we can separately identify strategic interactions, reactions to local and market level demographics, and operational advantages associated with larger stores and proprietary distribution systems. The Bayesian structure of the game allows us to account for different equilibria with the same covariates, due to the presence of unobserved types. More importantly, it allows us to model all 8,000 markets as variations in the play of a game with the same structure, but different conditioning variables. As the conditioning variables vary, we are able to trace out the equilibrium correspondence and identify the impact of several distinct factors. First, we find that firms choose strategies that are tailored to the demographics of the market they serve. Moreover, the impact of demographics corresponds closely to existing empirical studies of consumer preferences and conventional wisdom regarding search behavior. Second, we find that EDLP is favored by firms that operate larger stores and are vertically integrated into distribution. Again, this accords with conventional wisdom regarding the main operational advantages of EDLP. Finally, with regard to strategic interaction, we find that firms coordinate their actions, choosing pricing strategies that match their rivals. This identifies an aspect of firm behavior that has not been addressed in the existing literature: exactly how firms react to rival strategies.

Our main empirical results are presented in Table 6. The coefficients, which represent the parameters of the profit function represented in Equation (6), have the same interpretation as those of a standard MNL model: positive values indicate a positive impact on profitability, increasing the probability that the strategy is selected relative to the outside option (in this case, PROMO).

6.1. The Role of Demographics

The coefficients on consumer demographics are presented in the second and third sections of Table 6. With the exception of two MSA-level covariates, every demographic factor plays a significant role in the

Table 6 **Estimation Results**

	EDLP			HYBRID		
	Estimate	Std. err	T-stat	Estimate	Std. err	T-stat
Effect						
Intercept	−1.5483	0.2426	−6.3821	2.1344	0.2192	9.7372
Strategy variables						
$\hat{\rho}^{\mathrm{EDLP}}_{-i_c^l m}$	4.4279	0.1646	26.9010	−2.0924	0.1595	−13.1185
$\hat{\rho}^{\mathrm{PROMO}}_{-i_c^l m}$	−3.7733	0.1501	−25.1386	−6.3518	0.1351	−47.0155
MSA characteristics						
Size ('000 sq. miles)	0.0394	0.0848	0.4645	−0.0566	0.0804	−0.7039
Density (pop 10,000 per sq. mile)	−0.0001	0.0002	−0.4587	0.0006	0.0002	2.9552
Avg. food expenditure ($ '000)	−0.0375	0.0155	−2.4225	−0.0013	0.0141	−0.0904
Market variables						
Median household size	0.5566	0.1989	2.7983	0.2150	0.0900	2.3889
Median HH income	−0.0067	0.0019	−3.5385	0.0056	0.0017	3.2309
Proportion Black	0.6833	0.1528	4.4719	0.0139	0.1443	0.0963
Proportion Hispanic	0.5666	0.2184	2.5943	−0.0754	0.2033	−0.3708
Median vehicles in HH	−0.1610	0.0840	−1.9167	0.2263	0.1173	1.9292
Store characteristics						
Store size (sqft '000)	0.0109	0.0015	7.2485	0.0123	0.0014	8.8512
Vertically integrated	0.1528	0.0614	2.4898	0.0239	0.0550	0.4343
Chain characteristics						
Number of stores in chain	−0.0002	0.0001	−2.7692	0.0002	0.0001	3.5000
Chain effect	1.7278	0.0998	17.3176	2.8169	0.0820	34.3531
Chain/MSA effect	0.7992	0.0363	22.0408	0.9968	0.0278	35.8046

choice of EDLP as a pricing strategy. This is important from an econometric standpoint, since we use these very same factors to construct expectations in the first stage. In particular, the significance of the estimates means that we do not have to worry about collinearity. The statistical significance of the parameters is also substantively important. It suggests that the even after accounting for competitive and supply side (e.g. store/chain) characteristics, consumer demand plays a strong role in determining pricing strategy.

Focusing more closely on the demand related parameters, we find that (relative to PROMO), EDLP is the preferred strategy for geographic markets that have larger households ($\beta^{HH} = 0.5566$), more racial diversity in terms of African-American ($\beta^{BL} = 0.6833$) and Hispanic ($\beta^{HI} = 0.5666$) populations, lower income ($\beta^{INC} = -0.0067$), and fewer vehicles per household ($\beta^{VH} = -0.1610$). These results suggest that EDLP is mostly aimed at lower income consumers with larger families (i.e. more urbanized areas). Our findings are clearly consistent with the consumer segments that firms like Wal-Mart and Food Lion are widely perceived to target. It also accords quite well with the "fixed basket" model of shopping behavior (Bliss 1988, Bell and Lattin 1998), in which consumers who are more sensitive to the price of an overall basket of goods prefer EDLP. In particular, our results suggest that the consumers who are unable to substitute inter-temporally are disproportionately poor, nonwhite, and from larger families. On the other hand, we find that consumers who are most able to defer or stockpile purchases (wealthy suburbanites with greater access to transportation) tend to prefer PROMO or HYBRID pricing.

6.2. Firm and Store Level Characteristics

Turning next to chain and store level characteristics, we again find that most parameter estimates are statistically significant. These effects, which are in line with both theory and broad intuition, provide an additional empirical validation of our structural framework.

The last two sections of Table 6 reveal that stores choosing EDLP are both significantly larger ($\beta^{SS} = 0.0109$) and far more likely to be vertically integrated into distribution ($\beta^{VI} = 0.1528$). This is consistent with the view that EDLP requires substantial firm level investment, careful inventory management, and a deeper selection of products in each store in order to satisfy the demands of one-stop shoppers. It is also consistent with the logic of Lal and Rao (1997), whereby pricing strategy involves developing an overall positioning strategy, requiring complementary investments in store quality and product selection. Surprisingly, the total number of stores in the chain is negatively related to EDLP ($\beta^{ST} = -0.0002$), although this is difficult to interpret since almost all the large chains are vertically integrated into distribution (i.e. there are almost no large, nonvertically integrated firms). Finally, both the chain specific and chain/MSA random effects are highly significant, which is not surprising given the geographic patterns shown earlier.[19]

6.3. The Role of Competition: Differentiation or Coordination

By constructing a formal model of strategic interaction, we are able to address the central question posed in this paper—what is impact of competitive expectations on the choice of pricing strategy? Our conclusions are quite surprising. The first section of Table 6 reveals that firms facing competition from a high (expected) share of EDLP stores are far more likely to choose EDLP than either HYBRID or PROMO ($\hat{\alpha}_{k1} = 4.4279, \hat{\alpha}_{k2} = -3.7733$). The HYBRID case behaves analogously; when facing a high proportion of either EDLP or PROMO rivals, a store is least likely to choose HYBRID ($\hat{\alpha}_{k1} = -2.0924, \hat{\alpha}_{k2} = -6.3518$). In other words, we find *no evidence that firms differentiate themselves* with regard to pricing strategy. To the contrary, we find that rather than isolating themselves in strategy space, firms prefer to coordinate on a single pricing policy. Pricing strategies are strategic complements.

This coordination result stands in sharp contrast to most formal models of pricing behavior, which (at least implicitly) assume that these strategies are vehicles for differentiation. Pricing strategy is typically framed as a method for segmenting a heterogeneous market—firms soften price competition by moving further away from their rivals in strategy space. This is not the case for supermarkets. Instead of finding the anti-correlation predicted by these spatial models, we find evidence of associative matching, which usually occurs in settings with network effects or complementarities. This suggests that firms are able to increase the overall level of demand by matching their rivals' strategies, a possibility we discuss in more detail below. However, before discussing our coordination result in greater detail, we must address the issue of correlated unobservables.

[19] An earlier version of this paper also included the share of each firm's stores outside the local MSA that employ EDLP and PROMO pricing as additional regressors. Not surprisingly, a firm's propensity to follow a particular strategy at the level of the chain had a large and significant impact on its strategy in a particular store (and soaked up a lot of variance). While this suggests the presence of significant scale economies in implementing pricing strategies, as suggested by both Lattin and Ortmeyer (1991) and Hoch et al. (1994), we omitted it from the current draft to maintain the internal coherency of the underlying model (i.e. the simultaneity of actions). However, these results are available from the authors upon request.

Table 7 Robustness

Specification	Strategy	Strategy $\hat{\rho}^{\text{EDLP}}_{-i_c^{lm}}$	Variables $\hat{\rho}^{\text{PROMO}}_{-i_c^{lm}}$
Baseline	EDLP	4.4279 (0.1646)	−3.7733 (0.1501)
	HYBRID	−2.0924 (0.1595)	−6.3518 (0.1351)
MSA by MSA	EDLP	3.1867 (0.2522)	−3.2823 (0.1771)
	HYBRID	−3.4418 (0.2603)	−6.2746 (0.1701)
NPL	EDLP	1.7464 (0.1743)	−2.5699 (0.1723)
	HYBRID	−0.7365 (0.1770)	−4.9899 (0.1739)

The surprising nature of our coordination result demands careful consideration. Again, how can we be sure that firms are actually reacting to the actions of their rivals, rather than simply optimizing over some common but unobserved feature of the local market? Section 5.6 described two alternative strategies for dealing with the potential presence of common unobservables. The first method involved adding an MSA level fixed effect to the baseline specification. In practice, this requires estimating the first stage separately for each MSA (to ensure a consistent first stage) and then expanding the second stage likelihood to include an MSA fixed effect. The main coordination results are presented in the section of Table 7 titled MSA by MSA (the demographic and chain level covariates have been suppressed for brevity, but are available from the authors upon request). While the coefficients have changed slightly in magnitude, the main coordination result remains strong. The second method involved adding a cluster level random effect, and re-estimating the model using Aguirregabiria and Mira's (2006) NPL algorithm. These results are presented in the section titled NPL. Here we find that the magnitudes of the coefficients fall relative to both the baseline and MSA by MSA specifications, as one might expect if firms are indeed reacting to a common unobservable. However, the coordination effects are still large and significant: pricing strategies are indeed strategic complements.

But how important are these strategic effects? The parameter estimates from our baseline model can be used to gauge the relative influence that strategic interactions have on profits. Because individual covariates can influence profits either negatively or positively, a simple additive decomposition of profits by strategic and nonstrategic factors is inappropriate. To adjust for this, we adopted the method proposed in Silber et al. (1995), using the average squared contributions of each factor to construct a measure of the share of variance explained. This decomposition reveals that, on average, strategic factors explain about 20.3% of the variation in EDLP profits and 13.2% of the variation in HYBRID profits, quite substantial fractions. The remaining variance is explained by nonstrategic factors, including market characteristics, store and firm-level covariates, and the random effects that we have treated as private information.

In addition to this decomposition of profits, we also conducted a policy experiment aimed at highlighting the mechanism by which strategic effects influence pricing strategy. To do so, we simply shut off the strategic effects and compared the odds of choosing PROMO relative to EDLP under this counterfactual scenario to what we see in the data.[20] The results are striking. At the aggregate level, the true odds ratio was around 1.31, implying that PROMO is roughly 31% more likely to be chosen than EDLP. However, in the counterfactual scenario (without strategic effects) this drops to 4.1% (odds ratio = 1.041). This finding is notable since it offers an explanation for why EDLP did not become the dominant paradigm in supermarket pricing. To see why, notice first that even without strategic effects, the odds ratio was greater than one. This implies that there are market factors that, on average, lead a market to lean towards one strategy. With the same set of characteristics, the strategic effects induce a feedback effect that can cause the market to tip more significantly in that direction. While these are clearly aggregate trends, we observed similar phenomena in individual markets as well. Broadly speaking, strategic effects strengthen coordination in markets where one strategy is weakly dominant (under the counterfactual).

6.4. Discussion of Results

The Bayesian structure of our game allows us to represent a quite complex game using a relatively simple structure. By tracing out the equilibrium correspondence, we have found that firms favor particular strategies in certain markets, in ways that are consistent with existing theory. We have also found that certain types of firms favor particular strategies, also consistent with existing theory. Finally, we have found that firms are more likely to choose a particular strategy if they expect their rivals to do the same. This is a sharp departure from existing theory. It is worth emphasizing that reactions to market demographics and firm characteristics help explain *how* firms are able to coordinate on consistent strategies. However, they do not explain *why* they choose to do so. Coordination implies that firm's conditional choice probabilities act as strategic complements, meaning that their best response probability functions (9) are upward sloping. To support such complementarity, coordination must somehow increase the overall size of the perceived market. In most cases, this means drawing expenditures away from the outside good.

[20] Note that $(\hat{P}_{\text{PROMO}}/\hat{P}_{\text{EDLP}})$ is now our object of interest.

In the context of supermarket pricing, this suggests that coordination may actually increase the amount consumers are willing to spend on groceries, perhaps by drawing them away from substitutes like restaurants, convenience stores, and discount clubs. One way this might occur in practice is if consumers are more likely to trust retailers that provide a message that is consistent with those of their rivals. In other words, if one firm tells you that providing the highest value involves high price variation while another touts stable prices, you may be unwilling to trust either, shifting your business to a discount club or another retail substitute. While this intuition has yet to be formalized, it is consistent with the emphasis that Ortmeyer et al. (1991) place on maintaining pricing credibility. Another possibility, consistent with Lal and Rao (1997), is that price positioning is multi-dimensional and by coordinating their strategies stores can mitigate the costs of competing along several dimensions at once. Without a formal model of consumer behavior and detailed purchase data, we are unable to pin down the exact source of the complementarities we have documented here. However, we have provided strong empirical evidence regarding how firms actually behave. Understanding why firms find it profitable to coordinate their actions remains a promising area for future theoretical research.

The results presented above provide definitive answers to the three questions posed in the introduction of this paper. We have found that demand related factors (i.e. demographics) are important for determining the choice of pricing strategy in a market; store and firm level characteristics also play a central role. Both of these results are in line with the extant literature. However, our results concerning competitive expectations are in sharp contrast to prevailing theory in both economics and marketing and warrant further attention. The final section outlines a research agenda for extending the results found in this paper.

7. Conclusions and Directions for Future Research

This paper analyzes supermarket pricing strategies as discrete game. Using a system of simultaneous discrete choice models, we estimate a firm's optimal choice conditional on the underlying features of the market, as well as each firm's beliefs regarding its competitor's actions. We find evidence that firms cluster by strategy, rather than isolating themselves in product space. We also find that demographics and firm characteristics are strong determinants of pricing strategy. From a theoretical perspective, it is clear that we have yet to fully understand what drives consumer demand. The fact that firms coordinate with their rivals suggests that consumers prefer to receive a consistent message. While our results pertain most directly to supermarkets, it seems likely that other industries could behave similarly. Future research could examine the robustness of our findings by analyzing other retail industries, such as department stores or consumer electronics outlets.

In this paper, our primary focus was the construction and econometric implementation of a framework for analyzing best responses to rival pricing strategies. Our analysis describes the nature of strategic interactions, but does not delve into the details of why these strategies are dominant. Decomposing the *why* element of strategic coordination seems a fruitful area for research. We hasten to add that such research is needed not only on the empirical side but also on the theoretical front. Building theoretical models that allow for the possibility of both differentiation and coordination is a challenging but undoubtedly rewarding path for future research.

The tendency to coordinate raises the possibility that games such as this might support multiple equilibria. While this is not a concern in our current study, it could play a central role when conducting policy experiments or when analyzing settings in which demographics (or other covariates) cannot effectively facilitate coordination. Developing methods that are robust to such possibilities remains an important area for future research.

On the methodological front, our research also centers its attention on the discrete choice aspect of strategy. There are a number of issues that emerge once such strategic choices have been made such as the reaction of consumers (see e.g. Singh et al. 2006) and the overall demand faced by stores. Research that aims at incorporating such postgame outcome data into the analysis promises to offer newer and crisper insights into the nature of competition in the market.

Finally, in building our model of strategic interaction, we have assumed that firms interact in a static setting, making independent decisions in each store. A more involved model would allow chains to make joint decisions across all of their outlets and account for richer (dynamic) aspects of investment. Developing such a model is the focus of our current research.

Acknowledgments

The authors thank participants at the Supermarket Retailing Conference at the University of Buffalo, the 2006 BCRST Conference at the University of Toronto, the 2005 QME Conference at the University of Chicago, and the Supermarket Conference held at IFS London, as well as seminar participants at Duke, UCLA, Stanford, UT Dallas, and Yale. The authors also thank Victor Aguirregabiria, Pat Bajari, J. P. Dubé, Han Hong, Paul Nelson, and

Chris Timmins for their comments. All remaining errors are the authors' responsibility.

Appendix A. Survey Validity

All of the variables in the Trade Dimensions data, including the information on pricing strategy, are self-reported. This may raise some concerns regarding accuracy, especially given the high degree of local variation we observe in the data. Two questions naturally arise. First, are firms actually willing and able to set prices at such local levels? Second, do these self-reported strategies reflect actual differences in pricing behavior? We will address both issues in turn.

First, with regard to local pricing, we should note that supermarket firms clearly have the technological resources to set prices (and therefore pricing strategy) at a very local level. Indeed, Montgomery (1997) provides a novel method for profitably customizing prices at the store level, using widely available scanner data.[21] We contacted pricing managers at several major chains and other industry professionals regarding their ability to engage in such micro-marketing. Even on the condition of anonymity, they were extremely reluctant to discuss the details of their actual pricing strategies, but did acknowledge that they "certainly have the data and resources to do it." Furthermore, a consultant who was involved in several recent supermarket mergers confirmed that the extent of local pricing was a key factor in the approval process.[22]

A related issue is whether firms face significant pressure to maintain a consistent (pricing) image across stores. We suspect not. Unlike many other types of retail food services (e.g., fast food establishments), supermarket customers do the majority of their shopping in a single store.[23] Therefore, while consumers undoubtedly have strong preferences over the pricing strategy of their *chosen store*, they have little reason to care directly about the overall strategy of the *chain*. Of course, chains may have strong *operational* incentives (e.g. scale economies in distribution and advertising) to maintain a consistent strategy across several (not necessarily proximate) stores, which might lead them to adopt a common strategy in multiple outlets. Indeed, we are relying on just such incentives to provide the variation needed to identify the effect of strategic interactions (see §5.7). The point is that firms may indeed have both strong incentives and the ability to tailor pricing to the local environment.

The second question concerns the validity of the survey instrument itself. We note first that the survey was of store managers but administered by brokers (who explained the questions), providing an additional level of cross-validation. It is unlikely that the results reported below could be the product of systematic reporting error, as this would require coordination between tens of thousands of managers and hundreds of brokers to willfully and consistently mis-report their practices (for no obvious personal gain). However, to further allay such fears, we cross-verified the data ourselves using publicly available data from the Dominick's Finer Foods (DFF) supermarket chain in Chicago. In particular, we extracted store level prices from four major product categories for every store in the DFF data set and matched them up to the pricing classifications reported by Trade Dimensions. The vast majority of the Dominick's stores are identified as PROMO (93%), while the remainder are HYBRID, which is itself encouraging since Dominick's is known to be a PROMO chain. We then checked whether the incidence of promotions (i.e. whether a UPC was "on sale") varied across PROMO and HYBRID stores. In all four categories that we examined (Soft Drinks, Oatmeal, Paper Towels, and Frozen Juice), we found a significantly lower incidence of promotions at the HYBRID stores. The differences ranged from 8.1% in Soft Drinks (a very heavily promoted category) to 23.4% in Oatmeal. All differences were significant at the 1% level.

In addition, we also compared the HYBRID and PROMO stores for equality in the variance of the prices using standard folded—F tests. One would expect PROMO stores to have higher variances. For three of the four categories (Oatmeal, Paper Towels, and Frozen Juice) the variance in prices was indeed higher in the PROMO stores, validating the survey data. The difference was not significant for Soft Drinks category. We also repeated each analysis for only the highest selling UPC in each category and found qualitatively similar results. While these tests use only a few product categories from a single chain in a single market, the sharpness of the results should provide additional confidence in the integrity of our data.

Appendix B. Robustness Checks

B.1. Market Delineation and Definition

As noted earlier, our empirical analysis uses specific market definitions based on spatial cluster analysis. We verified the robustness of our results to alternative market definitions by repeating the analysis using ZipCodes, Counties, and MSAs. In all cases, the results were qualitatively similar. We also varied the number of clusters and did not find significant differences from the results reported above. Finally, we experimented with n-nearest neighbor methods (we tried 3 and 5 nearest neighbors of a focal store) and again found similar results.

B.2. Multiplicity

As we noted in the main text, consistent estimation of a static (or dynamic) game requires some form of uniqueness of equilibrium, either in the model or in the data.[24] Consistency of our baseline model requires that only one equilibrium be played in the data which, in our context, means

[21] While the emphasis there is on maintaining a consistent image, Montgomery argues that the potential gains to micro-marketing are quite significant. Setting different sales frequencies in different stores is simply an alternative method of micro-marketing.

[22] While detailed information on the degree of micro-marketing in the supermarket industry is not publicly available, explicit evidence of local pricing was a major issue in the proposed merger between Staples and Office Depot (Ashenfelter et al. 2006).

[23] According to the Food Marketing Institute, consumers allocate 78% of their overall budget to their primary store. Moreover, their secondary store is almost always part of a different chain.

[24] Uniqueness may fail to hold in many settings. Brock and Durlauf (2001) and Sweeting (2004) provide two such examples. Non-uniqueness can complicate policy experiments, which typically involve solving for a new equilibrium.

every location in every MSA. It is possible to relax this by estimating the first stage separately for each MSA, so the requirement becomes that a unique equilibrium be played in each MSA (we do not have enough data to estimate the first stage separately for each cluster, which would eliminate the problem entirely). The results of this procedure were very close to the baseline model. For brevity, we report only the coefficients on the strategy variables (see Table 7).

B.3. Format Characterization

In our baseline model, we assumed that firms care only about the share of their rivals that choose each strategy. An alternative, similar to what is done in the entry literature, is to assume that firms care instead about the number of rivals. We reestimated the baseline model using counts instead of shares and found qualitatively similar results.

B.4. Nonparametric Estimation of ρ_{-i}

As noted above, the ideal approach for estimating beliefs is nonparametric. However, the number of covariates we use precludes us from adopting such a strategy. To assess the robustness of our results, we used a bivariate thin-plate spline to model pricing strategies as nonparametric functions of the strategies chosen outside the MSA. Again, the main results were qualitatively similar to those presented above.

B.5. Nonlinearity of $f(\rho_{-i})$

To examine the potentially nonlinear relationship between payoffs (Π) and strategies (ρ_{-i}), we adopted a smoothing splines approach to modeling $f(\rho_{-il})$. In particular, we reestimated our model using a bivariate thin-plate spline, treating the functional relationship as

$$f_j(a_{-i_c^{l_m}}) = f(\rho^E_{-i_c^{l_m}}, \rho^P_{-i_c^{l_m}} \mid \varpi). \tag{18}$$

The qualitative results obtained using the linear specification continue to hold. For example, the probability of firms choosing EDLP increases with the proportion of competitors that also choose EDLP.

B.6. Error Structure

In our analysis we assumed that firm types (the ϵ_i's) were distributed Gumbel (Type I Extreme Value), allowing us to specify set of simultaneous multinomial logit choice probabilities for determining pricing policies. As an alternative specification, similar to the empirical application in Bajari et al. (2005), we also tested ordered logit/probit models in which the strategies were ranked ordered EDLP to PROMO. While qualitative findings were similar, these ordered specifications force a particular ranking of strategies that may not be warranted.

Appendix C. Nested Pseudo Likelihood (NPL) Algorithm

We assume that the common unobservables are jointly distributed with distribution function $F(\boldsymbol{\eta}_{l_m} \mid \Omega)$, where Ω is a set of parameters associated with F. To start the algorithm, let $\mathbf{P}_{l_m}$ be the set of strategy choice probabilities across players in a given local market l_m. Finally, let $\widehat{\mathbf{P}}^0_{l_m}$ be some (not necessarily consistent) estimator for $\mathbf{P}_{l_m}$.

In the rth iteration implement the following steps:

Step 1. Given $\widehat{\mathbf{P}}^{r-1}$ update $\boldsymbol{\Theta}$

$$\widehat{\boldsymbol{\Theta}}^r = \arg\max_{\boldsymbol{\Theta}} \sum_{m\in M}\sum_{c\in C} \ln \int \Bigg[\prod_{l_m\in L_m} \prod_{i_c^{l_m}\in N_c^{l_m}} \prod_{k\in K} \big[\Psi_{i_c^{l_m}}\big(a_{i_c^{l_m}} = k \mid \boldsymbol{\Theta}, \mathbf{P}_{l_m}, \mathbf{s}, \boldsymbol{\eta}_{l_m}(k)\big)\big]^{\delta_{i_c^{l_m}}(k)} \Bigg] dF(\boldsymbol{\eta}_{l_m} \mid \Omega)$$

Step 2. Update $\mathbf{P}_{l_m}$ using $\widehat{\Theta}^r$, setting

$$\widehat{P}^r_{i_c^{l_m}}(a_{i_c^{l_m}} = k) = \int \Psi_{i_c^{l_m}}\big(a_{i_c^{l_m}} = k \mid \widehat{\Theta}^r, \widehat{\mathbf{P}}^{r-1}_{l_m}, \mathbf{X}, \eta_{l_m}(k)\big) \cdot \Lambda\big(\eta \mid \widehat{\Theta}^r, \widehat{\mathbf{P}}^{r-1}_{l_m}, \mathbf{s}, \boldsymbol{\eta}_{l_m}(k), a_{i_c^{l_m}}\big)\, dF\big(\boldsymbol{\eta}_{l_m} \mid \widehat{\boldsymbol{\Omega}}^r\big),$$

where

$$\Lambda(\eta \mid \cdots) = \Bigg(\prod_{l_m\in L_m} \prod_{i_c^{l_m}\in N_c^{l_m}} \prod_{k\in K} \big[\Psi_{i_c^{l_m}}\big(a_{i_c^{l_m}} = k \mid \widehat{\Theta}^r, \widehat{\mathbf{P}}^{r-1}_{l_m}, \mathbf{s}, \eta_{l_m}(k)\big)\big]^{\delta_{i_c^{l_m}}(k)} \Bigg) \cdot \Bigg(\int \prod_{l_m\in L_m} \prod_{i_c^{l_m}\in N_c^{l_m}} \prod_{k\in K} \big[\Psi_{i_c^{l_m}}\big(a_{i_c^{l_m}} = k \mid \widehat{\Theta}^r, \widehat{\mathbf{P}}^{r-1}_{l_m}, \mathbf{s}, \boldsymbol{\eta}_{l_m}(k)\big)\big]^{\delta_{i_c^{l_m}}(k)} dF\big(\boldsymbol{\eta}_{l_m} \mid \widehat{\boldsymbol{\Omega}}^r\big) \Bigg)^{-1}$$

Step 3. If $\|\mathbf{P}^r_{l_m} - \mathbf{P}^{r-1}_{l_m}\|$ is smaller than a predetermined value, stop and choose $\widehat{\Theta}^{NPL} = \Theta^r$. If not, increment r and return to Step 1.

Note that there are a few key differences between our approach and the Nested Pseudo Likelihood (NPL) algorithm proposed by Aguirregabiria and Mira (2007) (henceforth AM). First, unlike AM, our game is static. This does not alter the main econometric properties of the NPL estimator, since a static game is simply a one-period subcase of a dynamic one. However, a natural consequence of the static setting is that the state variables do not transition over time, allowing us to extend the NPL approach to include continuous states. A more significant point of departure between our algorithm and the NPL is the inclusion of continuous heterogeneity. Since the evolution of the observed state variables naturally depends on the unobserved state variables, AM restricted their estimator to a finite support. In our case, the static nature of the problem, coupled with an independence assumption (η_{l_m} is orthogonal to s), allows us to simply integrate out over a continuous heterogeneity distribution. An attractive feature of the NPL algorithm is that it works even in the presence of inconsistent or poorly estimated initial probabilities. As long as the algorithm converges, it will do so to the root of the likelihood equations. In our experience, the procedure converged very quickly to the same fixed point for several different starting values.

References

Aguirregabiria, V., P. Mira. 2007. Sequential estimation of dynamic discrete games. *Econometrica* **75**(1) 1–53.

Ashenfelter, O., D. Ashmore, J. B. Baker, S. Gleason, D. S. Hosken. 2006. Econometric methods in merger analysis: Econometric analysis of pricing in FTC v. staples. *Internat. J. Econom. Bus.* **13**(2) 265–279.

Augereau, A., S. Greenstein, M. Rysman. 2006. Coordination versus differentiation in a standards war: 56K modems. *RAND J. Econom.* **37**(4) 887–909.

Bajari, P., C. L. Benkard, J. Levin. 2007. Estimating dynamic models of imperfect competition. *Econometrica* **75**(5) 1331–1370.

Bajari, P., H. Hong, J. Krainer, D. Nekipelov. 2005. Estimating static models of strategic interactions. Working paper, University of Michigan.

Bayer, P., C. Timmins. 2007. Estimating equilibrium models of sorting across locations. *Econom. J.* **117**(518) 353–374.

Bell, D., J. Lattin. 1998. Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Sci.* **17**(1) 66–88.

Bell, D., T. Ho, C. S. Tang. 1998. Determining where to shop: Fixed and variable costs of shopping. *J. Marketing Res.* **35**(3) 352–369.

Bliss, C. 1988. A theory of retail pricing. *J. Indust. Econom.* **36**(4) 375–391.

Brock, W., S. Durlauf. 2001. Discrete choice with social interactions. *Rev. Econom. Stud.* **62**(2) 235–260.

Dhar, S. K., S. J. Hoch. 1997. Why store brand penetration varies by retailer. *Marketing Sci.* **16**(3) 208–227.

Einav, L. 2003. Not all rivals look alike: Estimating an equilibrium model of the release date timing game. Working paper, Stanford University.

Ellickson, P. 2006. Quality competition in retailing: A structural analysis. *Internat. J. Indust. Organ.* **24**(3) 521–540.

Gowrisankaran, G., J. Krainer. 2004. The welfare consequences of ATM surcharges: Evidence from a structural entry model. Working paper, Washington University.

Harsanyi, J. 1973. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *Internat. J. Game Theory* **2**(1) 1–23.

Hoch, S., X. Dreze, M. Purk. 1994. EDLP, Hi-Lo, and margin arithmetic. *J. Marketing* **58**(4) 16–27.

Hotz, J., R. Miller. 1993. Conditional choice probabilities and the estimation of dynamic models. *Rev. Econom. Stud.* **60**(3) 497–531.

Kumar, N., R. Rao. 2006. Using basket composition data for intelligent supermarket pricing. *Marketing Sci.* **25**(2) 188–199.

Lal, R., R. Rao. 1997. Supermarket competition: The case of everyday low pricing. *Marketing Sci.* **16**(1) 60–81.

Lattin, J., G. Ortmeyer. 1991. A theoretical rationale for everyday low pricing by grocery retailers. Working paper.

Manski, C. 1993. Identification of endogenous social effects: The reflection problem. *Rev. Econom. Stud.* **60**(3) 531–42.

Messinger, P., C. Narasimhan. 1997. A model of retail formats based on consumers' economizing on shopping time. *Marketing Sci.* **16**(1) 1–23.

Montgomery, A. L. 1997. Creating micro-marketing pricing strategies using supermarket scanner data. *Marketing Sci.* **16**(4) 315–337.

Newey, W. K., D. McFadden. 1999. Large sample estimation and hypothesis testing. D. McFadden, R. Engle, eds. *Handbook of Econometrics*, Vol. 4, Chap. 36. Elsevier, North-Holland, Amsterdam, The Netherlands, 2113–2245.

Orhun, A. Y. 2005. Spatial differentiation in the supermarket industry. Working paper, University of California.

Ortmeyer, G., J. Quelch, W. Salmon. 1991. Restoring credibility to retail pricing. *Sloan Management Rev.* **33**(1) 55–66.

Rust, J. 1987. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica* **55**(5) 999–1033.

Seim, K. 2006. An empirical model of firm entry with endogenous product-type choices. *RAND J. Econom.* Forthcoming.

Shankar, V., R. Bolton. 2004. An empirical analysis of determinants of retailer pricing strategy. *Marketing Sci.* **23**(1) 28–49.

Silber J., P. Rosenbaum, R. Ross. 1995. Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics. *J. Amer. Statist. Assoc.* **90**(429) 7–18.

Singh V., K. Hansen, R. Blattberg. 2006. Market entry and consumer behavior: An investigation of a wal-mart supercenter. *Marketing Sci.* **25**(5) 457–476.

Smith, H. 2006. Supermarket choice and supermarket competition in market equilibrium. *Rev. Econom. Stud.* Forthcoming.

Sweeting, A. 2004. Coordination games, multiple equilibria, and the timing of radio commercials. Working paper, Northwestern University.

Varian, H. 1980. A model of sales. *Amer. Econom. Rev.* **70**(4) 651–659.

Watson, R. 2005. Product variety and competition in the retail market for eyeglasses. Working paper, University of Texas.

Zhu, T., V. Singh, M. Manuszak. 2005. Market structure and competition in the retail discount industry. Working paper, Carnegie Mellon University, Pittsburgh.