The Welfare Effects of Price Advertising with Basket Shopping: Structural Estimates from Supermarket Promotions

Gao, Cixiu

Publication date:
2015

The Welfare Effects of Price Advertising with Basket Shopping: Structural Estimates from Supermarket Promotions

Cixiu Gao*
Wuhan University

October, 2014

Abstract

Retailers send price cut information, or promotions, to attract consumers to the store where they would buy other items at high prices (one-stop shopping behavior). Such price advertising comes at enormous costs, and the economic efficiency of such markets is empirically unclear. This paper quantitatively shows the important role shopping transportation plays in the performance of a spatial supermarket oligopoly with price advertising. Equilibrium advertising intensities are found to be socially excessive, as price advertising erodes consumer surplus by triggering additional, perhaps long-distance shopping trips. Counterfactual simulations show how this social waste can be mitigated by (1) a promotion tax levied on retailers, and (2) on-line shopping.

Keywords: Price Advertising; Market Efficiency; Structural Estimation; Moment Inequality Approach; Counterfactual Simulation; Supermarket Retail Industry; Scanner Data.

JEL Classification Numbers: L1, M37.

*I am grateful to my dissertation committee members Rosanne Altshuler, Sanjib Bhuyan, Douglas Blair and Hilary Sigman for their comments and suggestions. I thank David Bell at Wharton School for providing the data. I also thank Colin Campbell, Ruby Henry, Mark Killingsworth, John Landon-Lane, Tom Prusa, Anne Piehl, and seminar participants at Rutgers. I obtained great support and suggestions from Tomas Sjöström. All errors are mine.

1 Introduction

This paper empirically examines the effects of costly information and transportation on market efficiency in the supermarket retail industry. One-stop shopping economizes on shopping trip, avoiding all-over-the-town transportation. Advertised price cuts, or promotions, are used by supermarkets as an instrument of competition, informing potential customers about price offers at a specific store location. Such price advertising comes at a cost. During 2007-2010, U.S. supermarkets spent about $800 million on price advertising per year.[1] An important aspect of industrial economics is to examine the market efficiency of price advertising.

Theories of informative price advertising do not provide unambiguous predictions of welfare effects. [2,3] Two effects of informative price advertising, the *demand-creating effect* and the *business-stealing effect*, have been recognized since Marshall (1919). In general, market efficiency with price advertising depends on the magnitudes of the two effects.[4] In the supermarket retail industry, there is a third, important, *transportation-erosion effect*. Transportation costs erode consumer surplus when price cuts trigger a shopping trip. Moreover, in a competition-intensified market, shoppers are more likely to shop at distant stores. The theoretical implications of this transportation cost have been recognized in the literature (e.g, Bester and Petrakis (1995)). However, the empirical significance of the interaction between informative price advertising and transportation erosion needs to be quantified.

I examine the performance of this spatial market by quantifying transportation erosion in a supermarket oligopoly, in addition to the demand-creating and the business-stealing effects. The key variables to be quantified are the cost of price advertising, shopping transportation costs, consumer surplus and the extent to which it is eroded by transportation. To do this, I estimate demand and the cost of advertising using a model that accounts for both consumer shopping behavior and retail pricing and advertising behavior – a Hotelling model with a few twists. In a Bayesian-Nash equilibrium, shoppers receive promotion information and choose

[1]Data scource: Kantar Media and http://online.wsj.com. See also Bolton et al. (2010) and Levy et al. (1997).

[2]This paper focuses on the effect of informative advertising which conveys price information only and distinguishes with the literature that examines welfare implication of 'persuasive' advertising, such as Dixit and Norman (1978), Stigler and Becker (1977) and Nichols (1985), where advertisements shift consumer preference.

[3]For example, Butters (1977) and Roy (2000) predict that equilibrium advertising is socially optimal; in models by Stegeman (1991) and Stahl and Dale (1994) private advertising is socially inadequate; Grossman and Shapiro (1984) argue that it could be either socially inadequate or excessive.

[4]On the one hand, because the firm cannot appropriate all of the social benefit created, the demand-creating effect suggests that equilibrium advertising is socially inadequate. On the other hand, the business-stealing externality among competing firms suggests that advertising may be socially excessive: the firm is motivated by the profit margin "stolen" from rivals, while social welfare is not impacted by the simple re-distribution of margins from one firm to another.

one optimal store to buy a bundle of products from; the competing retailers maximize the expected store-level profits by making promotion and pricing decisions for all products, facing the tradeoff between attracting extra store visits and paying additional promotion costs. The structural model is estimated using consumer scanner panel data in a two-year window from a mid-US metropolitan area, combined with the corresponding pricing and promotion information of the stores.

To know whether the current equilibrium advertising levels are socially excessive, I simulate market outcomes in the neighborhood of the current equilibrium: I numerically measure components of market surplus when there is one more, or less, unit of promotion (i.e., one more or less item promoted). If one less unit of promotion leads to higher aggregate surplus, it implies that the current equilibrium advertising is socially excessive.

My demand estimation follows fairly standard methods from the discrete choice literature. The more challenging estimation problems come from the firms' side. A well-known problem is that the firms' marginal costs (which include wholesale prices, distribution and shelving costs, etc) are unknown. This is solved using the firms' first-order condition at the observed pricing decisions following the classic IO literature (Bresnahan, 1987; Nevo, 2001; Porter, 1983). Another crucial step is to estimate the cost of price advertising (promotion). I tackle this problem by adopting the moment inequality approach (Pakes, 2010; Pakes et al., 2011).

The moment inequality approach is based on the necessary condition of profit maximization – the retailer chooses strategies that according to his expectations lead to profits at least as high as feasible alternatives. By estimating demand, I am able to predict how sales, and therefore profits, would have changed if the retailer had made slightly different decisions of advertising intensity. The moment inequality approach uses the difference between the actual and the counterfactual profits to obtain bounds for promotion costs. This approach partly mitigates the dimensionality issues caused by the large number of products (identified by SKU, stock keeping unit) of the multi-product retailers. Due to computational complexity, this step also involves some novel techniques which will be explained below.

My estimates and simulations suggest a few important points. (1) The size of transportation erosion is significant – in the current equilibrium, it erodes consumer surplus by about 36 percent. (2) The equilibrium advertising intensities are found to be socially excessive, because the social benefit created by the last unit of promotion, which is mainly sales expansion, does not outweigh the social costs, which to a large extent consists of shopping transportation costs. This finding naturally leads to experiments where transportation costs are somehow reduced, or completely removed. (3) In the extreme case where no transportation costs occur, the equilibrium promotion intensities are found to be socially inadequate, as social cost generated by the last unit of promotion is now small enough to be outweighed by the extra social benefit.

One possible way of improving market efficiency is to introduce an advertising tax levied on stores. The intuition is that facing expensive advertising, stores will promote less, which leads to effectively smaller likelihoods of distant shopping. In fact, as counterfactual simulation 1 finds, the equilibrium outcome is quite sensitive to the (exogenous) change in the cost of advertising: a 5% increase in promotion costs can improve market surplus by 5.75%. A more straightforward way of improving efficiency is to exogenously reduce transportation costs. In counterfactual scenario 2, where zero distances are traveled for shopping (e.g. through on-line shopping channels), the stores' local market power is removed, and a more rigorous competition is induced. With lower prices, more promotions and zero transportation erosion, market surplus can be improved by a considerabe amount, 115 percent.

The rest of the paper is organized as follows. Section 2 provides an overview of the supermarket retail industry. Section 3 presents the model of demand and supply. Section 4 describes the dataset. Section 5 explains the estimation procedure. Section 6 contains the welfare implications of the current equilibrium. Section 7 discusses the counterfactual simulations and Section 8 concludes.

2 The Supermarket Retail Industry

The grocery retailer is located at the end of the marketing chain, purchasing goods in bulk from manufacturers or wholesalers and directly servicing the final consumer. A grocery store is classified as a supermarket if its annual sales exceed $2 million; it emphasizes self-service and features dairy, meat, produce, dry grocery and home good departments. Through advertising and point-of-purchase material, retailers furnish information to customers about the prices of goods.

Grocery retailing is the largest retail sector in the U.S. economy and the most expensive segment of the grocery retailing system (Kohls and Uhl, 2001). The total supermarket sales exceed $602 billion in 2012 and consumer food expenditures account for 5.7% of disposable income in 2011. Competition among supermarket retailers is fierce – the net profit margin (after tax) in 2012 is only 1.5%.[5] In-store marketing activities (such as price promotion, display, food sampling, etc) and store branding largely account for a the gap between retail and wholesale prices.

Supermarket retailers use an arsenal of marketing instruments and sophisticated pricing

[5]Food Marketing Institute, http://www.fmi.org/research-resources/supermarket-facts.

strategies to attract customers and avoid being squeezed out of the market. They price each product as a component of a total mix of products offered by the store, often referred to as "basket pricing". Another strategy is temporary advertised price cuts, or variable price merchandising, used to differentiate their stores and attract consumers. Practically, the store manager chooses the set of promoted products and price cuts, and changes the set on a weekly basis. This strategy relies on the consumers' tendency towards one-stop shopping; thus low profits or losses on the featured items can be made up by purchases of the higher-profit items. For shoppers, prices are not the only determinant of store choice: factors such as geographical location of stores, product assortment, shopping experience and customer service, are also important.

The traditional brick-and-mortar supermarkets are being encroached by online grocery shopping. In contrast to the conventional wisdom that Internet grocery shopping only fills a small niche for high-income consumers who place a high value on their time and a low value on store experience, recent trends show that this new shopping channel is pervasive.[6] Moreover, on-line shopping exhibits great economies of scale in logistics. It is therefore valuable to make economic predictions for supermarket industry with this novel shopping channel.

3 Model

A few facts about the supermarket industry deviate from the common assumptions made in standard theories. These assumptions include single-product oligopoly firms and single-product shopping behavior,[7] consumers' unawareness of product availability unless informed by price advertising[8], and firms' optimization by choosing the market "reach" of advertising. In contrast, the supermarket industry is characterized by consumer basket shopping behavior and multi-product firms; product availabilities are usually well known to shoppers; and retailers make advertising decisions by selecting the set of promoted products.

To investigate the pricing strategy and market efficiency in the supermarket retail industry, I set out a model of consumer and firm behavior. The model assumes that in a Bayesian equilibrium shoppers choose the store that offers the greatest shopping utility given store characteristics and their price knowledge; stores maximize store-level profits by making pricing and promotion decisions. They can inform shoppers about price promotions (which items are promoted and how much they are priced) in order to compete over sales. The estimated shopping

[6] *Four Forces Shaping Competition in Grocery Retailing*, industry report, Booz & Company.

[7] One exception is the Hotelling model of two-product duopoly constructed by Lal and Matutes (1994).

[8] See, for example, Bagwell (2007), Butters (1977), Grossman and Shapiro (1984), and Stegeman (1991).

utility function and production functions will allow for counterfactual experiments in which shoppers reallocate themselves across stores and new promotional and pricing equilibria are computed.

3.1 Shopping Behavior

The model follows the discrete-choice literature and incorporates the store choice models developed by Bell et al. (1998) and Bell and Lattin (1998) that account for both store pricing decisions and geographical factor. The model assumes that prior to a shopping trip, a shopper h may receive promotion and price information from zero, one or more stores. Based on that information, the shopper constructs an expected *merchandising utility* of each store. The shopper also takes into account store valuations and shopping transportation costs, and chooses a single store that offers the greatest expected *shopping utility*.[9] Once in the chosen store when prices and merchandising activities are observed, for each product category the shopper chooses the optimal product (or not to purchase) that maximizes *category utility*. I first specify how shoppers make product choices within each category conditional on store choice, then describe the store choice decision making.

3.1.1 Within-category Product Choice

Let C denote the set of product categories. A category consists of a large number of related products. Examples of categories are ice-cream, frozen pizza and toothpaste. Let J_c denote the set of product alternatives of category $c \in C$. Once in the store and observing prices and other merchandising activities, for each category a shopper h chooses a product to maximize category utility. The product choices are independently made across categories. The outside option, no purchase from category c, is denoted 0_c. At the time of purchase, the indirect utility that shopper h obtains from product $j_c \in J_c \cup \{0_c\}$ from category c at time t in store s takes the form

$$w_{hst,j_c} = \chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c} + \epsilon_{hst,j_c}, \tag{1}$$

[9]The model does not take into account the "cherry-picker" behavior that a shopper would choose multiple stores in one shopping trip to assemble the bundle. Evidence shows that cherry pickers consist only a small fraction of consumers and that their negative contribution to store profitability is small (Fox and Hoch, 2005; Gauri et al., 2008a; Smith and Thomassen, 2012; Talukdar et al., 2008).

where p_{st,j_c} is the price; m_{st,j_c} is the promotion dummy; n_{st,j_c} is the dummy of in-store display,[10] y_{j_c} contains dummies of brand and package size; χ_c is the intrinsic utility of category c invariant over products within the category; ϵ_{hst,j_c} is an idiosyncratic shock assumed to follow a type I extreme value distribution, i.i.d. across products, categories, stores, shoppers, and periods. The parameters to be estimated are χ_c, α_c, $\beta_{c,1}$, $\beta_{c,2}$ and γ_c. Finally, the deterministic utility of the outside option, no purchase, is normalized to zero, thus $w_{hst,0_c} = \epsilon_{hst,0_c}$. Let ρ_{st,j_c} be the probability of choosing j_c, which is the probability that $j_c \in \arg\max w_{hst,j_c}, j_c \in J_c \cup \{0_c\}$. Following McFadden (1974), this probability is given by

$$\rho_{st,j_c} = \frac{\exp(\chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c})}{1 + \sum_{k_c \in J_c} \exp(\chi_c + \alpha_c p_{st,k_c} + \beta_{c,1} m_{st,k_c} + \beta_{c,2} n_{st,k_c} + \gamma_c y_{k_c})}, \tag{2}$$

and the expected utility of the category, *category utility*, is

$$v_{sct} = \log\Big(1 + \sum_{j_c \in J_c} \exp(\chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c})\Big). \tag{3}$$

There are three reasons to include brand and size dummies, y_{jc}. First, it improves model fit. Second, the brand-size combination captures unobserved product characteristics (e.g, quality). Therefore, the correlation between price and unobserved characteristics is accounted for and does not need instruments. [11] Third, the inclusion of brand and size dummies (their combination is sufficient to distinguish products of the same category) will not increase the number of coefficients as many as the number of choice alternatives. Thus it does not defeat the main motivation of the use of discrete-choice models.[12]

Let $\boldsymbol{x}_{st}$ denote merchandising decision that consists of pricing and promotion decisions for all products, $\boldsymbol{x}_{st} = (\boldsymbol{p}'_{st}, \boldsymbol{m}'_{st})'$, where $\boldsymbol{p}_{st}$ and $\boldsymbol{m}_{st}$ are vectors of price and promotion variables, respectively. The store merchandising utility is defined as the total category utility summing across all categories, given by

$$u_{st}(\boldsymbol{x}_{st}) = \sum_{c \in C} v_{sct}. \tag{4}$$

[10] In-store display, a kind of merchandising activity, is included in estimating consumer preferences, but is not treated as a choice variable in firm's problem. The merchandising activities thoughout this paper refer to pricing and promotions only.

[11] A potential correlation between price and unobserved characteristics may result from demand shocks, if the industry observes the shocks and account for them in pricing. In this model I assume out these time-specific shocks, as in Ho et al. (1998) and Bell et al. (1998).

[12] I also tried including SKU dummies, for the purpose of fully accounting for unobserved characteristics of each product, but regression results suggest that model fits are bad due to dimensionality.

The simple additive format of u_{st} makes the inclusion of category intrinsic utility χ_c clear: χ_c accounts for different "weights" of categories in store choice decision making (see below). A promoted product from a category with higher intrinsic utility is more effective in attracting customers.[13]

3.1.2 Store Choice

Prior to a shopping trip, for each store the shopper evaluates the merchandising utility and a purchase bundle, comprised of the optimal product of each category. The expected merchandising utility and the bundle depend on the shopper's price knowledge. It is assumed that the shopper passively receives promotion ads from stores and does not search, and that the probability of receiving an ad is independent across stores. Let ϕ_s denote the time-invariant probability of receiving promotion ads from store $s \in \{1, ..., S\}$. Let a dummy vector $ad_{ht} = (ad_{h1t}, ..., ad_{hst}, ..., ad_{hSt})'$ denote shopper h's ad exposure in t, satisfying $prob(ad_{hst} = 1) = \phi_s$.

The shopper has some prior knowledge of merchandising decision, a time-invariant distribution $F_s(\boldsymbol{x}_s)$. If she didn't receive promotion information prior to shopping from store s (uninformed, $ad_{hst} = 0$), she maintains the prior price information and forms expectation on product choices according to $F_s(\boldsymbol{x}_s)$. If she received promotion information from s (informed, $ad_{hst} = 1$), then she updates her knowledge conditional on promotion information, to $F_{st}(\boldsymbol{x}_{st}|\boldsymbol{x}_{st}^{prom})$, where the superscript *prom* denotes promoted items. The expected merchandising utilities of s, $\bar{u}_{st}(ad_{hst})$, perceived by uninformed and informed shoppers, respectively, are given by

$$\begin{aligned} \bar{u}_{hst}(ad_{hst} = 0) &= \int u_s(\boldsymbol{x}_s) dF_s(\boldsymbol{x}_s) \equiv \bar{u}_s, \\ \bar{u}_{hst}(ad_{hst} = 1) &= \int u_{st}(\boldsymbol{x}_{st}) dF_{st}(\boldsymbol{x}_{st}|\boldsymbol{x}_{st}^{prom}). \end{aligned} \tag{5}$$

Equation (5) implies that, as $\boldsymbol{x}_s$ is integrated out, the expected merchandising utility at a given store for an uninformed shopper is time-invariant, while for an informed shopper, the expected merchandising utility depends on promotion information received (price uncertainties of unpromoted items are integrated out). Furthermore, since $u_{st}(\cdot)$ is concave, price advertising can increase the store merchandising utility: $\bar{u}_{hst}(ad_{hst} = 1) > \bar{u}_{hst}(ad_{hst} = 0)$.

[13] The increase in shopping utility of store s due to a pure promotion can be written as

$$\frac{\Delta u_{st}}{\Delta m} = \exp(\chi_c) \cdot \frac{\exp(\alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c})}{1 + \sum_{J_c} \exp(\chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c})}.$$

Based on available price information, the shopper chooses a store to maximize shopping utility, which depends on the expected store merchandising utility $\bar{u}_{hst}$, store characteristics, and shopping transportation cost. The indirect utility function of shopper h at store s at time t takes the form:

$$U_{hst}(ad_{hst}) = \lambda_s + \iota \bar{u}_{hst}(ad_{hst}) + \kappa dist_{hs} + \zeta_{hst}, \tag{6}$$

where λ_s is the average store valuation that accounts for factors such as services and shopping environment; $\bar{u}_{hst}$ is the expected merchandising attractiveness at s that depends on ad exposure ad_{sht}; $dist_{hs}$ is the home-store distance of shopper h that allows the model to include geographic information specific to individual-store combination; ζ_{hst} is an idiosyncratic shock; ι and κ are parameters associated with expected store merchandising utility and home-store distance, respectively. The deterministic utility of the outside option, no shopping, is normalized to zero. Assuming ζ_{hst} follows type I extreme value distribution, i.i.d. across individuals, stores, and time, and following the discrete choice literature, shopper h will visit store s with probability

$$\eta_{hst}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}, ad_{ht}, dist_{hs}) = \frac{\exp\Big(\lambda_s + \iota \bar{u}_{hst}(ad_{hst}) + \kappa dist_{hs}\Big)}{1 + \sum_{q \in \{1,\dots,S\}} \exp\Big(\lambda_q + \iota \bar{u}_{hqt}(ad_{hqt}) + \kappa dist_{hq}\Big)}. \tag{7}$$

If the *cdf* of ad exposure is $\Omega(ad_{ht})$ and home-store distance follows a distribution $D(dist_{hs})$, the market share of s is

$$\bar{\eta}_{st}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}) = \int \int \eta_{hst}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}, ad_{ht}, dist_{hs}) d\Omega(ad_{ht}) dD(dist_{hs}). \tag{8}$$

Let MS be the market size. Let ρ_{st} be the vector of product choice probabilities, and mc_{st} be the vector of wholesale prices. The sales revenue R_{st}, which is the difference between total revenue and wholesale costs, but excluding promotion costs and fixed cost, is the following:

$$R_{st}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}) = MS \times \bar{\eta}_{st}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}) \times \rho_{st}(\boldsymbol{x}_{st})'(\boldsymbol{p}_{st} - mc_{st}). \tag{9}$$

3.2 Store Behavior

Retail stores simultaneously make pricing and promotion decisions for all products to maximize the expected store-level profit given their expectations on rivals' decisions [14] . For the ease of notation I drop the subscript for time t. A strategy played by store s is a mapping

[14] The assumption of store-level profit maximization follows models developed by Gauri et al. (2008b) and Hosken and Reiffen (2007), as opposed to category profit maximizing models, such as Bonnet et al. (2010), Bolton and Shankar (2003), and Bolton et al. (2010), Nevo (2001), and Villas-Boas (2007).

$\sigma_s : \mathcal{H}_s \to \mathcal{X}_s$ where $\mathcal{H}_s$ is the collection of store s's information sets, and $\mathcal{X}_s$ is the action set. The information at the time of decision is denoted H_s, where $H_s \in \mathcal{H}_s$. Let $\pi_s(\boldsymbol{x}_s, \boldsymbol{x}_{-s})$ be the profit of s. The store's rational strategy is $\boldsymbol{x}_s \in \arg\max E[\pi(\boldsymbol{x}_s, \boldsymbol{x}_{-s})|H_s]$, where $\boldsymbol{x}_s \in \mathcal{X}_s$. Formally, the store's problem can be written as [15]

$$\boldsymbol{x}_s = (\boldsymbol{p}_s', \boldsymbol{m}_s')' \in \arg\max E\left[\pi_s((\boldsymbol{p}_s', \boldsymbol{m}_s')', \boldsymbol{x}_{-s})|H_s\right]. \tag{10}$$

I assume a unit promotional cost θ_s will be incurred for each promoted product. The expected profit is the expected sales revenue minus the total promotion costs and fixed cost:

$$E\left[\pi_s(\boldsymbol{x}_s, \boldsymbol{x}_{-s})|H_s\right] = E\left[R_s(\boldsymbol{x}_s, \boldsymbol{x}_{-s})|H_s\right] - \theta_s \cdot (\mathbf{1}_J \cdot \boldsymbol{m}_s) - FC_s, \tag{11}$$

where $E\left[R_s(\boldsymbol{x}_s, \boldsymbol{x}_{-s})|H_s\right]$ is the expected revenue with wholesale costs subtracted (equation 9); $\mathbf{1}_J$ is a vector of ones; $\mathbf{1}_J \cdot \boldsymbol{m}_s$ represents the total number of promotions; and FC_s is the fixed cost.

The firm's problem can be decomposed into a discrete promotion decision making problem and a sub-problem of pricing conditional on promotion. In the sub-problem, assume the existence of an interior solution, $\boldsymbol{p}_s^*(\boldsymbol{m}_s)$. From (9) and (11), conditional on $\boldsymbol{m}_s$ the first-order condition with respect to $\boldsymbol{p}_s$ is

$$\begin{aligned} &\frac{\partial}{\partial \boldsymbol{p}_s} E[\pi_s((\boldsymbol{p}_s, \boldsymbol{m}_s), \boldsymbol{x}_{-s})|H_s] = \frac{\partial}{\partial \boldsymbol{p}_s} E[R_s((\boldsymbol{p}_s, \boldsymbol{m}_s), \boldsymbol{x}_{-s})|H_s] \\ &= 0 \\ &= E\left[\frac{\partial \bar{\eta}_s}{\partial \boldsymbol{p}_s} \cdot \rho_s'(\boldsymbol{p}_s - mc_s) + \bar{\eta}_s \cdot \left[\frac{\partial \rho_s}{\partial \boldsymbol{p}_s}\right](\boldsymbol{p}_s - mc_s) + \bar{\eta}_s \cdot \rho_s \middle| H_s\right]\Bigg|_{\boldsymbol{p}_s = \boldsymbol{p}_s^*(\boldsymbol{m}_s)}. \end{aligned} \tag{12}$$

Thus the store's problem can be transformed to a promotion optimization with the optimal price vector satisfying (12). Hereafter, I use $E\left[\pi_s(\boldsymbol{m}_s, \boldsymbol{x}_{-s})|H_s\right]$ to denote the store's objective function, omitting the implicit optimal price variable.

A necessary equilibrium condition is that the strategy played by the agent is at least as

[15] Practically, I solve for the optimal $(\boldsymbol{p}_s', \boldsymbol{m}_s')'$ subject to the restriction that only discounted prices can be promoted: $\underline{p}_{s,j_c} \leq p_{s,j_c} \leq \overline{p}_{s,j_c}$, if $m_{s,j_c} = 0$ and $\underline{p}_{s,j_c} \leq p_{s,j_c} \leq \widehat{p}_{s,j_c}$, if $m_{s,j_c} = 1$, where $\underline{p}_{s,j_c}$ and $\overline{p}_{s,j_c}$ are the bounds of an un-prompted price, and $\widehat{p}_{s,j_c}$ is the upper bound of a promoted price satisfying $\widehat{p}_{s,j_c} < \overline{p}_{s,j_c}$. This restriction reflecting the coordination of pricing and promotion decisions is implied by the observed price histories (by the econometrician). It also helps makes sure that the price vector numerically solved is of a reasonable magnitude.

good as any alternative. That is, the optimal choice of $\boldsymbol{m}_s$ satisfies

$$E\left[\pi_s(\boldsymbol{m}_s, \boldsymbol{x}_{-s})|H_s\right] \geq E\left[\pi_s(\boldsymbol{m}_s^{'}, \boldsymbol{x}_{-s})|H_s\right], \tag{13}$$

for all $\boldsymbol{m}_s^{'} \neq \boldsymbol{m}_s$. From equation (9), this implies the following condition:

$$\begin{aligned} &E\left[\Delta R_s(\boldsymbol{m}_s, \boldsymbol{m}_s^{'}, \boldsymbol{x}_{-s})|H_s\right] \equiv E\left[R_s(\boldsymbol{m}_s, \boldsymbol{x}_{-s})|H_s\right] - E\left[R_s(\boldsymbol{m}_s^{'}, \boldsymbol{x}_{-s})|H_s\right] \\ &\geq \theta_s \cdot \mathbf{1}_J \cdot (\boldsymbol{m}_s - \boldsymbol{m}_s^{'}). \end{aligned} \tag{14}$$

The above inequality implies that the unit promotion cost, θ_s, can be estimated by computing the difference between actual and counterfactual expected sale revenues generated by observed and alternative promotion decisions. To recover the counterfactual expected revenue $E\left[R_s(\boldsymbol{m}_s^{'}, \cdot)|H_s\right]$, a price vector associated with the alternative promotion decision, $\boldsymbol{p}_s^*(\boldsymbol{m}_s^{'})$, must be found using the first-order condition in (12).

4 Data

To carry out the empirical investigation, I use a dataset of individual scanner panel data across 24 product categories originally obtained from IRi (a retail market research company). The data was drawn from the metro area of a large U.S. city, and covers a 104-week period from June 1991 to June 1993. The market has 548 households of a total population 1267 and five retail stores. The dataset contains two components, household level data and store level data. The household level data includes records of 81,105 unique shopping trips over the period. For each household in a given week, it provides information on whether the household shops, which store is visited if it shops, which items are purchased, and how much is paid. The store-level component contains a history of merchandising activities, including prices, promotions, and in-store displays, during the 104-week window. The dataset also contains proxy measures for the distance to each store for each of the 548 households, using the households' and stores' five-digit zip codes. Since it is difficult to isolate the market of the five competing stores in the extent of geographical area or customer identity, I approximate market size, MS, by comparing the total quantities sold by the stores to the quantities purchased by the tracked households.[16]

Two of the five stores in this market explicitly advertise as operating an "every-day-low-price"(EDLP) format. The third store uses a "high-price-low-price" (HiLo) strategy with fre-

[16]For each category, the average consumption rate implied by the tracked purchase histories in the two-year period is computed. Then the ratio between tracked households' consumption rate and stores' sell rate, averaging over categories, derives the market size. The market size is estimated to be 54,535 households.

quent price adjustments. The remaining two stores are high tier (HT) retailers from the same chain. The five stores are denoted EDLP1, EDLP2, HiLo, HT1, and HT2, respectively. Figure 1 shows the geographical location of stores that was first published in Bell et al. (1998). The summary statistics of pricing and promotions of the stores, including average price levels, the average frequency of promotions, price cuts, and deep price cuts, are shown in Table 1. Market Share in the table refers to the proportion of store visits at a specific store, as opposed to the "usual" market share that is computed using quantities sold. The Average Price Level is indicated by the average price index, computed as the ratio between period-t price of a product and its regular price, weighted by market share. Deep Price Cuts are price reductions at least 15% below regular price. The statistics are consistent with stores' price positioning: EDLP stores have lower prices, HiLo stores offer more (deep) price cuts and promotions, and HT stores provide less frequent promotions and higher price levels.

In this model, since each SKU (stock keeping unit) is treated as a separate product and the total number of products is very large (6,364), the firm's profit maximizing problem becomes extremely complex. For this reason, a special effort was made to select categories and products. First, I select categories that are frequently bought, given information on quantity sold, while keeping some variety. 18 categories out of 24 were processed for the purpose of this study: Bacon, Butter, Breakfast Cereal, Toothpaste, Ground Coffee, Crackers, Laundry Detergent, Eggs, Hot Dogs, Ice Cream, Peanuts, Frozen Pizza, Potato Chip, Soap, Tissue Paper, Paper Towel, and Yogurt. Second, for each selected category, I eliminate items with small market share.[17] This reduces the number of items within each category from a range of 47 to 729 to a smaller range of 16 to 38, and the total number of products from 6364 to 474.

Statistics related to shopping trips are shown in the bottom half of Table 1. On average, shoppers visit grocery stores 1.56 time a week, spending 37 dollars per visit. The mean home-store distance is 2.7 miles, while the mean of the actual travel distance is 1.47 miles, implying a tendency to choose closer stores. Realizing that households may visit multiple stores in a given week, I keep the observation with the greatest amount of transaction in that week and remove others, in order to be compatible with the logit model. Smaller transactions are treated as unplanned or urgent purchases.

[17] Depending on category, I set the threshold of "small" market share to be 0.5 to 2 percent, balancing between the efficiency of logit regression and product variety.

5 Estimation

The goal of estimation is to find the promotion cost parameters, $\theta_s, s \in \{1, ..., S\}$. This requires first estimating demand and the wholesale prices that will be used to recover profits generated by alternative decisions. My estimation of the behavioral model will implement three major methodologies. First, the demand system will be estimated using standard logit regressions and simulation methods.[18] Second, the wholesale costs are estimated based on the store's first-order condition at observed merchandising decisions. Third, the promotional cost parameters are estimated using the moment inequality method.

5.1 Demand

The demand estimation contains two stages. In stage one, I estimate parameters associated with within-category product choice, $\Theta_1 = (\chi', \alpha', \beta_1', \beta_2', \gamma')'$, using logit regressions conditional on observed purchases. Among these parameters, α, β_1, β_2 and γ can be identified from market shares within each product category. The parameter of category-intrinsic utility, χ, is identified from purchase incidence. The probability of outside choice here accounts for the events that shoppers pay store visits but make no purchase from a given category.

In stage two, parameters related to store choices, $\Theta_2 = (\kappa, \iota, \lambda)'$, and ad exposure ($\phi$) are jointly estimated by maximizing the likelihood of the observed store choices given stage-one estimates, $\hat{\Theta}_1$. The *cdf* of prior knowledge, $\hat{F}_s(\boldsymbol{p}_s, \boldsymbol{m}_s)$, is approximated by the empirical distribution; the updated price knowledge, $\hat{F}_{st}(\boldsymbol{x}_{st}|\boldsymbol{x}_{st}^{prom})$, is simplified as follows: the promoted products' prices equal the advertised price, and the un-promoted products' prices equal their regular prices. Finally, the distribution of ad exposure $\Omega(ad)$ remains to be empirically specified. There are 2^S mutually different ad exposure statuses. Let $\mathcal{AD}$ denote the set of all possible statuses. Assuming shoppers are independently exposed to ads from different stores, the probability of status $ad = (ad_1, ..., ad_S)'$ is

$$prob(ad) = \prod_s \Big(ad_s \cdot \phi_s + (1 - ad_s)(1 - \phi_s) \Big). \tag{15}$$

The log-likelihood function of store choice is

$$l(\phi, \Theta_1, \Theta_2) = \sum_t \sum_h \sum_{ad \in \mathcal{AD}} prob(ad) \cdot \log \left(\sum_s \eta_{hst}(\boldsymbol{x}_t, ad, dist_{hs}; \Theta_1, \Theta_2) \cdot store_{hst} \right), \tag{16}$$

[18] Bell et al. (1998) and I use the same raw dataset to estimate store choice. But there is no price information updating in their mode.

where $store_{hst}$ equals 1 if store s is visited by h in t, and 0 otherwise. Given the parameter estimates of within-category choice preference, $\hat{\Theta}_1$, the identified ϕ and Θ_2 are the parameters that jointly maximize the store-choice likelihood:

$$(\phi, \Theta_2) \in \arg\max l(\phi, \hat{\Theta}_1, \Theta_2). \tag{17}$$

The store choice likelihood needs to be constructed by integrating over F_s, as store choice probability depends on price knowledge (see equations (5) and (7)). Practically, I compute this likelihood function using simulation by randomly drawing prices from $\hat{F}_s(\boldsymbol{x}_s)$ or $\hat{F}_{st}(\boldsymbol{x}_{st}|\boldsymbol{x}_{st}^{prom})$.

Besides jointly estimating ϕ and $\Theta_2 = (\lambda, \iota, \kappa)$, I estimate Θ_2 under the following two alternative assumptions to see how store choice estimates may be biased when restrictions are imposed to shoppers' price knowledge: (1) shoppers have perfect knowledge about promotion and price information ($\phi_s = 1, \forall s$); (2) shoppers have no better knowledge than the prior distribution ($\phi_s = 0, \forall s$). Under (1), the regressors are a constant, $dist_{hs}$, and u_{st} that is constructed with observed merchandising decisions and $\hat{\Theta}_1$. Under (2), since there is no time variation in store utility, the regressors include a constant, $dist_{hs}$, and the expected merchandising utility $\bar{u}_s$ constructed by simulation.

5.2 Supply

5.2.1 Wholesale Costs

To recover the counterfactual profits under alternative promotion decisions, the marginal cost vector (mc_s) must be known. However, I do not observe wholesale prices, distribution and shelving costs, or any other data that can be used to approximate this variable. Following the empirical I.O. literature (Bresnahan, 1987; Nevo, 2001; Porter, 1983), I estimate mc_s using the first-order condition of the firm's problem conditional on the observed merchandising decisions. Suppose the wholesale cost vector takes the form:

$$mc_{st} = mc_s + \tau_{st}, \tag{18}$$

where mc_s is the vector of the mean wholesale cost vector to be estimated, and τ_{st} is a vector of unobservable (to the econometrician) disturbances but are known by the store at the point of decision making, satisfying $E[\tau_{st}] = 0$. Sources of this cost disturbance may include variations in manufacturer's price and logistic costs. The first-order condition in equation (12) implies

$$mc_s + \tau_{st} = \boldsymbol{p}_{st} + E\left[\left[\left[\frac{\partial \bar{\eta}_{st}}{\partial \boldsymbol{p}_{st}}\right] \cdot \rho_{st} + \bar{\eta}_{st} \cdot \left[\frac{\partial \rho_{st}}{\partial \boldsymbol{p}_{st}}\right]\right]^{-1} (\bar{\eta}_{st} \cdot \rho_{st}) \middle| H_{st} \right]. \tag{19}$$

The mean wholesale cost vector mc_s is estimated by taking the average of (19) using the observed prices and product choice probabilities, and demand estimates. The numerical procedure includes integrating over the distributions of ad_{st}, $dist_{hs}$, and $F_{-s}(\boldsymbol{p}_{-s}, \boldsymbol{m}_{-s})$.

5.2.2 Market Share

Using the discrete distribution of ad exposure status, the market share in (8) becomes

$$\bar{\eta}_{st}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}, \phi) = \int \sum_{ad \in AD} prob(ad)\eta_{sht}(\boldsymbol{x}_{st}, \boldsymbol{x}_{-st}, ad, dist_{hs})dD(dist_{hs}). \tag{20}$$

5.2.3 Promotion Decisions

The goal in this section is to estimate the unit cost of promotion, θ_m, using equilibrium revenue and counterfactual revenue generated by alternative promotion decision $\boldsymbol{m}'_{st}$. I use moment inequality method, which allows me to circumvent the dimensionality issue and preserve the discrete nature of the variable. My estimation methods draw from Pakes (2010) and Pakes et al. (2011) and are similar to applications such as Ho (2009), Ishii (2011), and Katz (2007). Identification of the parameters is based on the necessary condition for a Bayes-Nash equilibrium that a store's expected profit generated by the observed choice is greater than counterfactual profits generated by alternative choices. The difference between the actual and the counterfactual profits provides the boundaries of promotional costs. The large size of the product space makes it possible to construct a sufficient number of alternative promotion decisions.

Following the literature of moment inequality approach, the cost function for promotion cost takes the form:

$$\theta_{st} = \theta_s + \tilde{\theta}_{st}, \tag{21}$$

where θ_s is the mean promotion cost of store s to be estimated; $\tilde{\theta}_{st}$ captures cost variations known to the store but not to the econometrician, and $\sum_t \tilde{\theta}_{st} = 0$. The promotion cost may vary due to variations in labor cost of the marketing team, advertising contracting between store and media, etc. The inequality condition in (14) implies that

$$E\left[\Delta R_{st}(\boldsymbol{m}_{st}, \boldsymbol{m}_{st}^{'}, \boldsymbol{x}_{-st})|H_{st}\right] \geq (\theta_s + \tilde{\theta}_{st}) \cdot (\mathbf{1}_J \cdot (\boldsymbol{m}_{st} - \boldsymbol{m}_{st}^{'})). \tag{22}$$

I consider small deviations from the observed promotions as alternatives, that is, holding everything else constant, increase or decrease $\boldsymbol{m_{st}} = \sigma(H_{st})$ for all H_{st} by one unit, so that $\mathbf{1}_J \cdot (\boldsymbol{m}_{st} - \boldsymbol{m}_{st}^{'}) = \pm 1$. This implies two classes of counterfactuals: to drop a promotion of a promoted item, and to add a promotion to an unpromoted item, keeping promotion decisions of all other items unchanged. Note that in the counterfactual, the deviated item will be repriced subject to the discount price constraint. I discuss the two classes of counterfactuals as follows.

Counterfactual 1. Drop the promotion of item j_c if it is currently promoted, i.e, $\boldsymbol{m}_{st}^{'} = \boldsymbol{m}_{st} - \boldsymbol{e}_{st,j_c}$ with $m_{st,j_c} = 1$, where $\boldsymbol{e}_{st,j_c}$ is a vector of length J with the j_cth element equals one and others zero. Dropping the promotion saves the promotional cost but results in a smaller expected revenue as it reduces the store's attractiveness. The equilibrium condition requires that the cost saved must not exceed the decrease in expected revenue:

$$E\left[\Delta R_{st}(\boldsymbol{m}_{st}, \boldsymbol{m}_{st}^{'}, \boldsymbol{x}_{-st})|H_{st}\right] \geq \theta_s + \tilde{\theta}_{st}. \tag{23}$$

Counterfactual 2. Add a promotion to a non-promoted item, so $\boldsymbol{m}_{st}^{''} = \boldsymbol{m}_{st} + \boldsymbol{e}_{st,j_c}$ with $m_{st,j_c} = 0$. In equilibrium the additional cost should not cover the increment in the expected revenue resulting from the extra promotion:

$$E\left[\Delta R_{st}(\boldsymbol{m}_{st}, \boldsymbol{m}_{st}^{''}, \boldsymbol{x}_{-st})|H_{st}\right] \geq -(\theta_s + \tilde{\theta}_{st}). \tag{24}$$

Suppose in observation t, the number of promoted items is $J_{st,1}$ and the number of unpromoted items is $J_{st,2}$. The sample analogue of inequalities (23) and (24) are [19]

$$\begin{aligned} \theta_s &\leq \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{st,1}}\sum_{\boldsymbol{m}_s^{'}}\Delta R_{st}(\boldsymbol{m}_{st}, \boldsymbol{m}_{st}^{'}, \boldsymbol{x}_{-st}) \equiv UB_s, \\ \theta_s &\geq -\frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{st,2}}\sum_{\boldsymbol{m}_s^{''}}\Delta R_{st}(\boldsymbol{m}_{st}, \boldsymbol{m}_{st}^{''}, \boldsymbol{x}_{-st}) \equiv LB_s. \end{aligned} \tag{25}$$

[19]Note that the cost disturbance $\tilde{\theta}_{st}$ on the right hand side of (23) and (24) are averaged out: for example, the right hand side of (23) becomes $\frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{st,1}}\sum_{\boldsymbol{m}_s^{'}}(\theta_s+\tilde{\theta}_{st}) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{st,1}}J_{st,1}\times(\theta_s+\tilde{\theta}_{st}) = \frac{1}{T}\sum_{t=1}^{T}(\theta_s+\tilde{\theta}_{st}) = \theta_s + \frac{1}{T}\sum_{t=1}^{T}\tilde{\theta}_s = \theta_s$. Thus LB_s and UB_s are consistent estimates of the bounds.

Confidence intervals of $(1-\alpha)$ level are constructed as in Pakes et al. (2011). The interval is the set of parameters that satisfy the sample moment restrictions with probability $(1-\alpha)$.

5.2.4 Computational Issues

Dimensionality

To compute expected profits, the optimal price vector under alternative promotion decisions must be found. This is to find the optimal prices in the sub-problem in (12) conditional on $\boldsymbol{m}'$ constructed from observed promotions. Moreover, in the counterfactual experiments where model parameters are exogenously changed (promotion costs and transportation costs), the new $\boldsymbol{p}$ and $\boldsymbol{m}$ must be jointly solved in order to find new equilibrium outcomes. However, the dimensionality issue and the discrete nature of promotion decisions make it practically impossible to solve for $\boldsymbol{p}$ and $\boldsymbol{m}$ using standard algorithms: first, searching for the optimal $\boldsymbol{p}$ in the continuous space conditional on promotion is itself time-exhausting and inefficient; second, searching for $\boldsymbol{m}$ is of complexity J^2 if the number of product items is J. [20]

I use principal component technique and factor analysis to deal with the first issue, and an "ordered" promotion decision rule for the second problem. Though the number of products is greatly reduced using the method discussed in Section 3, jointly solving for 474 prices in the continuous space is still a big challenge. The principal component technique is used to compress the large dimensional variable into a vector of much smaller dimension, thus the profit optimization problem can be solved in the reduced space. The principal component analysis on price variations shows that the first 12 components account for 80 percent of the overall price variations in the data. I project the price vector into the reduced space using a linear transformation consisting of the first 12 singular vectors (the loading coefficients), so that the search of optimal price is in the space of 12 dimensions instead of 474. Once the shorter optimal price vector is found, by solving a simple restricted linear programming problem, the real price vector is recovered using the second linear transformation, obtained using factor analysis, into the original space with 474 dimensions.

Next, I reduce the number of products in the choice set of promotion. Data shows that many of the 474 products considered are rarely promoted. Unfortunately, after removing these products from the choice set, the number of alternatives is still large. I compress the choice set by selecting items that are relatively frequently promoted (at least two standard deviations higher than the mean frequency of promotion). There are 52 products in this set.

[20]Heuristic algorithms such as genetic algorithm are available in solving the mixed-integer optimization problem, but they tend to be time-consuming when the number of integer variables to solve for is large.

For promotion decision making, I use a new algorithm to searching for the optimal $\boldsymbol{m}$, aiming to effectively reduce time consumption. Suppose the number of items considered for promotion is N ($N = 52$). First of all, recover the wholesale cost shocks τ_{st} that induce the time-varying store decisions. For each store, find the first optimal item of promotion that generates the greatest profit. This procedure nests finding the optimal price vector using the first-order condition with respect to $\boldsymbol{p}$ (equation (12)) and the recovered wholesale cost shock in t, τ_{st}. Then, conditional on the first promotion, choose the second optimal item of promotion, and record the increase (decrease) in profit. Iterate this procedure until the profit increase from the n^{th} to the $(n+1)^{th}$ promotion is less than the unit promotion cost (assuming discrete concavity of profit function). The optimization of price vector, using the method described above, is nested in each iteration. The computational complexity is N in the first iteration, $N-1$ in the second iteration, and so on. Therefore, the algorithm largely reduces computational complexity from 2^N, if search over all alternatives in the choice set, to at most $N(N+1)/2$.

Solving For Equilibria Under New Model Parameters

In counterfactual experiments, new Bayesian-Nash equilibria under changed model parameters must be found. One cannot use the distributions of store actions in the old equilibrium $F_s(\boldsymbol{x}_s;\Theta), s = 1, ..., S$ as the consistent belief upon which to compute the new distributions of optimal actions under the changed parameters $F_s(\boldsymbol{x}_s;\Theta'), s = 1, ..., S$, because the belief has to be consistent with the new action distributions. For this reason, I use an intreated algorithm that updates stores' optimal choices given the new parameter, then updates their belief using the updated store choices, until the iterated "market outcome" converges. Practically, I use the convergence of expected store utilities $\bar{u}_s$ to proxy the convergence of the market outcome. The algorithm nests the numerical method that jointly solves for $\boldsymbol{p}$ and $\boldsymbol{m}$ mentioned above. Also, since the new Bayesian-Nash equilibrium requires the same set of wholesale cost shocks, the distributions of shocks must be recovered. Formally,

1. Recover wholesale cost shocks $\hat{\tau}_{st}$ using (19) given estimated average wholesale costs $\hat{mc}_s$ and demand estimates.

2. Compute the expected store merchandising utilities of the old equilibrium, $(\bar{u}_1, ..., \bar{u}_S)^0$ using observed pricing and promotion decisions.

3. In each iteration $k \geq 1$,

   (a) Given expected store merchandising utilities $(\bar{u}_1, ..., \bar{u}_S)^{k-1}$, update the *distribution* of optimal actions of store $s = 1, ..., S$ using the recovered shocks $\hat{\tau}_{st}$. The new distribution of action profile is $(F_1(\boldsymbol{x}_1;\Theta'),...,F_S(\boldsymbol{x}_S;\Theta'))^k$, where $\boldsymbol{x}_s \sim F_s(\boldsymbol{x}_s;\Theta')$.

(b) Update the expected store merchandising utilities $(\bar{u}_1,...,\bar{u}_S)^k$ using the updated action distributions, $(F_1(\boldsymbol{x}_1;\Theta'),...,F_S(\boldsymbol{x}_S;\Theta'))^k$.

4. Iterate step 3 until the expected store merchandising utilities converge, $\left|(\bar{u}_1,...,\bar{u}_S)_t^{k+1} - (\bar{u}_1,...,\bar{u}_S)_t^k\right| < \epsilon$, where $\epsilon > 0$. The distribution of the new equilibrium action profile is $(F_1(\boldsymbol{x}_1;\Theta'),...,F_S(\boldsymbol{x}_S;\Theta'))^K$ where K is the number of iterations at convergence.

5. Finally, the welfare measures of the new equilibrium are the surpluses integrating over the distributions $F(\cdot)$. For example, $W(\Theta') = \int W(\boldsymbol{x}_1(\Theta'),...,\boldsymbol{x}_S(\Theta'))dF_1...dF_S$, where $(\boldsymbol{x}_1(\Theta'),...,\boldsymbol{x}_S(\Theta')) \sim (F_1(\boldsymbol{x}_1;\Theta'),...,F_S(\boldsymbol{x}_S;\Theta'))^K$.

5.3 A Summary of Estimation Procedures

To summarize my empirical implementation, I provide a roadmap to what needs to be accomplished in this section:

1. For each product category, estimate the parameters associated with within-category product choice given observed purchases, $\Theta_1 = (\chi', \alpha', \beta_1', \beta_2', \gamma')'$. This is stage one demand estimation;

2. Using step-one estimates $\hat{\Theta}_1$, jointly estimate parameters associated with store choice $\Theta_2 = (\kappa, \iota, \lambda)'$ and ads exposure ϕ. This is stage two demand estimation;

3. Using $\hat{\Theta}_1$, $\hat{\Theta}_2$ and $\hat{\phi}$, estimate wholesale cost vector mc_s;

4. Construct actual revenue $R(\boldsymbol{m},\cdot)$ and counterfactual revenue $R(\boldsymbol{m'},\cdot)$ at the observed firm choices;

5. Estimate promotion cost θ_s by finding the difference between $R(\boldsymbol{m},\cdot)$ and $R(\boldsymbol{m'},\cdot)$.

6 Results

6.1 Within-category Choice

I estimate demand in order to predict sales and profits generated by alternative pricing decisions. Table 2 displays the results of stage-1 demand estimation by regressing product choice probabilities on observable marketing activities and product characteristics. The regressors of column (i) includes prices, display and feature dummies only. Column (ii) also includes brand and size dummies. All price coefficients are of negative sign, and feature and display dummies affect utility positively. The estimate variances are shown in parenthesis. All estimates are significant at the 5% level. In column (ii), when brands and package sizes are controlled for, the price coefficients increase in absolute value (except Butter and Eggs), indicating that the unobserved characteristics correlate with price and that failure to account for the correlation would result in biased estimates of price parameters. This parallels the demand estimation results in Nevo (2001) and Hendel and Nevo (2006) where the inclusion of brand dummies, which fully accounts for the mean unobserved characteristics, leads to more negative coefficients of price. As for Butter and Eggs, the reason that price coefficients do not turn more negative when brand and size dummies are included might be the nature of the two categories: products are much less differentiated and the differences in unobserved characteristics are small.

I use intercept χ to measure the intrinsic category utilities. They also serve as "weights" in forming store attractiveness: frequently bought categories weigh more in store choice consideration. They cannot be identified from observed purchases only, as they are common for all products of the same category. They are identified using both observed purchases and outside-choice observations (i.e., shoppers who enter the store but didn't purchase the category). The estimates of χ varies significantly across categories. A small value of χ_c implies the purchase incidence of the category is low.

To see the effect of price promotion in driving sales, I compute the percentage change in choice probability at a price cut, averaging across products. The price cut takes the value of 15 percent of its regular price (deep price cut).The percentage change of choice probability is computed as follows:

$$\frac{\Delta\rho_{j_c}}{\rho_{j_c}} = \frac{1}{\rho_{j_c}}\left(\frac{\partial\rho_{j_c}}{\partial p_{j_c}} \times 0.15p_{j_c} + \frac{\Delta\rho_{j_c}}{\Delta m_{j_c}}\right) = (-\alpha_c \times 0.15p_{j_c} + \beta_{c,1})(1 - \rho_{j_c}).$$

Table 3 provides the maximum, the mean, and the standard deviation of the percentage change in market share in response to promoted price cuts for each category at the five stores. The results show that promoted price cuts are quite effective in driving sales: on average they cause 1 to 8 percent increases in market shares of the promoted item; for categories of Detergents, Hot Dogs, Tissue Papers, and Yogurt, they can cause an increase of one fourth to one third. Promotion with deep price cuts are considerably effective in these categories, partly because of the logit setting and the large number of items considered: since the percentage change is greater when the market share is smaller, a promoted price cut is more effective in categories with larger variety (Hot Dogs, Yogurt). In categories with much fewer products and less product differentiations, like Butter and Sugar, this effect is smaller. Another factor is the value of preference estimates that determine price and promotion sensitivities. Categories with higher α and/or β_1 (Detergent and Tissue Paper) tend to have large quantity effect.

6.2 Store Choice

Parameters of ad exposure probabilities and coefficients associated with store preferences (ϕ, λ, κ, ι) are jointly estimated by maximizing the likelihood of observed store choices using simulation and numerical search. The simulation here is to compute the store choice likelihood as discussed in Section 5.1, and numerical search is conducted to find the parameter value that maximizes the simulated log-likelihood. The process requires the expected marketing attractiveness constructed using the estimates obtained from stage-1 estimation, $(\alpha, \beta_1, \beta_2, \chi)$. The results are displayed in the first column of Table 4. Their variances are obtained by bootstrapping and are shown in parentheses. The estimated ad exposure probabilities ϕ_s are all significant at the 5% level, ranging from 0.03 to 0.19. As for the substitution pattern between merchandising attractiveness and travel distance, my estimates imply that a shopper would be indifferent between enjoying an additional promotion with a 15 percent price cut and traveling another 0.008 to 0.023 miles.

Besides jointly estimating ϕ and other store choice parameters, as mentioned above I estimate (λ, ι, κ) under restrictions $\phi_s = 0$ and $\phi_s = 1$, respectively. The results are displayed in the second and the third columns of Table 4. As the bottom row of Table 4 shows, these two hypotheses ($\phi_s = 0$ and $\phi_s = 1$) are rejected by likelihood ratio tests. When restrictions on ϕ are imposed, I still obtain negative distance coefficients, and the marketing attractiveness enters store attractiveness positively. As expected, the parameter associated with sensitivity to u_s, ι, is underestimated (overestimated) when restriction of $\phi = 1$ ($\phi = 0$) is imposed. However, the extent to which estimate of distance sensitivity is biased is more complicated. If a distant store offers frequent promotions which result in a large variation in u_s, restriction of $\phi = 0$ would underestimate κ, as the store visits actually attracted by promotion information is explained by a smaller travel sensitivity. If frequent promotions are offered by relatively nearby stores, κ would be overestimated when imposing $\phi = 0$. The biased estimates under the restriction $\phi = 0$ would imply a greater sensitivity to travel distance. The results indicate that the first scenario fits the market investigated: the distant store decides to offer a great number of promotions to avoid being squeezed out of the market. The last column of the table contains the estimates under the alternative hypothesis ($\phi_s = \phi$), which aims to test whether shoppers have the same ad exposure to all stores. Again, this hypothesis is rejected, implying that exposure to (or "reach" of) advertising may differ across stores.

To measure how effective a price promotion is in driving store visits, and, in stealing rivals' business, I simulate the percentage changes in self- and cross-store choice probabilities, when one store promotes a 15 percent price cut. The percentage changes in store choice probabilities are computed as

$$\frac{\Delta\eta_s}{\eta_s} = \frac{1}{\eta_s}(\frac{\partial\eta_s}{\partial p_{s,j_c}} \times 0.15p_{s,j_c} + \frac{\Delta\eta_s}{\Delta m_{s,j_c}}) = \int\int\int \iota \times \rho_{j_c}(-\alpha_c \times 0.15p_{s,j_c} + \beta_{c,1})(1-\eta_s) \times dFd\Omega dD,$$

$$\frac{\Delta\eta_q}{\eta_q} = -\frac{1}{\eta_q}(\frac{\partial\eta_q}{\partial p_{s,j_c}} \times 0.15p_{s,j_c} + \frac{\Delta\eta_q}{\Delta m_{s,j_c}}) = \int\int\int -\iota \times \rho_{j_c}(-\alpha_c \times 0.15p_{s,j_c} + \beta_{c,1}) \times \eta_s \times dFd\Omega dD,$$

where $\frac{\Delta\eta_s}{\eta_s}$ is the percentage change in self market share; $\frac{\Delta\eta_q}{\eta_q}$ is the percentage change in a rival's market share. The changes are averaged across products. I first compute the current market share implied by the estimates, then simulate the percentage changes using the estimated store choice parameters. The results are displayed in Table 5. They show that price promotion drives a 0.12 to 0.47 percent increase in self store visit probability, and causes a 0.02 to 0.52 percent decrease in rivals' market share. These numbers imply 7 to 42 extra store visits generated by a promoted price cut, with 2 to 13 customers stolen from each of the rival stores (computed at the estimated market size).

The parameter of ad exposure ϕ plays a crucial role in store competition: a store is able to attract a large amount of additional customers, either switched from rivals stores or non-shopping, if a good portion of them are able to respond to promotion information. As the results show, the magnitude of store choice semi-elasticities are closely related to the value of ad exposure probability. The store choice probability is the most elastic at EDLP2 for which ad exposure is the greatest, and least at HT1 for which ad coverage is the smallest. According to standard logit analysis, where the probability of being informed is assumed to be one, the elasticity of the choice alternative with the lowest choice probability (η_s) is the highest. However, the semi elasticity in here depends not only on its market share but also on the proportion of informed consumers (ϕ_s), as store choice probabilities of the uninformed consumers won't change.

6.3 Promotion Costs

The promotional costs are estimated by comparing the actual profits generated by the actual merchandising decisions, and the alternative profits caused by small deviations in $\boldsymbol{m}$. For the store-side observations that span over 104 weeks, the number of feature promotions varies quite a bit across stores and weeks (from 15 to 136), as does the total number of items on the shelf (1,399 to 2,356). The number of deviations is computed based on these two numbers at each store in each week.

Table 6 displays the estimates of promotion cost (per promotion per week) at each of the five stores. Because the distributions of lower and upper bounds overlap, points estimates are obtained. The points estimates of this cost range from about \$50 to \$214. The magnitudes of estimated bounds imply a substantial dispersion of promotional cost across stores. The cost is only \$50 at EDLP1, while it could be as high as 200 dollars or higher at HiLo and HT1. The 95% confidence intervals (bound estimates obtained when 95% of the inequalities are satisfied) imply wider ranges of this cost. The wide dispersion in promotion costs may suggest a dispersion in the efficiency of the marketing division at different stores.

To check if the estimates are reasonable in size, I compute stores' average total weekly and yearly expenditures on promotion, given the observed frequency of promotions (Table 1), and compare them with reported data. The lower bound estimates imply that for the stores investigated in this study, promotions would at least cost \$1,052 to \$4,500 per week, or \$54.7 k to \$234 k per year. The national average yearly ad spending per supermarket in 1993 is about \$13,324.[21] However, this national average would be too low, because there are reasons to believe that the promotion cost in the metropolitan area where the data is collected would be much higher than the national average. On the other hand, as equation (9) shows, the magnitudes of bound estimates are closely related to the approximated market size, which serves as a scalar in the process of estimation. Since market size is approximated by the average consumption rate of observed product categories of the tracked households, the bound estimates of promotion cost could be improved if better knowledge about this parameter is provided. For example, a larger household sample size and a broader range of categories.

[21]This spending is obtained using the total U.S. supermarket ad spending in 2012 (about \$800 million) divided by the number of supermarkets (37,053) and deflated back to 1993.

7 Market Efficiency

7.1 Current Equilibrium

Although theory does not provide unambiguous predictions on welfare, two effects of price advertising, the demand-creating effect and the business-stealing effect, are well understood. Price advertising announces product existence and reduces price uncertainty and therefore creates demand. But since the firm that provides additional advertising is unable to appropriate all of the resulting market surplus, the private advertising level tends to be socially inadequate. The business-stealing effect implies that advertising may be excessive (Tirole, 1988) because a firm is motivated by the profit margin that it would enjoy on a "stolen" consumer, while social welfare is not impacted by the redistribution of margins from one firm to another.

Advertising efficiency in the supermarket industry is complicated by a number of facts. First, price announcements create a significant amount of new demand when consumers are not aware of product availability, as in Bagwell (2007), Butters (1977), Grossman and Shapiro (1984), and Stegeman (1991). In the supermarket industry, however, product availabilities are typically well known. In the current model, new demand could be created if (1) the additional advertising reaches a consumer who wouldn't have shopped otherwise, but now decides to visit that store (the *new* consumers), or (2) it reaches a consumer who would have shopped at a rival store and now purchases a bigger bundle at a remote store (the *switched* consumers). Second, the business-stealing externality among competing supermarkets is greater than its counterpart in the single-product scenario. Due to basket shopping behavior, the firm undertaking price advertising is motivated by the profit margin of a product bundle, not just the margin of the promoted item. Third, market surplus created will be eroded by the increased transportation cost, in both cases where new demand is created: in (1), the cost of the new shopping trip made by the new shopper; in (2), the cost of the longer shopping distance of the switched consumer (theoretically, transportation costs would also be saved if consumers switch to some closer store).

In sum, market efficiency in this industry is complex. I attack this problem by numerically examining all components of market surplus at and in the neighborhood of the current equilibrium. Practically, I allow for a small deviation of one store's promotion level from the current equilibrium while keeping other stores' promotions unchanged, and simulate the changes and levels of total surplus and its components. If the additional unit of promotion improves (harms) market surplus, then the private promotion levels are socially inadequate (excessive). The optimal price vector conditional on the additional promotion will be updated. Section 5.2 discusses the computational challenges and solutions. In the following section I detail the measures of market surplus components for this model.

7.1.1 Surplus Measures

The market surplus for the retail market is the sum of total producer surplus and total consumer surplus, induced by store decisions $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_s, ..., \boldsymbol{x}_S)$, integrating over the distributions of store actions $F(\boldsymbol{x})$:

$$W = \int \sum_s PS_s(\boldsymbol{x})dF(\boldsymbol{x}) + \int \sum_h CS_h(\boldsymbol{x})dF(\boldsymbol{x}) \tag{26}$$

and its empirical analogue is

$$W = \frac{1}{T}\sum_{t=1}\sum_s PS_{st}(\boldsymbol{x}_t) + \frac{1}{T}\sum_{t=1}\sum_h CS_{ht}(\boldsymbol{x}_t) = PS(\boldsymbol{x}) + CS(\boldsymbol{x}), \tag{27}$$

where PS_{st} is the individual producer surplus in t, and CS_{ht} is the individual consumer surplus in t. Both are measured in dollars.

The individual producer surplus PS_{st} is the expected payoff (excluding fixed cost) of store s in period t induced by action profile $\boldsymbol{x}_t = (\boldsymbol{x}_{1t}, ..., \boldsymbol{x}_{St})$. It is computed using parameter estimates and store decisions as follows:

$$PS_{st}(\boldsymbol{x}) = \frac{MS}{H} \times \sum_h \sum_{ad_{ht}\in AD} prob(ad_{ht}) \cdot \hat{\eta}_{hst}(\boldsymbol{x}_t, ad_{ht}, dist_{hs}) \cdot \hat{\rho}_{st} \cdot (\boldsymbol{p}_{st} - \widehat{mc}_s) - \hat{\theta}_s \cdot (\mathbf{1}_J \cdot \boldsymbol{m}_{st}). \tag{28}$$

Consumer surplus of consumer h is her expected gain from a shopping trip. Notice that the utility from a shopping trip must be rescaled in terms of dollars. First I compute the surplus generated from purchase,

$$CS^{purchase}_{hst} = \sum_{c\in C} \frac{1}{|\widehat{\alpha}_c|}\hat{v}_{sct},$$

where the inverse of $\widehat{\alpha}_c$ is used to transform utility to purchase surplus measured in dollars. Then I use a scalar, $\bar{\alpha}$, to linearly transform the expected utility gain from a shopping trip to a dollar-measured surplus. $\bar{\alpha}$ is the ratio between expected merchandising utility and surplus from purchase, averaging across time and stores:

$$|\bar{\alpha}| = \frac{1}{S}\frac{1}{T}\sum_s\sum_t \frac{\widehat{\bar{u}}_{st}}{CS^{purchase}_{hst}}.$$

Finally, individual consumer surplus is given by

$$CS_{ht}(\boldsymbol{x}_t) = \frac{1}{|\bar{\alpha}|}\frac{1}{|\hat{\iota}|}\log\left(1 + \sum_s \exp\Big(\hat{\lambda}_s + \hat{\iota}\hat{\widehat{u}}_{hst}(ad_{hst}) + \hat{\kappa}dist_{hs}\Big)\right). \tag{29}$$

I decompose the change in market surplus. Let $W(\boldsymbol{x})$ denote the equilibrium market surplus induced by the equilibrium decisions $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_s, ..., \boldsymbol{x}_S)$, and let $W(\boldsymbol{x}^{'})$ denote the surplus induced by the deviation $\boldsymbol{x}^{'} = (\boldsymbol{x}_1, ..., \boldsymbol{x}^{'}_s, ..., \boldsymbol{x}_S)$, where $\boldsymbol{x}^{'}_s$ contains slightly increased or decreased promotion level $\boldsymbol{m}^{'}_s = \boldsymbol{m}_s \pm e_{j_c}$ and the associated optimal prices, while $\boldsymbol{x}_{q\neq s}$ are the same as in the equilibrium. Now consider s offering an extra unit of promotion. The change in market surplus induced by the deviation is $\Delta W^s = W(\boldsymbol{x}^{'}) - W(\boldsymbol{x})$.

$$\begin{aligned}
\Delta W^s =& W(\boldsymbol{x}^{'}) - W(\boldsymbol{x}) \\
=& (\Delta PS_s + \sum_{q\neq s}\Delta PS_q) + \Delta CS \\
=& \Delta E[R_s] - \theta_s + \sum_{q\neq s}\Delta E[R_q] + \Delta CS \\
=& \Delta E[R_s] - \theta_s + \sum_{q\neq s}\Delta E[R_q] + (\Delta\widetilde{CS} - \Delta TrC) \\
=& \Delta\widetilde{W} - \Delta TrC.
\end{aligned} \tag{30}$$

In the fourth line of (30), ΔW^s is decomposed into five parts: the change in own revenue $E[R_s]$, a unit cost of promotion θ_s, the change in rivals' expected revenues $\sum_{q\neq s} E[R_q]$, the change in consumer surplus generated by purchase bundles $\Delta\widetilde{CS}$ and the change in transportation cost ΔTrC. Here I distinguish consumer surplus excluding and including transportation: a surplus excluding transportation cost represents the surplus created by purchase bundles which measures the "immediate" demand creation, denoted by $\widetilde{CS}$, whereas a surplus including the cost is the net surplus after transportation erosion, denoted by CS, and $CS = \widetilde{CS} - TrC$. It is clear that $\Delta\widetilde{CS}$ represents the demand-creating effect of the deviation, ΔTrC represents the transportation erosion and $\sum_{q\neq s}\Delta E[R_q]$ represents the business-stealing effect.

7.1.2 Implications of the Current Equilibrium

Table 7 describes the current equilibrium, where market surplus W is composed into consumer surplus CS, profit of each of the five stores PS_s and transportation erosion (costs) TrC. In addition, Table 7 provides the variables describing consumer and store choices, including store choice probabilities, average price indexes, and advertising intensities measured by the number of promotions. [22]

Table 7 show that in the current equilibrium, transportation erosion takes a considerable portion of consumer surplus that would have been gained from purchase bundles ($2,672.79k excluding transportation cost and $1,702.28k including the cost), which implies that about 36% of consumer surplus gained from purchase has been eroded by transportation.

Table 8 illustrates market efficiency of the current equilibrium. I simulate the changes in market surplus ΔW, in consumer surplus ΔCS, in self revenue $\Delta E[R^s]$ and in each rival's revenue $\Delta E[R^q]$, when each of the five stores makes a hypothetical deviation by offering one more or one less unit of promotion, holding actions of other stores constant (the subscript s denotes the deviating store). In the computation the prices under the deviation will be re-optimized and profits are calculated accordingly. The top and bottom of Table 8 sections display simulation results for one more and one less promotion, respectively. In the top half of the table the diagonal numbers are positive and the non-diagonal numbers are negative, indicating that when the deviating store offers one more promotion, it increases its own expected revenue and reduces the expected revenues of rival stores; and we see the reverse when the deviating store withdraws the last promotion, as shown in the bottom half of the table. The changes in market surplus ΔW following the deviation are computed based on the point estimates of $\hat{\theta}_s$ in Table 6. I find that $\Delta W < 0$ (> 0) when one more (less) promotion is offered (ΔW is negative when HT1 or HT2 drops one promotion) which means that the equilibrium promotion levels are socially excessive: an additional promotion will not expand quantities sufficiently enough to offset the extra social cost, while withdrawing one will save the society more than the loss from declined sales. In general, the social cost of the one extra unit of promotion is just $\hat{\theta}_s$, but in the supermarket industry this social cost includes another component, the transportation costs due to store switching.

First of all, the third line of (30) indicates that in the neighborhood of equilibrium the sign of ΔW roughly depends on the values of $\sum_{q \neq s} \Delta E[R^q]$ and ΔCS, since the change in the deviating store's revenue $\Delta E[R^s]$ is very close to θ_s at the discrete optimization. In many models they respectively correspond to the business-stealing effect and the demand-creating effect, and whether in a particular market advertising is excessive or inadequate boils down to the comparison between their magnitudes. However, the transportation erosion complicates the comparison, which is why we need to further decompose the demand creation into two parts, $\Delta \widetilde{CS}$ that measures the "immediate" demand-creating effect, and the changes in transportation costs ΔTrC.

[22] The average price index is computed as the average ratio between the optimal (profit-maximizing) price and its observed regular price, weighted by within-category market share of each product.

As stated before, the incentive of price advertising with basket shopping is big because by offering an additional promotion, the store is able to appropriate not only the profit marginal of the promoted item but also the margins of other high-priced items in the same purchase bundle, which means a high loss in rivals' profit $\sum_{q \neq s} \Delta E[R^q]$. On the other hand, the demand-creating effect $\Delta\widetilde{CS}$ is very limited in the supermarket setting where the newly-created demand results from the price reduction itself and product existence is typically well known. Besides, the demand created will be eroded by transportation: since a transportation cost must be paid for when shopping, a consumer wouldn't shop unless the transportation cost can be at least offset by the surplus gain generated from the purchase bundle. This erosion applies to the new shoppers and the switched shoppers: for a new shopper, a transportation cost must be paid for traveling to store s; for a switched shopper, if she can obtain extra surplus from the bundle purchased at s, this extra surplus will be eroded by the increased transportation cost if she has to travel a longer distance (of course, for some switched shoppers, store s happens to be closer). Thus, the net demand creation would be even smaller if ΔTrC is positive (ΔTrC will not be bigger than $\Delta\widetilde{CS}$ if they are both positive, because consumers are at least as good as in equilibrium), which makes demand-creating effect less possible to outweigh the business-stealing effect.

In sum, shopping transportation costs are found to play a crucial role in the welfare implications of price advertising. They not only cause surplus erosion but also leas to worse erosion when firms offer better deals. This seems to suggest that if competition among the firms is intensified in some way, transportation would cause higher efficiency loss. In the next section I provide counterfactual experiments where (1) promotion costs are (exogenously) slightly changed, and (2) transportation costs are eliminated, to further examine whether and how market efficiency could be improved.

7.2 Counterfactual Experiments

7.2.1 Small Changes in Unit Promotion Costs

The results of the last section imply that the inefficiency of price advertising could be due to the higher transportation erosion. One possible way to reduce this erosion is to have the stores offer less promotions (and perhaps lower prices) so that consumers will shop at closer stores with higher probabilities. I obtain this goal by increasing all stores' unit promotion costs θ_s by the same percentage. It should be clarified that, given the overall increase in unit promotion costs, whether the new optimal price levels will be higher or lower is theoretically ambiguous. The numerical simulation, however, indicates that the raised promotion costs lead to less promotions and higher price levels. The counterfactual simulation shows that, counterintuitively, the increased promotion costs does not reduce market surplus; it results in not only lower promotion expenses but also smaller transportation costs.

Practically, due to the discrete nature of the promotion decision, I allow for an overall 5 percent deviation in θ_s to induce some changes in store decisions and, in turn, market outcome. For each experiment, I solve for each store's optimal price and promotion vectors given the new promotion cost, then compute surpluses, market shares, price levels, promotion intensities and profits induced by stores' optimal decisions.

Table 9 reports the market outcome variables and their respective percentage changes compared to the base case. When the price of price advertising decreases by 5%, retailers respond by intensifying promotions and pricing lower. We see that the market surplus has decreased by 170.64, when transportation erosion is taken account (After Transportation Erosion). This is a counterintuitive result because in general when one of the production inputs become cheaper the market surplus should increases. In this market, however, the intensified retail market attracts more store visits and thus higher transportation erosions (increase by 216.46). Consumers are better off because of the competition with CS increases by 1.27. Notice that consumers are at least as good as before despite of internalizing the higher transportation costs, as they can always stick to their store choices before the change. Stores now make smaller profits due to the intensified competition: they respond to the 5% decrease in θ_s by promoting more and pricing lower; as the promotion intensity has increased by more than 5%, promotion expenses have increased. As a result, despite of higher store visits (and higher quantities sold), PS has decreased by 171.91. After all, due to the large raise in transportation costs, the increase in CS is too small to outweigh the decrease in PS, and market surplus has decreased. This numerical outcome is in line with the theoretical finding in (Bester and Petrakis, 1995) that inefficiency would occur because of higher transportation costs paid to commute to a distant retailer that offers advertised low price. In contrast, when transportation erosion is not accounted (Before Transportation Erosion[23]), we see the usual outcome that market surplus has increased, because the increment in consumer surplus is large enough to offset the reduction in PS. Again, we see that almost all gain in consumer surplus has been eroded by transportation.

There are two sources of the increased transportation costs. One source is the new consumers created by the more intense promoting activities and lower prices. The shopping probability (the sum of store choice probabilities over all stores) increases by 1×10^{-4}, which means 0.01 percent of households has become new shoppers. The other possible source is the old shoppers,

[23]"Before Transportation Erosion" only means that the transportation costs occurred are not accounted in the surplus calculations; shoppers still take store locations into consideration when making store choice decisions.

if they now travel longer distances. Because the percentage increase in transportation cost is 0.02 percent, it is clear that consumers now travel slightly longer distances per trip. This means that the old consumers must have switched to some further-away stores in the new equilibrium.

When the unit promotion costs increase by 5%, we see the reverse. Market surplus has increased by $99.32, since the saving in transportation costs outweighs the decrease in surplus due to quantity shrinking. Again, departing from usual observations, the less-intensive competition (less promotions and higher price levels) generates higher market surplus.

As a result, the welfare implication of price advertising in this particular market departs from the usual conclusion that informative advertising is always welfare-improving; the welfare-harming effect of price advertising, the transportation erosion, plays a key role. Therefore, policy makers can improve the market efficiency by taxing price advertising in the supermarket retail market, where there exists erosions, transportation costs or other kinds.

7.2.2 On-line Shopping

In this experiment I simulate market outcome where transportation costs are eliminated. Since this roughly resembles the on-line grocery shopping, this counterfactual helps provide predictions on the new shopping regime.

To simplify online shopping behavior, I make three assumptions. First, product bundles are delivered by the store with free shipping[24], and thus home-store distance does not affect store choice. Second, all shoppers read promotion circulars posted by all stores and evaluate the expected store utilities before making store choice; once the store choice is made, the bundle of product choices is suboptimal conditional on store choice. Third, the unobserved store valuations that still affect store choice in online grocery shopping regime (though in-store experience may not related) are of the same quality as those in the traditional shopping regime. These assumptions allow me to have the following parameter setting: coefficient associated with transportation cost κ is set to zero, all shoppers are informed by promotions so $\phi_s = 1$, and store dummies λ_s remain the same. Notice that by setting κ to zero, the stores' local market power due to the spatial factor is removed.

How would the market perform under this shopping regime? Firstly, it is obvious that the new regime immediately avoids transportation erosion and therefore improves consumer surplus. Secondly, spatial models of firm competition predict that geographical locations are anticompetitive, because each firm naturally possesses some market power over customers who live close. For example, Hotelling's location model suggest that when the two firms' locations

[24]In reality the shipping fee may related to home-store distance. I assume shipping is free for simplicity.

are fixed but are able to set price, the geographical locations give firms market power. Therefore, I expect that when stores lose their neighbor shoppers competition among stores will be intensified. This is on-line shopping's indirect contribution to market surplus, as opposed to the direct effect from transportation avoidance. However, the overall effect on social welfare is still ambiguous: the stores could compete aggressively by spending much greater in price advertising which would outweigh the sum of the above two positive effects. After all, the welfare implication depends on the magnitudes of two major parts: (1) savings from transportation avoidance, (2) the demand creation effect due to intensified competition, which can be decomposed into consumer and store components. These components are quantified by the counterfactual simulation.

As for computation, the counterfactual outcome is simulated by finding a new equilibrium, in which allocation of consumers no longer depends on shopping distance and stores' decisions are adjusted accordingly. One technical difficulty of finding the new equilibrium is that in the new equilibrium agents' consistent belief about stores' actions will differ from the price distribution in the current equilibrium. This means that the price distribution in the base case, or the approximated distribution using observations, cannot be used for simulating the new outcome. Instead, a new price distribution as an approximation for consistent belief must be found. Starting from the base-case equilibrium, I numerically compute the counterfactual equilibrium by iterating market evolution until it converges. The criteria for convergence is that the expectation of merchandising utility, u_s, is sufficiently close to the value in the last iteration.

The simulation result is contained in Table 10. The top section lists consumer surplus, producer surplus, market surplus, and the probability of shopping. As expected, comparing to the base case, the new shopping regime intensifies competition and attracts more store visits. Prices are driven down by 5% to 12%, promotion intensities increase by 9% to 24%, and shopping probability increases by 37.75%. Because of the greater store visits, stores' profits largely increase from \$20.17k to \$27.15k, despite of the greater promotion costs and lower price in the more rigorous competition,

The evidence of local market power removal can be found in the stores' market shares. In this counterfactual, the five stores' market shares range from 11.94% to 29.10%, implying a smaller divergence compared to the base case, where they range from 10.14% to 32.55%. Actually, all stores that used to have an above 20% share now have a smaller portion of the market, and stores with a below 20% share now possess a greater portion of the market.

Consumer surplus has largely been improved from \$1,702.28k to \$3681.77k,. The elimination of transportation immediately saves them \$970.52 k, which is what they used to spend in the base case. The shoppers' benefit from the intensified competition can be obtained by comparing the counterfactual CS to the base-case $\widetilde{CS}$ (CS Before Transportation Erosion in Table 7), which is \$1,008.98k or an increase of 37.75% . Finally, the numerical results show that the new shopping regime is welfare improving: social welfare has been improved by about 115%, from \$1,722.45k to \$3,708.92k [25], 59.6% of which is contributed by the intensified competition.

8 Conclusion

The empirical significance of the interaction between informative price advertising and transportation erosion has not been well studied in the previous literature. This paper quantitatively establishes the importance of transportation costs to the economics of supermarket shopping. In particular, this paper quantifies the demand-creating effect, the business-stealing effect and the transportation erosion that determine the market performance of price advertising in a supermarket retailing oligopoly.

Its market efficiency is examined by simulating market outcomes in the neighborhood of the current equilibrium, and numerically measuring components of market surplus in the observed and the counterfactual outcomes. The decomposition of the market surplus allows me to quantify the three effects. My model accounts for both consumer shopping behavior and retailer merchandising behavior. Using scanner data of consumer shopping and store merchandising information, I structurally estimate demand and retailers' cost functions. The key to measuring efficiency is the estimation of promotion costs, which is done by adopting the moment inequality method. The techniques used here also include numerical methods that aim to mitigate large-dimensionality issues.

These structural estimates allow for simulations at and in the neighborhood of the equilibrium equilibrium. The numerical results highlight the empirical significance of transportation costs in a supermarket retail market with a geographical nature. Equilibrium promotion intensities are found to be socially excessive. In contrast to the usual result that cheaper production inputs lead to higher market surplus, the two counterfactual outcomes at varied promotion costs show the reverse: lower promotion costs actually increase advertising intensities, resulting in more long-distance shopping trips; though promotion is now cheaper and more demand is created, the total benefit is too small to offset the increased transportation costs. In contrast,

[25] My computation does not take delivery costs into account, but the welfare improving feature of online shopping would not be significantly affected even if delivery fee is strictly positive. By economies of scale the delivery cost of a grocery bundle can be very small, if orders are delivered using efficient logistic system, such as van trucks.

slightly higher promotion costs can improve efficiency by reducing advertising intensities and shopping frequencies.

In the second counterfactual simulation, zero transportation costs (roughly resembling the on-line shopping channel) remove supermarkets' local market power and therefore induce a more rigorous competition. In this case, both transportation erosion and stores' local market power being eliminated, market surplus is improved substantially, by 115 percent. Unfortunately, shipping costs are unknown to the researcher. For simplicity, the cost of shipping the goods from store to consumer is assumed zero. But a safe conjecture would be that the (social) cost of shipping would be significantly cheaper than that of shopping trips, as they exhibit great economies of scale.

References

Bagwell, K., 2007: The economic analysis of advertising. *Handbook of industrial organization*, **3**, 1701–1844.

Bell, D., T. Ho, and C. Tang, 1998: Determining where to shop: Fixed and variable costs of shopping. *Journal of Marketing Research*, 352–369.

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

Bester, H. and E. Petrakis, 1995: Price competition and advertising in oligopoly. *European Economic Review*, **39 (6)**, 1075–1088.

Bolton, R. and V. Shankar, 2003: An empirically derived taxonomy of retailer pricing and promotion strategies. *Journal of Retailing*, **79 (4)**, 213–224.

Bolton, R., V. Shankar, and D. Montoya, 2010: Recent trends and emerging practices in retailer pricing. *Retailing in the 21st Century*, 301–318.

Bonnet, C., P. Dubois, S. B. Villas Boas, and D. Klapper, 2010: Empirical evidence on the role of non linear wholesale pricing and vertical restraints on cost pass-through. *The Review of Economics and Statistics*, **(0)**.

Bresnahan, T. F., 1987: Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, 457–482.

Butters, G., 1977: Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, 465–491.

Dixit, A. and V. Norman, 1978: Advertising and welfare. *The Bell Journal of Economics*, 1–17.

Fox, E. and S. Hoch, 2005: Cherry-picking. *Journal of Marketing*, 46–62.

Gauri, D., K. Sudhir, and D. Talukdar, 2008a: The temporal and spatial dimensions of price search: Insights from matching household survey and purchase data. *Journal of Marketing Research*, **45 (2)**, 226–240.

Gauri, D., D. Talukdar, and B. Ratchford, 2008b: Empirical investigation of the impact of loss leader promotion on store and category performance in grocery industry. Tech. rep., Working Paper, Syracuse University.

Grossman, G. M. and C. Shapiro, 1984: Informative advertising with differentiated products. *The Review of Economic Studies*, **51 (1)**, 63–81.

Hendel, I. and A. Nevo, 2006: Measuring the implications of sales and consumer inventory behavior. *Econometrica*, **74 (6)**, 1637–1673.

Ho, K., 2009: Insurer-provider networks in the medical care market. *American Economic Review, 99(1): 393-430.*

Ho, T., C. Tang, and D. Bell, 1998: Rational shopping behavior and the option value of variable pricing. *Management Science*, **44 (12-Part-2)**, S145–S160.

Hosken, D. and D. Reiffen, 2007: Pricing behavior of multiproduct retailers. *The BE Journal of Theoretical Economics*, **7 (1)**.

Ishii, J., 2011: Compatibility, competition, and investment in network industries: Atm networks in the banking industry. *manuscript, Stanford University.*

Katz, M., 2007: Supermarkets and zoning laws. *Unpublished Ph.D. Dissertation, Harvard University.*

Kohls, R. L. and J. N. Uhl, 2001: *Marketing of Agricultural Products*. Prentice Hall.

Lal, R. and C. Matutes, 1994: Retail pricing and advertising strategies. *Journal of Business*, 345–370.

Levy, D., M. Bergen, S. Dutta, and R. Venable, 1997: The magnitude of menu costs: direct evidence from large us supermarket chains. *The Quarterly Journal of Economics*, **112 (3)**, 791–824.

Marshall, A., 1919: *Industry and Trade: A Study of Industrial Technique and Business Organization and of Their Influences on the Conditions of Various Classes and Nations*. MacMillan and Co., London.

McFadden, D., 1974: Conditional logit analysis of qualitative choice behavior.

Nevo, A., 2001: Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 307–342.

Nichols, L. M., 1985: Advertising and economic welfare. *The American Economic Review*, **75 (1)**, 213–218.

Pakes, A., 2010: Alternative models for moment inequalities. *Econometrica*, **78 (6)**, 1783–1822.

Pakes, A., J. Porter, K. Ho, and J. Ishii, 2011: Moment inequalities and their application. *Unpublished Manuscript, Harvard University*.

Porter, R. H., 1983: A study of cartel stability: The joint executive committee, 1880-1886. *The Bell Journal of Economics*, 301–314.

Roy, S., 2000: Strategic segmentation of a market. *International Journal of Industrial Organization*, **18 (8)**, 1279–1290.

Smith, H., 2004: Supermarket choice and supermarket competition in market equilibrium. *The Review of Economic Studies*, **71 (1)**, 235–263.

Smith, H. and O. Thomassen, 2012: A Model of Demand for Grocery Categories Mimeo, oxford University.

Stahl, I. and O. Dale, 1994: Oligopolistic pricing and advertising. *Journal of economic theory*, **64 (1)**, 162–177.

Stegeman, M., 1991: Advertising in competitive markets. *The American Economic Review*, 210–223.

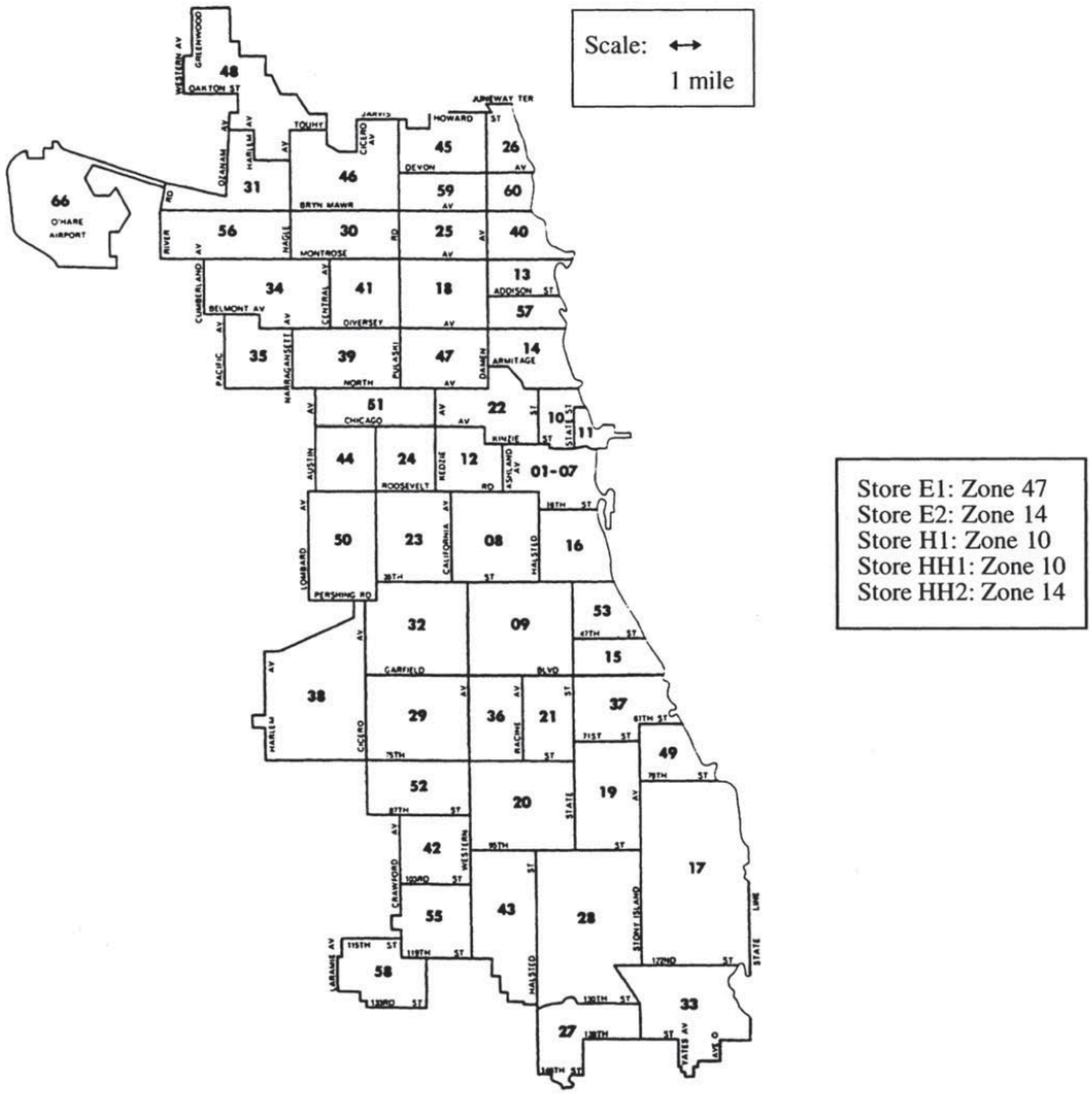
Stigler, G. J. and G. S. Becker, 1977: De gustibus non est disputandum. *The American Economic Review*, **67 (2)**, 76–90.

Talukdar, D., D. Gauri, and D. Grewal, 2008: Extreme cherry pickers: Prevalence, profile and profit impact in the frequently purchased goods market. Tech. rep., Working Paper, SUNY Buffalo.

Tirole, J., 1988: *The Theory of Industrial Organization: Jean Tirole*. MIT press.

Villas-Boas, S. B., 2007: Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, **74 (2)**, 625–652.

Figure 1: Locations of Stores*



* This figure is from Bell et al. (1998). The store codes in the legend, E1, E2, H1, HH1, HH2, correspond to the codes used in this paper EDLP1, DELP2, HiLo, HT1, HT2, respectively.

Table 1: Summary Statistics

Store Pricing and Promotions

			Weekly Frequency		
Store	Market Share[1]	Average Price Index[2]	Promotions	Price Cuts	Deep Price Cuts[3]
EDLP1	0.2147	0.8334	41.53	54.03	24.71
EDLP2	0.2679	0.8327	46.46	59.64	26.34
HILO	0.3255	0.8602	58.38	73.37	48.66
HT1	0.0905	0.9124	32.06	44.91	17.97
HT2	0.1014	0.9253	39.37	34.15	18.02

Households

	Mean	Std.
Family Size	2.31	1.37
Basket Spending ($)	37.04	32.34
Trips per Week	1.56	1.05
Home-Store Distance(miles)	2.70	2.47
Shopping Trip Distance	1.47	2.38

[1]Market share are calculated based on store visits.
[2]The average price index is computed as the ratio between period-t price of a product and its regular price, weighted by market share.
[3]Deep price cut is a price cut with at least 15% reduction.

Table 2: Within-Category Preferences

Category	i				ii			
	α_c	$\beta_{1,c}$	$\beta_{2,c}$	χ_c	α_c	$\beta_{1,c}$	$\beta_{2,c}$	χ_c
Bacon	-0.5551	1.1236	0.9911	1.6394	-0.8038	1.0573	0.6185	0.4638
	(0.0235)	(0.0540)	(0.0390)	(0.0480)	(0.0424)	(0.0588)	(0.0469)	(0.0346)
Butter	-0.5240	1.1684	1.2192	0.0893	-0.4879	1.1435	1.1582	1.1447
	(0.0223)	(0.0298)	(0.0284)	(0.0147)	(0.0285)	(0.0301)	(0.0291)	(0.0098)
Cereal	-0.0769	1.3594	0.9480	2.3350	-0.2272	1.3092	0.8724	0.2462
	(0.0174)	(0.0556)	(0.0688)	(0.0320)	(0.0227)	(0.0465)	(0.0586)	(0.0277)
Toothpaste	-0.4913	0.9557	0.8493	2.1228	-0.9273	0.9013	0.7265	0.2291
	(0.0334)	(0.0514)	(0.0805)	(0.0223)	(0.0678)	(0.0551)	(0.0227)	(0.0201)
Coffee	-0.0478	1.5691	1.4485	1.6968	-0.0730	1.3580	0.9752	2.7209
	(0.0088)	(0.0091)	(0.0397)	(0.0017)	(0.0142)	(0.0152)	(0.0416)	(0.0128)
Crackers	-0.2953	1.6941	0.8704	1.8944	-0.8026	1.5336	0.7612	0.0902
	(0.0277)	(0.0388)	(0.0493)	(0.0359)	(0.0432)	(0.0329)	(0.0691)	(0.0280)
Detergent	-0.1384	1.6808	1.0894	1.8578	-0.3069	1.4024	0.8965	0.6072
	(0.0053)	(0.0348)	(0.0434)	(0.0117)	(0.0149)	(0.0364)	(0.8965)	(0.0627)
Eggs	-1.0955	2.0032	1.5630	0.9806	-1.0732	1.8571	1.5075	1.3176
	(0.0412)	(0.1689)	(0.0585)	(0.0206)	(0.0738)	(0.1678)	(0.0613)	(0.0326)
Hot Dogs	-0.3778	1.1609	0.8919	0.4280	-0.6395	0.9984	0.6394	0.0949
	(0.0112)	(0.0399)	(0.0286)	(0.0117)	(0.0215)	(0.0411)	(0.0306)	(0.0217)
Ice Cream	-0.6180	0.9821	0.6960	1.7651	-0.9029	0.9349	0.3529	1.9362
	(0.0210)	(0.0643)	(0.0424)	(0.0101)	(0.0264)	(0.0647)	(0.0453)	(0.0195)
Peanuts	-0.3537	1.4422	0.9202	0.0796	-1.0187	1.2167	0.7213	2.0246
	(0.0234)	(0.0741)	(0.8818)	(0.0380)	(0.0297)	(0.0793)	(0.9312)	(0.0289)
Frozen Pizza	-0.1347	1.0583	1.0286	0.6923	-0.6724	0.7734	0.7552	0.3428
	(0.0119)	(0.0401)	(0.0359)	(0.0260)	(0.0368)	(0.0407)	(0.0428)	(0.0688)
Potato Chips	-0.2372	1.4522	1.1792	0.1154	-0.5471	1.6276	1.3291	1.1007
	(0.0255)	(0.0458)	(0.0512)	(0.0145)	(0.0311)	(0.0416)	(0.0313)	(0.0079)
Soap	-0.2783	1.1905	0.9849	0.2428	-0.6982	1.1275	0.8904	0.6679
	(0.0173)	(0.0503)	(0.0725)	(0.0208)	(0.0199)	(0.0524)	(0.0471)	(0.0527)
Sugar	-0.1537	1.3432	1.9931	2.0586	-0.9092	0.7830	0.6784	1.6425
	(0.0174)	(0.0552)	(0.0545)	(0.0134)	(0.0909)	(0.0622)	(0.0663)	(0.0390)
Tissue paper	-0.5665	0.8383	1.2240	1.7371	-0.8071	1.3745	1.0506	0.6147
	(0.0131)	(0.0253)	(0.0257)	(0.0105)	(0.0376)	(0.0268)	(0.0284)	(0.0200)
Paper Towel	-0.2751	1.1172	1.1819	0.7927	-0.9016	1.2179	1.0552	1.5670
	(0.0148)	(0.0272)	(0.0287)	(0.0284)	(0.0241)	(0.0285)	(0.0320)	(0.0297)
Yogurt	-0.1100	0.8034	1.1706	2.3756	-0.2283	0.7361	1.2024	1.3485
	(0.0193)	(0.0564)	(0.0378)	(0.0656)	(0.0197)	(0.0572)	(0.0426)	(0.0234)

The regressors of column i results include price, promotion and display dummies as explanatory variables, while in column ii they also includes brand and package size dummies.
All estimates are significant at at least 5% level.

Table 3: Elasticity of Price Cuts

Category	max	mean	std. dev.
Bacon	10.3913	2.9082	1.8645
Butter	5.2082	1.6199	0.7912
Cereal	12.2025	3.8505	1.7071
Toothpaste	17.1323	2.0892	1.9320
Coffee	18.6311	5.2345	2.7412
Crackers	16.3184	2.8645	1.6128
Detergent	35.0113	8.2882	6.0732
Eggs	19.5092	3.6860	3.1485
Hot Dogs	24.0420	3.6996	1.9806
Ice Cream	19.3625	4.9626	3.8713
Peanuts	12.3572	3.1889	1.7739
Frozen Pizza	12.6547	1.9869	1.3792
Potato Chips	11.6735	2.1346	1.7916
Soap	11.0077	3.2694	1.9840
Sugar	4.1252	1.6681	1.1158
Tissue Paper	30.2039	2.1968	4.0207
Paper Towel	19.8574	2.5594	3.3111
Yogurt	25.0165	2.3285	4.1306

*The elasticity of price cut are defined as percentage increase of within-category product choice probability due to an observed price cut, given by

$$\frac{\Delta\rho_{j_c}}{\rho_{j_c}} = \frac{1}{\rho_{j_c}}\left(\frac{\partial\rho_{j_c}}{\partial p_{j_c}}\times 0.15 p_{j_c} + \frac{\Delta\rho_{j_c}}{\Delta m_{j_c}}\right) = (-\alpha_c \times 0.15 p_{j_c} + \beta_{c,1})(1-\rho_{j_c}).$$

The percentage increase is calculated for each promoted item across all observed t.

Table 4: Store Preference

	joint	H_0: $\phi_s = 0$	H_0: $\phi_s = 1$	H_0: $\phi_s = \phi$
κ(distance)	-0.4659	-0.3789	-0.8977	-0.4328
	(0.1072)	(0.0833)	(0.0892)	(0.1377)
ι(merchandising utility)	0.0397	0.0986	0.0008	0.0432
	(0.0029)	(0.0021)	(0.0001)	(0.0238)
λ_s(store dummy)	0.5583	3.7267	0.4791	2.7982
	(0.0477)	(1.4742)	(0.0835)	(1.1721)
	0.2593	3.9117	0.6344	4.8903
	(0.0769)	(1.2008)	(0.1742)	(1.1905)
	0.1148	3.6297	0.6730	4.2877
	(0.0496)	(0.9073)	(0.2746)	(2.0033)
	-0.7954	2.7240	-0.3411	3.5786
	(0.1724)	(1.1355)	(0.7829)	(0.9982)
	-1.0659	2.6685	-0.4491	3.2090
	(0.3233)	(1.1084)	(0.0672)	(1.7953)
ϕ_s(prob. of ad exposure)	0.0350**			0.7230*
	(0.0102)			(0.2447)
	0.1958**			
	(0.0473)			
	0.1784**			
	(0.0509)			
	0.0277**			
	(0.0076)			
	0.0389**			
	(0.0122)			
Log-likelihood	-4.3941×10^4	-4.4118×10^5	-8.0716×10^5	-4.2946×10^4
D-statistic		7.9448×10^5***	1.5264×10^6***	1,990***

***significant at 1% level, **significant at 5% level, * significant at 10% level.
The null hypothesis (H_0) is respectively no promotion ad exposure from all stores ($\phi_s = 0$) , full ad exposure from all stores($\phi_s = 1$), and equal ad exposure from all stores ($\phi_s = \phi$).
Under the alternative hypothesis (joint), store preference coefficients and ad exposure probabilities are jointly estimated.

Table 5: Sensitivity of Store choice to Promoted Price Cuts

	EDLP1	EDLP2	HiLo	HT1	HT2
market share	0.1042	0.1515	0.2165	0.1381	0.0776
promotion offered by					
EDLP1	0.24	-0.05	-0.06	-0.11	-0.09
EDLP2	-0.16	0.47	-0.52	-0.02	-0.09
HiLo	-0.10	-0.06	0.36	-0.02	-0.07
HT1	-0.18	-0.10	-0.06	0.12	-0.07
HT2	-0.11	-0.07	-0.05	-0.02	0.18

The sensitivity of store choice is measured by the percentage changes in store choice probabilities (η) due to one promoted price cut of product at self or a rival store, given by

$$\frac{\Delta\eta_s}{\eta_s} = \frac{1}{\eta_s}(\frac{\partial\eta_s}{\partial p_{s,j_c}} \times 0.15p_{s,j_c} + \frac{\Delta\eta_s}{\Delta m_{s,j_c}}) = \int\int\int \iota \times \rho_{j_c}(-\alpha_c \times 0.15p_{s,j_c} + \beta_{c,1})(1-\eta_s) \times dFd\Omega dD,$$

$$\frac{\Delta\eta_s}{\eta_q} = -\frac{1}{\eta_q}(\frac{\partial\eta_q}{\partial p_{s,j_c}} \times 0.15p_{s,j_c} + \frac{\Delta\eta_q}{\Delta m_{s,j_c}}) = \int\int\int -\iota \times \rho_{j_c}(-\alpha_c \times 0.15p_{s,j_c} + \beta_{c,1}) \times \eta_s \times dFd\Omega dD.$$

Table 6: Estimates of Marginal Promotion Costs (θ)

	θ_s			95% Confidence Interval	
	LB	UB	point estimate	LB	UB
store					
EDLP1	3.90	6.12	5.01	[0.41, 12.56]	[1.06, 14.38]
EDLP2	6.98	18.61	12.79	[0.48, 29.67]	[5.93, 61.01]
HiLo	11.57	28.29	21.43	[0.90, 43.34]	[1.75, 81.48]
HT1	26.20	15.25	21.39	[0.49, 197.11]	[4.82, 42.83]
HT2	12.44	16.61	14.52	[0.03, 180.07]	[5.05, 46.73]

Per promotion per week. Unit: dollar.

Table 7: Current Equilibrium (Base Case)

	Transportation Erosion			
	After		Before	
CS	1,702.28		2,672.79	
PS	20.17		20.17	
W	1,722.45		2,692.97	
TrC	970.52			
Shopping Prob	0.6092			
store	Promo Freq	Average Price Index	Store Choice Prob.	PS_s
EDLP1	24.23	0.8450	0.0879	1.89
EDLP2	27.44	0.8943	0.1280	3.39
HiLo	30.46	0.8474	0.1968	7.83
HT1	20.43	0.9396	0.1266	4.52
HT2	22.75	0.9142	0.0699	2.53

Unit of all surpluses: 1 dollar.

Table 8: Market Efficiency

	$\Delta E[R]$					$\pm\hat{\theta}_s$	ΔCS	$\Delta\widetilde{CS}$	ΔTrC	ΔW	$\Delta\widetilde{W}$
	EDLP1	EDLP2	HiLo	HT1	HT2						
onemorepromotion											
EDLP1	4.63	-0.74	-1.08	-0.64	-0.58	-5.01	0.05	8.06	8.11	-3.39	4.67
EDLP2	-0.79	11.48	-3.79	-2.25	-2.31	-12.80	0.07	11.74	11.68	-10.40	1.27
HiLo	-0.58	-1.91	18.42	-4.69	-1.48	-21.43	0.07	11.09	11.03	-10.10	0.93
HT1	-0.48	-1.59	-6.59	16.51	-1.24	-21.39	0.04	6.30	6.27	-14.09	-7.81
HT2	-0.54	-2.01	-2.59	-1.54	12.75	-14.53	0.03	5.20	5.17	-8.43	-3.26
onelesspromotion											
EDLP1	-5.67	0.89	1.29	0.77	0.69	5.01	-0.07	-11.56	-11.49	2.90	-8.59
EDLP2	1.90	-27.77	9.03	5.42	5.49	12.80	-0.10	-18.04	-17.95	6.77	-11.18
HiLo	0.71	2.35	-22.59	5.79	1.83	21.43	-0.08	-12.92	-12.84	7.94	-4.90
HT1	0.61	1.99	8.31	-24.45	1.57	21.39	-0.07	-11.47	-11.40	8.71	-2.69
HT2	1.12	4.19	5.33	3.13	-26.10	14.53	-0.06	-10.31	-10.26	2.15	-8.11

Decomposing the change in market surplus: $\Delta W = \Delta E[R^s] + \sum_{q\neq s} \Delta E[R^q] + \Delta CS = (\Delta E[R^s] \pm \theta_s) + \sum_{q\neq s} \Delta E[R^q] + (\Delta\widetilde{CS} - \Delta TrC) = \Delta\widetilde{W} - \Delta TrC$.

Unit: dollar.

Table 9: Counterfactual 1 – Slight Changes in Advertising Costs

an overall 5% decrease			
	Transportation Erosion		
	After		Before
ΔCS	1.27		217.72
ΔPS	-171.91		-171.91
ΔW	-170.64		45.81
ΔTrC	216.46		
Δ Shopping Prob	0.0001		

	Promo Freq	Average Price Index	ΔPS_s
EDLP1	25.05	0.8260	12.90
EDLP2	28.14	0.8503	16.35
HiLo	33.03	0.8127	-176.67
HT1	21.81	0.9210	39.88
HT2	23.75	0.9052	-64.38

an overall 5% increase			
	Transportation Erosion		
	After		Before
ΔCS	-0.78		-130.54
ΔPS	100.10		100.10
ΔW	99.32		-30.44
ΔTrC	-129.77		
Δ Shopping Prob	-0.0001		

	Promo Freq	Average Price Index	ΔPS_s
EDLP1	23.34	0.8838	81.90
EDLP2	26.25	0.9042	104.45
HiLo	28.07	0.8902	243.39
HT1	17.72	0.9608	-448.31
HT2	21.47	0.9425	118.65

Unit of all surpluses: dollar.

Table 10: Counterfactual 2 – On-line Shopping Channel

	Counterfactual	Base Case	
		After Erosion	Before
CS	3,681.77 (+116.28%)	1,702.28	2,672.79
PS	27.15 (+34.61%)	20.17	20.17
W	3708.92 (+115.33%)	1,722.45	2,692.97
TrC	0	970.52	
Shopping Prob	0.8392 (+37.75%)	0.6092	

	Market Share	Promo Freq	Av. Price Ind.	Store Choice Prob	PS_s
EDLP1	18.39%	27.62	0.7728	0.1543	3.32
EDLP2	21.85%	34.02	0.7937	0.1834	4.86
HiLo	29.10%	33.20	0.8067	0.2442	9.71
HT1	18.72%	23.15	0.8344	0.1571	5.61
HT2	11.94%	24.82	0.8221	0.1002	3.63

Unit of all surpluses: dollar