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

Cixiu Gao*
Wuhan University

October, 2014

Abstract

This essay empirically examines the welfare effects of informative price advertising (promotions) in 
the supermarket retail industry, using structural estimation and individual scanner data. The welfare 
implications of promotions are determined by three effects: the demand-creating effect, the business 
stealing effect, and the transportation erosion effect. Transportation erosion is often ignored in the 
literature, but is found to be important in this empirical study. Using a spatial model accounting for 
consumer shopping behavior and retailer pricing behavior, structural estimates of consumer demand 
and the marginal costs of promotion are respectively obtained following the discrete choice literature 
and the moment inequality approach. The structural estimates from this model allow me to numer-
ically examine all components of market surplus and predict counterfactual outcomes. The results 
show a considerable erosion of consumer surplus due to transportation, and that private promotion 
intensities are socially excessive in the current equilibrium, but socially inadequate if transportation 
erosion were absent. Counterfactual experiments show that in this particular market intensified com-
petition may actually lead to lower market surplus because longer distances are traveled for better 
deals, outweighing the surplus gain due to quantity expansion; slightly higher promotion costs, which 
may be due to an advertising tax, improve market efficiency since retailers offer less promotions and 
consumers shop at closer stores. In the other counterfactual experiment with zero transportation costs 
(roughly resembles on-line shopping), I found that the market surplus is significantly improved (115%) 
with this modern shopping channel.

Keywords: Price Advertising; Market Efficiency; Structural Estimation; Moment Inequality Ap-
proach; 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 effect of costly information on market outcomes in the supermarket retail industry. 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. 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 advertising² do not provide unambiguous predictions of welfare effects.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 from purchase bundles. Moreover, if a consumer is attracted by price information at a distant store, the transportation cost is higher than if she shops at a local store. Thus, in a competition-intensified market where shoppers are more likely to travel longer distances, transportation costs erode the consumer surplus worse. 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 market performance of a supermarket oligopoly by simulating market outcomes following counterfactual increases and decreases in promotion intensities, and numerically measuring components of market surplus in the observed and counterfactual outcomes. This allows me to quantify the demand-creating effect, the business-stealing effect, and transportation erosion. My model accounts for both consumer shopping behavior and retail merchandising behavior. In a Bayesian-Nash

---

¹Data source: Kantar Media and http://online.wsj.com. See also Bolton et al. (2010) and Levy et al. (1997).
²This 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.
³For 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.
⁴On 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.
equilibrium, shoppers choose one optimal store to buy a bundle of products from; the competing retailers maximize 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.

Using scanner data of consumer shopping and store merchandising information, I structurally estimate consumer preferences and retailers’ cost functions. These structural estimates will allow me to simulate the equilibrium and counterfactual outcomes. My estimation of consumer preference follows fairly standard methods from the discrete choice literature. The more challenging estimation problems come from the firms’ side. A well-known problem that many empirical studies encountered is that the firms’ marginal costs (here, the wholesale prices) 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). A more difficult but crucial step is to estimate the (unobserved) marginal promotion costs based on the moment inequality approach (Pakes, 2010; Pakes et al., 2011). This estimation encounters several challenges which I solve by introducing some novel techniques explained below.

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 different decisions. The moment inequality approach uses the difference between the actual and the counterfactual profits to obtain bounds for promotion costs. This approach helps to circumvent dimensionality issues caused by the large number of products (identified by SKU, stock keeping unit) of the multi-product retailers.

Due to the large number of goods and the discrete nature of the promotion decisions, my estimation and simulation encounter two major challenges: (1) the search for the optimal large-dimensional price vector using regular methods is extremely inefficient; and (2) the search for the optimal discrete promotion decision is practically impossible. I adopt the techniques of principal component analysis and factor analysis to search for the optimal price vector, and a new algorithm that largely reduces computational complexity to solve for the promotion decision.

---

5The discrete and multi-product nature of supermarket retailers was noted by Smith (2004) and Smith and Thomassen (2012), but in these papers either the firm’s problem is absent or the large-dimensional price vector is collapsed into a price index.
The simulation results in the neighborhood of the observed equilibrium show that (1) transportation costs are very significant – in the current equilibrium, they erode consumer surplus by about 36 percent; (2) the private promotion intensities are socially excessive, because the social benefit created by the marginal unit of promotion (due to the increase in demand) is outweighed by social costs, to a large extent consisting of transportation costs; (3) however, when transportation costs are not accounted for in the same calculation, as if such erosion could be compensated in some way, private promotion intensities are found to be socially inadequate; (4) a reduction in promotion costs implies smaller market surplus, because the increased promotion intensity makes consumers become more likely to shop at distant stores and the higher social costs (mainly transportation costs) outweigh the social benefits (which here are due to both the demand created and the direct effect of the reduced promotion costs).

To further investigate the market performance, I simulate two counterfactual outcomes at varied promotion costs. I find that (1) slightly higher promotion costs, which could be due to an advertising tax, can actually improve efficiency by reducing the number of long store trips; (2) zero transportation costs (roughly resembling the on-line shopping channel) remove supermarkets’ local market power and therefore induce a more rigorous competition, which, however, no longer suffers from transportation erosion: market surplus is improved by a considerably large amount, 115 percent, compared to the current equilibrium.

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 food 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, and dry grocery 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%. In-store marketing activities (such as price promotion, display, food sampling, etc) and store branding largely account for the gap between retail and wholesale prices.

Supermarket retailers use an arsenal of marketing instruments and sophisticated pricing 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. It is therefore important to make economic predictions for supermarket industry with this new shopping regime.

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

---

7 Four Forces Shaping Competition in Grocery Retailing, industry report, Booz & Company.
behavior,\(^8\) consumers’ unawareness of product availability unless informed by price advertising\(^9\), 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 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.\(^{10}\) 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.

---

\(^8\)One exception is the Hotelling model of two-product duopoly constructed by Lal and Matutes (1994).

\(^9\)See, for example, Bagwell (2007), Butters (1977), Grossman and Shapiro (1984), and Stegeman (1991).

\(^{10}\)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).
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

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

(1)

where $p_{st,jc}$ is the price; $m_{st,jc}$ is the promotion dummy; $n_{st,jc}$ is the dummy of in-store display; $y_{jc}$ contains dummies of brand and package size; $\chi_c$ is the intrinsic utility of category $c$ invariant over products within the category; $\epsilon_{hst,jc}$ 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 $\chi_c$, $\alpha_c$, $\beta_{c,1}$, $\beta_{c,2}$ and $\gamma_c$. Finally, the deterministic utility of the outside option, no purchase, is normalized to zero, thus $w_{hst,0c} = \epsilon_{hst,0c}$. Let $\rho_{st,jc}$ be the probability of choosing $j_c$, which is the probability that $j_c \in \text{arg max} \ w_{hst,jc}, j_c \in J_c \cup \{0_c\}$. Following McFadden (1974), this probability is given by

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

(2)

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

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

(3)

There are three reasons to include brand and size dummies, $y_{jc}$. First, it improves model fit.

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 throughout this paper refer to pricing and promotions only.
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.\footnote{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).} 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.\footnote{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.}

Let \( x_{st} \) denote merchandising decision that consists of pricing and promotion decisions for all products, \( x_{st} = (p_{st}', m_{st}')' \), where \( p_{st} \) and \( m_{st} \) 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

\[
\begin{align*}
 u_{st}(x_{st}) = \sum_{c \in C} v_{sc}. \tag{4}
\end{align*}
\]

The simple additive format of \( u_{st} \) makes the inclusion of category intrinsic utility \( \chi_c \) clear: \( \chi_c \) 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.\footnote{The increase in shopping utility of store \( s \) due to a pure promotion can be written as

\[
\frac{\Delta u_{st}}{\Delta m} = \exp(\chi_c) \cdot \frac{\exp(\alpha_c p_{st,jc} + \beta_c,1 m_{st,jc} + \beta_c,2 n_{st,jc} + \gamma_c y_{jc})}{1 + \sum_{j,c} \exp(\chi_c + \alpha_c p_{st,jc} + \beta_c,1 m_{st,jc} + \beta_c,2 n_{st,jc} + \gamma_c y_{jc})}. \tag{5}
\]}

### 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 \( \phi_s \) denote the time-invariant probability of receiving promotion...
ads from store $s \in \{1, \ldots, S\}$. Let a dummy vector $ad_{ht} = (ad_{ht1}, \ldots, ad_{htst}, \ldots, ad_{htSt})'$ denote shopper $h$’s ad exposure in $t$, satisfying $\text{prob}(ad_{ht} = 1) = \phi_s$.

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

$$\bar{u}_{ht}(ad_{ht} = 0) = \int u_s(x_s)dF_s(x_s) \equiv \bar{u}_s,$$

$$\bar{u}_{ht}(ad_{ht} = 1) = \int u_{st}(x_{st})dF_{st}(x_{st}|x_{st}^{prom}).$$

Equation (6) implies that, as $x_s$ 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 un-promoted items are integrated out). Furthermore, since $u_{st}(\cdot)$ is concave, price advertising can increase the store merchandising utility: $\bar{u}_{ht}(ad_{ht} = 1) > \bar{u}_{ht}(ad_{ht} = 0)$.

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

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

where $\lambda_s$ is the average store valuation that accounts for factors such as services and shopping environment; $\bar{u}_{ht}$ 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; $\zeta_{hst}$ is an idiosyncratic shock; $\iota$ and $\kappa$ are parameters associated with expected store merchandising utility and home-store distance, respectively. The de-
terministic utility of the outside option, no shopping, is normalized to zero. Assuming $\zeta_{hst}$ 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

$$
\eta_{hst}(x_{st}, x_{-st}, ad_{ht}, dist_{hs}) = \frac{\exp(\lambda_s + \nu_h u_{hst}(ad_{ht}) + \kappa dist_{hs})}{1 + \sum_{q \in \{1,...,S\}} \exp(\lambda_q + \nu_h u_{hqt}(ad_{hqt}) + \kappa dist_{hq})}.
$$

(8)

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

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

(9)

Let $MS$ be the market size. Let $\rho_{st}$ 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:

$$
R_{st}(x_{st}, x_{-st}) = MS \times \bar{\eta}_{st}(x_{st}, x_{-st}) \times \rho_{st}(x_{st})'(p_{st} - mc_{st}).
$$

(10)

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. For the ease of notation I drop the subscript for time $t$. A strategy played by store $s$ is a mapping $\sigma_s : H_s \rightarrow X_s$ where $H_s$ is the collection of store $s$’s information sets, and $X_s$ is the action set. The information at the time of decision is denoted $H_s$, where $H_s \in H_s$. Let $\pi_s(x_s, x_{-s})$ be the profit of $s$. The store’s rational strategy

---

15 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).
is \( x_s \in \arg \max E[\pi(x_s, x_{-s})|H_s] \), where \( x_s \in X_s \). Formally, the store’s problem can be written as \(^{16}\)

\[
x_s = (p_s', m_s')' \in \arg \max E [\pi_s((p_s', m_s')', x_{-s})|H_s].
\] (11)

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

\[
E[\pi_s(x_s, x_{-s})|H_s] = E[R_s(x_s, x_{-s})|H_s] - \theta_s \cdot (1_J \cdot m_s) - FC_s,
\] (12)

where \( E[R_s(x_s, x_{-s})|H_s] \) is the expected revenue with wholesale costs subtracted (equation 10); \( 1_J \) is a vector of ones; \( 1_J \cdot m_s \) 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, \( p_s^*(m_s) \). From (10) and (12), conditional on \( m_s \) the first-order condition with respect to \( p_s \) is

\[
\frac{\partial}{\partial p_s} E[\pi_s((p_s, m_s), x_{-s})|H_s] = \frac{\partial}{\partial p_s} E[R_s((p_s, m_s), x_{-s})|H_s]
\] = 0

\[
= E \left[ \frac{\partial \bar{\eta}_s}{\partial p_s} \rho_s (p_s - mc_s) + \bar{\eta}_s \cdot \left[ \frac{\partial \rho_s}{\partial p_s} \right] (p_s - mc_s) + \bar{\eta}_s \cdot \rho_s \right]_{p_s = p_s^*(m_s)}.
\] (13)

Thus the store’s problem can be transformed to a promotion optimization with the optimal price vector satisfying (13). Hereafter, I use \( E[\pi_s(m_s, x_{-s})|H_s] \) 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 good as

---

\(^{16}\) Practically, I solve for the optimal \((p_s', m_s')\)' subject to the restriction that only discounted prices can be promoted: \( \overline{p}_{s,jc} \leq p_{s,jc} \leq \underline{p}_{s,jc} \), if \( m_{s,jc} = 0 \) and \( \overline{p}_{s,jc} \leq p_{s,jc} \leq \underline{p}_{s,jc} \), if \( m_{s,jc} = 1 \), where \( \overline{p}_{s,jc} \) and \( \underline{p}_{s,jc} \) are the bounds of an un-promoted price, and \( \overline{p}_{s,jc} \) is the upper bound of a promoted price satisfying \( \overline{p}_{s,jc} < \overline{p}_{s,jc} \). 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.
any alternative. That is, the optimal choice of $m_s$ satisfies

$$E[\pi_s(m_s, x_{-s})|H_s] \geq E[\pi_s(m'_s, x_{-s})|H_s],$$

(14)

for all $m'_s \neq m_s$. From equation (10), this implies the following condition:

$$E[\Delta R_s(m_s, m'_s, x_{-s})|H_s] \equiv E[R_s(m_s, x_{-s})|H_s] - E[R_s(m'_s, x_{-s})|H_s] \geq \theta_s \cdot 1_J \cdot (m_s - m'_s).$$

(15)

The above inequality implies that the unit promotion cost, $\theta_s$, 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 $E[R_s(m'_s, \cdot)|H_s]$, a price vector associated with the alternative promotion decision, $p^*_s(m'_s)$, must be found using the first-order condition in (13).

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 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 a total 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. 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
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 frequent 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 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.\(^{18}\) 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

\(^{17}\)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.

\(^{18}\)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.
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.

5 Estimation

The goal of estimation is to find the promotion cost parameters, \( \theta_s, s \in \{1, \ldots, S\} \). 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.\(^{19}\) 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, \( \alpha, \beta_1, \beta_2 \) and \( \gamma \) can be identified from market shares within each product category. The parameter of category-intrinsic utility, \( \chi \), 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, \( \hat{F}_s(p_s, m_s) \), is approximated by the empirical distribution; the updated price knowledge, \( \hat{F}_{st}(x_{st}|x_{st}^{prom}) \), 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 \( AD \) denote the set of all possible statuses. Assuming shoppers are independently exposed

\(^{19}\)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.
to ads from different stores, the probability of status \( ad = (ad_1, ..., ad_S)' \) is

\[
prob(ad) = \prod_s \left( ad_s \cdot \phi_s + (1 - ad_s)(1 - \phi_s) \right).
\]

The log-likelihood function of store choice is

\[
l(\phi, \Theta_1, \Theta_2) = \sum_t \sum_h \sum_{ad \in AD} prob(ad) \cdot \log \left( \sum_s \eta_{hs}(x_{st}, ad, dist_{hs}; \Theta_1, \Theta_2) \cdot store_{hs} \right),
\]

where \( store_{hs} \) 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 \( \phi \) and \( \Theta_2 \) are the parameters that jointly maximize the store-choice likelihood:

\[
(\phi, \Theta_2) \in \arg \max l(\phi, \hat{\Theta}_1, \Theta_2).
\]

The store choice likelihood needs to be constructed by integrating over \( F_s \), as store choice probability depends on price knowledge (see equations (6) and (8)). Practically, I compute this likelihood function using simulation by randomly drawing prices from \( \hat{F}_s(x_s) \) or \( \hat{F}_{st}(x_{st}|x_{prom}) \).

Besides jointly estimating \( \phi \) and \( \Theta_2 = (\lambda, \iota, \kappa) \), I estimate \( \Theta_2 \) 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 (\( \phi_s = 1, \) all \( s \)); (2) shoppers have no better knowledge than the prior distribution (\( \phi_s = 0, \) all \( s \)). 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 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 wholesale cost vector \( (mc_s) \) must be known. However, I do not observe wholesale prices or 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:

\[
mc_{st} = mc_s + \tau_{st},
\]

where \( mc_s \) is the vector of the mean wholesale cost vector to be estimated, and \( \tau_{st} \) 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, delivery cost, and packing cost. The first-order condition in equation (13) implies

\[
mc_s + \tau_{st} = p_{st} + E \left[ \frac{\partial \bar{\eta}_{st}}{\partial p_{st}} \cdot \rho_{st} + \bar{\eta}_{st} \cdot \left[ \frac{\partial \rho_{st}}{\partial p_{st}} \right] \right]^{-1} (\bar{\eta}_{st} \cdot \rho_{st}) \bigg| H_{st}. \]

The mean wholesale cost vector \( mc_s \) is estimated by taking the average of (20) 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 \( F_s(p_s, m_s) \).

5.2.2 Market Share

Using the discrete distribution of ad exposure status, the market share in (9) becomes

\[
\bar{\eta}_{st}(x_{st}, x_{-st}, \phi) = \int \sum_{ad \in AD} \text{prob(ad)} \eta_{ast}(x_{st}, x_{-st}, ad, dist_{hs}) dD(dist_{hs}).
\]

5.2.3 Promotion Decisions

The goal in this section is to estimate the unit cost of promotion, \( \theta_m \), using equilibrium revenue and counterfactual revenue generated by alternative promotion decision \( m'_{st} \). 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:

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

where $$\theta_s$$ 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 (15) implies that

$$E \left[ \Delta R_{st}(m_{st}, m'_{st}, x_{-st})|H_{st} \right] \geq (\theta_s + \tilde{\theta}_{st}) \cdot (1_J \cdot (m_{st} - m'_{st})).$$

(23)

I consider small deviations from the observed promotions as alternatives, that is, holding everything else constant, increase or decrease $$m_{st} = \sigma(H_{st})$$ for all $$H_{st}$$ by one unit, so that $$1_J \cdot (m_{st} - m'_{st}) = \pm 1$$. 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, $$m'_{st} = m_{st} - e_{st,j_c}$$ with $$m_{st,j_c} = 1$$, where $$e_{st,j_c}$$ is a vector of length $$J$$ with the $$j_c$$th 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:

$$E \left[ \Delta R_{st}(m_{st}, m'_{st}, x_{-st})|H_{st} \right] \geq \theta_s + \tilde{\theta}_{st}.$$  

(24)
Counterfactual 2. Add a promotion to a non-promoted item, so \( m''_{st} = m_{st} + e_{st,JC} \) with \( m_{st,JC} = 0 \). In equilibrium the additional cost should not cover the increment in the expected revenue resulting from the extra promotion:

\[
E \left[ \Delta R_{st}(m_{st}, m''_{st}, x_{-st}) | H_{st} \right] \geq -(\theta_s + \tilde{\theta}_{st}). \tag{25}
\]

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 (24) and (25) are \(^{20}\)

\[
\theta_s \leq \frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_{st,1}} \sum_{m_s'} \Delta R_{st}(m_{st}, m_s'_{st}, x_{-st}) \equiv UB_s,
\]

\[
\theta_s \geq -\frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_{st,2}} \sum_{m_s''} \Delta R_{st}(m_{st}, m_{st}'', x_{-st}) \equiv LB_s.
\]

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 (13) conditional on \( m' \) constructed from observed promotions. Moreover, in the counterfactual experiments where model parameters are exogenously changed (promotion costs and transportation costs), the new \( p \) and \( m \) 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 \( p \) and \( m \) using standard algorithms: first, searching for the optimal \( p \) in the continuous space conditional on promotion is itself time-exhausting and inefficient; second, searching for \( m \) is of complexity \( J^2 \) if the number of product

\(^{20}\)Note that the cost disturbance \( \tilde{\theta}_{st} \) on the right hand side of (24) and (25) are averaged out: for example, the right hand side of (24) becomes \( \frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_{st,1}} \sum_{m_s'} (\theta_s + \tilde{\theta}_{st}) = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{J_{st,1}} J_{st,1} \times (\theta_s + \tilde{\theta}_{st}) = \frac{1}{T} \sum_{t=1}^{T} (\theta_s + \tilde{\theta}_{st}) = \theta_s + \frac{1}{T} \sum_{t=1}^{T} \tilde{\theta}_s = \theta_s \). Thus \( LB_s \) and \( UB_s \) are consistent estimates of the bounds.
items is $J$. 21

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.

For promotion decision making, I use a new algorithm to searching for the optimal $m$, 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 $\tau_{st}$ 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 $p$ (equation (13)) and the recovered wholesale cost shock in $t$, $\tau_{st}$. 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

---

21Heuristic 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.
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 New Equilibria Under Changed 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 \( F_s(x_s; \Theta), s = 1, ..., S \) as the consistent belief upon which to compute the new distributions of optimal actions under the changed parameters \( F_s(x_s; \Theta'), s = 1, ..., S \), 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 \( p \) and \( m \) 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 (20) given estimated average wholesale costs \( \hat{\text{mc}}_s \) 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 \((F_1(x_1; \Theta'), ..., F_S(x_S; \Theta'))^k \), where \( x_s \sim F_s(x_s; \Theta') \).

   (b) Update the expected store merchandising utilities \((\bar{u}_1, ..., \bar{u}_S)^k \) using the updated action distributions, \((F_1(x_1; \Theta'), ..., F_S(x_S; \Theta'))^k \).
4. Iterate step 3 until the expected store merchandising utilities converge, 
\[ \left| (\bar{u}_1, ..., \bar{u}_S)^{k+1} - (\bar{u}_1, ..., \bar{u}_S)^k \right| < \epsilon, \text{ where } \epsilon > 0. \]

The distribution of the new equilibrium action profile is \( (F_1(x_1; \Theta'), ..., F_S(x_S; \Theta'))^K \) 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, 
\[
W(\Theta') = \int W(x_1(\Theta'), ..., x_S(\Theta'))dF_1 ... dF_S, \text{ where } (x_1(\Theta'), ..., x_S(\Theta')) \sim (F_1(x_1; \Theta'), ..., F_S(x_S; \Theta'))^K.
\]

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, \( \Theta_1 = (\chi, \alpha', \beta', \beta_1', \beta_2', \gamma')' \). 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 \( \phi \). 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 \( R(m, \cdot) \) and counterfactual revenue \( R(m', \cdot) \) at the observed firm choices;

5. Estimate promotion cost \( \theta_s \) by finding the difference between \( R(m, \cdot) \) and \( R(m', \cdot) \).

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 regression in 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 $\chi$ 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 $\chi$ varies significantly across categories. A small value of $\chi_c$ 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 due to simultaneous promotion and 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:

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

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 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 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 $\alpha$ and/or $\beta_1$ (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 ($\phi, \lambda, \kappa, \iota$) 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 $\phi$ 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 travelling another 0.008 to 0.023 miles.

Besides jointly estimating $\phi$ and other store choice parameters, as mentioned above I estimate ($\lambda, \iota, \kappa$) with restrictions $\phi = 0$ and $\phi = 1$, respectively. The results are displayed in the second and the third columns of Table 4. When restrictions on $\phi$ are imposed, I still obtain negative distance coefficients, and the marketing attractiveness enters store attractiveness positively. As expected, the parameter associated with sensitivity to $u_s$, $\iota$, is underestimated (overestimated) when restriction of $\phi = 1$ ($\phi = 0$) is imposed. However, how the 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 $\kappa$, 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, $\kappa$ 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 subjects investigated: the distant store decides to offer a great number of promotions to avoid being squeezed out of the market. As the bottom row of Table 4 shows, these two hypotheses \((\phi = 0\) and \(\phi = 1\)) are rejected by likelihood ratio tests. The last column of the table contains the estimates with alternative hypothesis that for all \(s\), \(\phi_s = \phi\), to test whether shoppers have the same exposure probability to all stores. The likelihood ratio test rejects the alternative hypothesis, which may imply that the "reach" of promotion advertising differs across stores, and/or shoppers have preference over stores’ ads.

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 store \(s\) offers a promoted 15 percent price cut. The percentage changes in store choice probabilities are computed using simulation as

\[
\frac{\Delta \eta_s}{\eta_s} = \frac{1}{\eta_s} \left( \frac{\partial \eta_s}{\partial p_{s,jc}} \times 0.15 p_{s,jc} + \frac{\Delta \eta_s}{\Delta m_{s,jc}} \right) = \int \int \int \iota \times \rho_{jc} (-\alpha_c \times 0.15 p_{s,jc} + \beta_{c,1}) (1 - \eta_s) \times dF d\Omega dD,
\]

\[
\frac{\Delta \eta_q}{\eta_q} = \frac{1}{\eta_q} \left( \frac{\partial \eta_q}{\partial p_{s,jc}} \times 0.15 p_{s,jc} + \frac{\Delta \eta_q}{\Delta m_{s,jc}} \right) = \int \int \int -\iota \times \rho_{jc} (-\alpha_c \times 0.15 p_{s,jc} + \beta_{c,1}) \times \eta_s \times dF d\Omega dD,
\]

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 rival store, computed using the estimated market size.

The parameter of ad exposure \(\phi\) 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 \((\eta_s)\) is the highest. However, the semi elasticity in here depends not only on its market share but also on the proportion of informed consumers \((\phi_s)\), 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 \(m\). 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. 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,

\(^{22}\)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.
as equation (10) 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.

7 Welfare

7.1 Market Efficiency of the 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 Welfare Measures

The market surplus for the retail market is the sum of total producer surplus and total consumer surplus, induced by store decisions \( x = (x_1, \ldots, x_s, \ldots, x_S) \), integrating over the distributions of store actions \( F(x) \):

\[
W = \int \sum_s PS_s(x)dF(x) + \int \sum_h CS_h(x)dF(x)
\]  

and its empirical analogue is

\[
W(x) = \frac{1}{T} \sum_{t=1}^{T} \sum_s PS_{st}(x_t) + \frac{1}{T} \sum_{t=1}^{T} \sum_h CS_{ht}(x_t) = PS(x) + CS(x),
\]

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 \( x_t = (x_{1t}, \ldots, x_{St}) \). It is computed using parameter estimates and
store decisions as follows:

\[ PS_{st}(x) = \frac{MS}{H} \times \sum_{h} \sum_{ad_{ht} \in AD} \text{prob}(ad_{ht}) \cdot \hat{\eta}_{hst}(x_t, ad_{ht}, dist_{hs}) \cdot \hat{\rho}_{st} \cdot (p_{st} - \hat{m}_c) - \hat{\theta}_s \cdot (1_J \cdot m_{st}). \]  

(29)

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,

\[ CS_{hst}^{\text{purchase}} = \sum_{c \in C} \frac{1}{|\alpha_c|} \hat{v}_{ct}, \]

where the inverse of \( \hat{\alpha}_c \) 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:

\[ |\bar{\alpha}| = \frac{1}{ST} \sum_s \sum_t \frac{\hat{u}_{st}}{CS_{hst}^{\text{purchase}}}. \]

Finally, individual consumer surplus is given by

\[ CS_{ht}(x_t) = \frac{1}{|\alpha|} \frac{1}{|\hat{\alpha}|} \log \left( 1 + \sum_s \exp(\hat{\lambda}_s + \hat{\nu}_{hst}(ad_{ht}) + \hat{\kappa}_{dist_{hs}}) \right). \]

(30)

I decompose the change in market surplus. Let \( W(x) \) denote the equilibrium market surplus induced by the equilibrium decisions \( x = (x_1, \ldots, x_s, \ldots, x_S) \), and let \( W(x') \) denote the surplus induced by the deviation \( x' = (x_1, \ldots, x'_s, \ldots, x_S) \), where \( x'_s \) contains slightly increased or decreased promotion level \( m'_s = m_s \pm e_j \) and the associated optimal prices, while \( x_{q \neq s} \) 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 $\Delta W^s = W(x') - W(x)$. 

$$
\Delta W^s = W(x') - W(x) = (\Delta PS_s + \sum_{q \neq s} \Delta PS_q) + \Delta CS
$$

$$
= \Delta E[R_s] - \theta_s + \sum_{q \neq s} \Delta E[R_q] + \Delta CS
$$

$$
= \Delta E[R_s] - \theta_s + \sum_{q \neq s} \Delta E[R_q] + (\Delta \bar{CS} - \Delta TrC)
$$

$$
= \Delta \bar{W} - \Delta TrC.
$$

In the fourth line of (31), $\Delta W^s$ is decomposed into five parts: the change in own revenue $E[R_s]$, a unit cost of promotion $\theta_s$, the change in rivals’ expected revenues $\sum_{q \neq s} E[R_q]$, the change in consumer surplus generated by purchase bundles $\Delta \bar{CS}$ and the change in transportation cost $\Delta TrC$. 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 $\bar{CS}$, whereas a surplus including the cost is the net surplus after transportation erosion, denoted by $CS$, and $CS = \bar{CS} - TrC$. It is clear that $\Delta \bar{CS}$ represents the demand-creating effect of the deviation, $\Delta TrC$ represents the transportation erosion and $\sum_{q \neq s} \Delta E[R_q]$ represents the business-stealing effect.

### 7.1.2 Simulation Results of Surpluses

As for numerical simulations, first of all I compute surplus components of the current equilibrium as the base case using observed store behavior and demand estimates. Other market outcome variables include the probability of shopping, and variables that describe each store’s behavior and profit: the total number of promotions, the average price index, store choice probability, and profit. 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. The results are displayed in Table 7.

To quantify the transportation erosion in price advertising, I compute consumer surpluses including
and excluding transportation cost. The results in 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.

Next, I simulate the changes in $\Delta W$, $\Delta CS$, $\Delta E[R^s]$ and $\sum_{q \neq s} \Delta E[R^q]$, 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. The prices under the deviation will be re-optimized and profits are calculated accordingly. Results are reported in Table 8, in which the top and bottom sections are for one extra 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 $\Delta W$ 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 ($\Delta W$ 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 components, the transportation costs due to store switching.

First of all, the third line of (31) indicates that in the neighborhood of equilibrium the sign of $\Delta W$ roughly depends on the values of $\sum_{q \neq s} \Delta E[R^q]$ and $\Delta CS$, since the change in the deviating store’s revenue $\Delta E[R^s]$ is very close to $\theta_s$ 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, $\Delta \tilde{CS}$ that measures the “immediate” demand-creating effect, and the changes in transportation costs $\Delta TrC$.

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 $\Delta \tilde{CS}$ 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 $\Delta TrC$ is positive ($\Delta TrC$ will not be bigger than $\Delta \tilde{CS}$ 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 lead 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) there are slightly manipulated promotion costs, and (2) transportation costs are eliminated, in order to see whether and how much 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 $\theta_s$ 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 $\theta_s$ 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 $\theta_s$ 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\textsuperscript{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 $1 \times 10^{-4}$, which means 0.01 percent of households has become new shoppers. The other possible source is the old shoppers, 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 Online Grocery 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\textsuperscript{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

\textsuperscript{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.

\textsuperscript{24}In reality the shipping fee may related to home-store distance. I assume shipping is free for simplicity.
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 $\kappa$ is set to zero, all shoppers are informed by promotions so $\phi_s = 1$, and store dummies $\lambda_s$ remain the same. Notice that by setting $\kappa$ 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. Second, 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 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 \$3,681.77k. The elimination of transportation immediately saves them \$970.52k, 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 \( \tilde{CS} \) (\( CS \) 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

This paper establishes the importance of transportation costs to the economics of supermarket shopping. This subject has been little studied in the previous literature. The empirical significance of the interaction between informative price advertising and transportation erosion has not been understood very well. This paper quantifies the demand-creating effect, the business-stealing effect and the

\(^{25}\)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.
transportation erosion that determine the welfare effects of price advertising.

The market performance of a supermarket oligopoly is examined by simulating market outcomes in the neighborhood of the current equilibrium, and numerically measuring components of market surplus in the observed and counterfactual outcomes. The decomposition of the market surplus allows me to quantify the three effects. As explained in Section 2, the supermarket industry has many unique characteristics that distinguish it from other oligopolistic industries. 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 consumer preferences and retailers’ cost functions. The key to measuring the welfare effects is estimating the stores’ unit promotion costs, which is done using the moment inequality method. These structural estimates allow for simulating the equilibrium and counterfactual outcomes.

The numerical results highlight the empirical significance of transportation costs in the supermarket retail market. Equilibrium promotion intensities are socially excessive, because the social benefits (due to the demand-creating effect) created by the last unit of promotion are too small to offset the social costs (mainly transportation costs). It is interesting to observe that the results would have been reversed - promotion levels would have been found to be socially inadequate - if transportation costs had not been taken into account.

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: the increased promotion intensities (and lower price levels) following slightly lower promotion costs make consumers more likely to shop at distant stores, and the higher social costs (mainly transportation costs) offset the social benefits (which are here due to both the demand created and the direct effect of the reduced promotion costs). Further surplus decomposition shows that higher demand has been created and surplus has been transferred from stores to consumers; but consumers who internalize transportation costs travel longer distances, eroding most of the surplus transferred from stores. This finding implies that slightly higher promotion costs, which could be due to an advertising tax, can actually improve efficiency by reducing shopping distances.

In the last 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 which market surplus is no longer eroded by transportation. Due to the eliminations of both transportation and stores’ local market power, market surplus is improved by a very large amount (31 percent).

This paper also introduced some new methods for dealing with large-dimensional and discrete choice problems. I use principal component technique and factor analysis to deal with the continuous large-dimension choice making (price vector), and a new algorithm that largely reduces the computation complexity for the discrete large-dimension choice making (promotion vector). The principal component technique is used to compress the large dimensional price vector into a vector of much smaller dimension, so that the profit optimization problem can be solved in the reduced space. The real price vector will then be recovered from the second linear transformation into the original space using factor analysis. Solving for the optimal discrete promotion vector would be practically infeasible due to the large number of products involved. I simplify the problem by assuming there is an order in which promoted products are chosen (conditional on the last chosen product). When the increment in profit associated with a chosen product is less than the unit promotion cost, the algorithm stops. The final promotion vector may not be fully optimal because in the algorithm each promotion decision is conditional on previous promotion decisions, whereas the fully optimal promotion decisions would be made simultaneously. However, since the complexity of this algorithm is only $N(N+1)/2$ if the length of the discrete vector is $N$, it makes solving for large-dimensional discrete vector (up to thousands of elements) practically feasible. These new methods may be helpful in future structural estimations involving continuous and discrete large-dimensional choice variables.
References


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


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

<table>
<thead>
<tr>
<th>Store Pricing and Promotions</th>
<th>Weekly Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Promotions</td>
</tr>
<tr>
<td>EDLP1</td>
<td>41.53</td>
</tr>
<tr>
<td>EDLP2</td>
<td>46.46</td>
</tr>
<tr>
<td>HILO</td>
<td>58.38</td>
</tr>
<tr>
<td>HT1</td>
<td>32.06</td>
</tr>
<tr>
<td>HT2</td>
<td>39.37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Households</th>
<th>Mean</th>
<th>Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Size</td>
<td>2.31</td>
<td>1.37</td>
</tr>
<tr>
<td>Basket Spending ($)</td>
<td>37.04</td>
<td>32.34</td>
</tr>
<tr>
<td>Trips per Week</td>
<td>1.56</td>
<td>1.05</td>
</tr>
<tr>
<td>Home-Store Distance(miles)</td>
<td>2.70</td>
<td>2.47</td>
</tr>
<tr>
<td>Shopping Trip Distance</td>
<td>1.47</td>
<td>2.38</td>
</tr>
</tbody>
</table>

<sup>1</sup>Market share are calculated based on store visits.
<sup>2</sup>The average price index is computed as the ratio between period-t price of a product and its regular price, weighted by market share.
<sup>3</sup>Deep price cut is a price cut with at least 15% reduction.
<table>
<thead>
<tr>
<th>Category</th>
<th>i</th>
<th></th>
<th></th>
<th></th>
<th>ii</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha_c$</td>
<td>$\beta_{1,c}$</td>
<td>$\beta_{2,c}$</td>
<td>$\chi_c$</td>
<td>$\alpha_c$</td>
<td>$\beta_{1,c}$</td>
<td>$\beta_{2,c}$</td>
<td>$\chi_c$</td>
</tr>
<tr>
<td>Bacon</td>
<td>-0.5551</td>
<td>1.1236</td>
<td>0.9911</td>
<td>1.6394</td>
<td>-0.8038</td>
<td>1.0573</td>
<td>0.6185</td>
<td>0.4638</td>
</tr>
<tr>
<td>Butt</td>
<td>-0.5240</td>
<td>1.1684</td>
<td>1.2192</td>
<td>0.0893</td>
<td>-0.4879</td>
<td>1.1435</td>
<td>1.5822</td>
<td>1.1447</td>
</tr>
<tr>
<td>Cereal</td>
<td>-0.0769</td>
<td>1.3594</td>
<td>0.9480</td>
<td>2.3350</td>
<td>-0.2272</td>
<td>1.3092</td>
<td>0.8724</td>
<td>0.2462</td>
</tr>
<tr>
<td>Toothpaste</td>
<td>-0.4913</td>
<td>0.9557</td>
<td>0.8493</td>
<td>2.1228</td>
<td>-0.9273</td>
<td>0.9013</td>
<td>0.7265</td>
<td>0.2291</td>
</tr>
<tr>
<td>Coffee</td>
<td>-0.0478</td>
<td>1.5691</td>
<td>1.4485</td>
<td>1.6968</td>
<td>-0.0730</td>
<td>1.3580</td>
<td>0.9752</td>
<td>2.7209</td>
</tr>
<tr>
<td>Crackers</td>
<td>-0.2953</td>
<td>1.6941</td>
<td>0.8704</td>
<td>1.8944</td>
<td>-0.8026</td>
<td>1.5336</td>
<td>0.7612</td>
<td>0.0902</td>
</tr>
<tr>
<td>Detergent</td>
<td>-0.1384</td>
<td>1.6808</td>
<td>0.4933</td>
<td>0.3593</td>
<td>-0.3069</td>
<td>1.4024</td>
<td>0.8965</td>
<td>0.0672</td>
</tr>
<tr>
<td>Eggs</td>
<td>-1.0955</td>
<td>2.0032</td>
<td>0.5630</td>
<td>0.9806</td>
<td>-1.0732</td>
<td>1.8571</td>
<td>1.5075</td>
<td>1.3176</td>
</tr>
<tr>
<td>Hot Dogs</td>
<td>-0.3778</td>
<td>1.1609</td>
<td>0.8919</td>
<td>0.4280</td>
<td>-0.6395</td>
<td>0.9984</td>
<td>0.6394</td>
<td>0.0949</td>
</tr>
<tr>
<td>Ice Cream</td>
<td>-0.6180</td>
<td>0.9821</td>
<td>0.6960</td>
<td>1.7651</td>
<td>-0.9029</td>
<td>0.9349</td>
<td>0.3529</td>
<td>1.3632</td>
</tr>
<tr>
<td>Peanuts</td>
<td>-0.3537</td>
<td>1.4422</td>
<td>0.9202</td>
<td>0.0796</td>
<td>-1.0187</td>
<td>1.2167</td>
<td>0.7213</td>
<td>2.0246</td>
</tr>
<tr>
<td>Frozen Pizza</td>
<td>-0.1347</td>
<td>1.0583</td>
<td>0.0286</td>
<td>0.0117</td>
<td>-0.0215</td>
<td>0.0411</td>
<td>0.0306</td>
<td>0.0217</td>
</tr>
<tr>
<td>Potato Chips</td>
<td>-0.2372</td>
<td>1.4522</td>
<td>1.1792</td>
<td>0.1154</td>
<td>-0.5471</td>
<td>1.6276</td>
<td>1.3291</td>
<td>1.1007</td>
</tr>
<tr>
<td>Soap</td>
<td>-0.2783</td>
<td>1.1905</td>
<td>0.9849</td>
<td>0.2428</td>
<td>-0.6982</td>
<td>1.1275</td>
<td>0.8904</td>
<td>0.6679</td>
</tr>
<tr>
<td>Sugar</td>
<td>-0.1537</td>
<td>1.3432</td>
<td>0.9931</td>
<td>2.0586</td>
<td>-0.9092</td>
<td>0.7830</td>
<td>0.6784</td>
<td>1.6425</td>
</tr>
<tr>
<td>Tissue paper</td>
<td>-0.5665</td>
<td>0.8383</td>
<td>0.1220</td>
<td>1.7371</td>
<td>-0.8071</td>
<td>1.3475</td>
<td>1.0506</td>
<td>0.6147</td>
</tr>
<tr>
<td>Paper Towel</td>
<td>-0.2751</td>
<td>1.1172</td>
<td>1.1819</td>
<td>0.7927</td>
<td>-0.9016</td>
<td>1.2719</td>
<td>1.0552</td>
<td>1.5670</td>
</tr>
<tr>
<td>Yogurt</td>
<td>-0.1100</td>
<td>0.8034</td>
<td>1.1706</td>
<td>2.3756</td>
<td>-0.2283</td>
<td>0.7361</td>
<td>1.2024</td>
<td>1.3485</td>
</tr>
</tbody>
</table>

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 the 5% level.
Table 3: Effects of Promoted Price Cuts – Percentage Increase* in Product Choice Prob.($\rho$)

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

*The percentage increase is

$$\frac{\Delta \rho_{jc}}{\rho_{jc}} = \frac{1}{\rho_{jc}} \left( \frac{\partial \rho_{jc}}{\partial p_{jc}} \times 0.15p_{jc} + \frac{\Delta \rho_{jc}}{\Delta m_{jc}} \right) = (-\alpha_c \times 0.15p_{jc} + \beta_c)(1-\rho_{jc})$$
Table 4: Store Preference

<table>
<thead>
<tr>
<th></th>
<th>$H_1$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>joint</td>
<td>$\phi = 0$</td>
<td>$\phi = 1$</td>
<td>$\phi_s = \phi$</td>
</tr>
<tr>
<td>$\kappa$(distance)</td>
<td>-0.4659</td>
<td>-0.3789</td>
<td>-0.8977</td>
<td>-0.4328</td>
</tr>
<tr>
<td></td>
<td>(0.1072)</td>
<td>(0.0833)</td>
<td>(0.0892)</td>
<td>(0.1377)</td>
</tr>
<tr>
<td>$\iota$(merchandising utility)</td>
<td>0.0397</td>
<td>0.0986</td>
<td>0.0008</td>
<td>0.0432</td>
</tr>
<tr>
<td></td>
<td>(0.0029)</td>
<td>(0.0021)</td>
<td>(0.0001)</td>
<td>(0.0238)</td>
</tr>
<tr>
<td>$\lambda_s$(store dummy)</td>
<td>0.5583</td>
<td>3.7267</td>
<td>0.4791</td>
<td>2.7982</td>
</tr>
<tr>
<td></td>
<td>(0.0477)</td>
<td>(1.4742)</td>
<td>(0.0835)</td>
<td>(1.1721)</td>
</tr>
<tr>
<td></td>
<td>0.2593</td>
<td>3.9117</td>
<td>0.6344</td>
<td>4.8903</td>
</tr>
<tr>
<td></td>
<td>(0.0769)</td>
<td>(1.2008)</td>
<td>(0.1742)</td>
<td>(1.1905)</td>
</tr>
<tr>
<td></td>
<td>0.1148</td>
<td>3.6297</td>
<td>0.6730</td>
<td>4.2877</td>
</tr>
<tr>
<td></td>
<td>(0.0496)</td>
<td>(0.9073)</td>
<td>(0.2746)</td>
<td>(2.0033)</td>
</tr>
<tr>
<td></td>
<td>-0.7954</td>
<td>2.7240</td>
<td>-0.3411</td>
<td>3.5786</td>
</tr>
<tr>
<td></td>
<td>(0.1724)</td>
<td>(1.1355)</td>
<td>(0.7829)</td>
<td>(0.9982)</td>
</tr>
<tr>
<td></td>
<td>-1.0659</td>
<td>2.6685</td>
<td>-0.4491</td>
<td>3.2090</td>
</tr>
<tr>
<td></td>
<td>(0.3233)</td>
<td>(1.1084)</td>
<td>(0.0672)</td>
<td>(1.7953)</td>
</tr>
<tr>
<td>$\phi_s$(prob. of ad exposure)</td>
<td>0.0350</td>
<td></td>
<td>0.7230</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0102)</td>
<td></td>
<td>(0.2447)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1958</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0473)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1784</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0509)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0389</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>$-4.3941 \times 10^4$</td>
<td>$-4.4118 \times 10^5$</td>
<td>$-8.0716 \times 10^5$</td>
<td>$-4.2946 \times 10^4$</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>$7.9448 \times 10^5^*$</td>
<td>$1.5264 \times 10^6^*$</td>
<td></td>
<td>$1.990^*$</td>
</tr>
</tbody>
</table>

*significant at the 1% level.

$U_{hst} = \lambda_s + \iota_{hst} + \kappa_{dist_{hst}} + \zeta_{hst}.$
Table 5: Effects of Promoted Price Cuts – Percentage Changes* in Store Choice Prob.($\eta$)

<table>
<thead>
<tr>
<th></th>
<th>EDLP1</th>
<th>EDLP2</th>
<th>HiLo</th>
<th>HT1</th>
<th>HT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>market share</td>
<td>0.1042</td>
<td>0.1515</td>
<td>0.2165</td>
<td>0.1381</td>
<td>0.0776</td>
</tr>
<tr>
<td>promotion offered by</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDLP1</td>
<td>0.24</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.11</td>
<td>-0.09</td>
</tr>
<tr>
<td>EDLP2</td>
<td>-0.16</td>
<td>0.47</td>
<td>-0.52</td>
<td>-0.02</td>
<td>-0.09</td>
</tr>
<tr>
<td>HiLo</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.36</td>
<td>-0.02</td>
<td>-0.07</td>
</tr>
<tr>
<td>HT1</td>
<td>-0.18</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.12</td>
<td>-0.07</td>
</tr>
<tr>
<td>HT2</td>
<td>-0.11</td>
<td>-0.07</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.18</td>
</tr>
</tbody>
</table>

*The percentage changes in store choice probabilities ($\eta$) due to a promoted price cut of product is

$$\frac{\Delta \eta_s}{\eta_s} = \frac{1}{\eta_s} \left( \frac{\partial \eta_s}{\partial p_{s,j,c}} \times 0.15p_{s,j,c} + \frac{\Delta \eta_s}{\Delta m_{s,j,c}} \right) = \int \int \int t \times \rho_{j,c} (-\alpha_c \times 0.15p_{s,j,c} + \beta_{c,1}) (1 - \eta) \times \frac{dFd\Omega dD}{\Sigma \eta_s \times dFd\Omega dD},$$

$$\frac{\Delta \eta_q}{\eta_q} = -\frac{1}{\eta_q} \left( \frac{\partial \eta_q}{\partial p_{s,j,c}} \times 0.15p_{s,j,c} + \frac{\Delta \eta_q}{\Delta m_{s,j,c}} \right) = \int \int \int -t \times \rho_{j,c} (-\alpha_c \times 0.15p_{s,j,c} + \beta_{c,1}) \times \eta \times \frac{dFd\Omega dD}{\Sigma \eta_s \times dFd\Omega dD}.$$
Table 7: Current Equilibrium (Base Case)

<table>
<thead>
<tr>
<th></th>
<th>Transportation Erosion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>After</td>
<td>Before</td>
</tr>
<tr>
<td>CS</td>
<td>1,702.28</td>
<td>2,672.79</td>
</tr>
<tr>
<td>PS</td>
<td>20.17</td>
<td>20.17</td>
</tr>
<tr>
<td>W</td>
<td>1,722.45</td>
<td>2,692.97</td>
</tr>
<tr>
<td>Total Transportation Cost (x10^3)</td>
<td>970.52</td>
<td></td>
</tr>
<tr>
<td>Shopping Prob</td>
<td>0.6092</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Promo Freq</th>
<th>Average Price Index</th>
<th>Store Choice Prob.</th>
<th>PS_s</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDLP1</td>
<td>24.23</td>
<td>0.8450</td>
<td>0.0879</td>
</tr>
<tr>
<td>EDLP2</td>
<td>27.44</td>
<td>0.8943</td>
<td>0.1280</td>
</tr>
<tr>
<td>HiLo</td>
<td>30.46</td>
<td>0.8474</td>
<td>0.1968</td>
</tr>
<tr>
<td>HT1</td>
<td>20.43</td>
<td>0.9396</td>
<td>0.1266</td>
</tr>
<tr>
<td>HT2</td>
<td>22.75</td>
<td>0.9142</td>
<td>0.0699</td>
</tr>
</tbody>
</table>

Unit: 1,000 dollars.

Table 8: Market Efficiency

<table>
<thead>
<tr>
<th>m' = m + e</th>
<th>∆E[R]</th>
<th>±θ_s</th>
<th>∆CS</th>
<th>∆̃CS</th>
<th>∆TrC</th>
<th>∆W</th>
<th>∆̃W</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDLP1</td>
<td>4.63</td>
<td>-0.74</td>
<td>-1.08</td>
<td>-0.64</td>
<td>-0.58</td>
<td>0.05</td>
<td>8.06</td>
</tr>
<tr>
<td>EDLP2</td>
<td>-0.79</td>
<td>11.48</td>
<td>-3.79</td>
<td>-2.25</td>
<td>-2.31</td>
<td>-12.80</td>
<td>0.07</td>
</tr>
<tr>
<td>HiLo</td>
<td>-0.58</td>
<td>-1.91</td>
<td>18.42</td>
<td>-4.69</td>
<td>-1.48</td>
<td>-21.43</td>
<td>0.07</td>
</tr>
<tr>
<td>HT1</td>
<td>-0.48</td>
<td>-1.59</td>
<td>-6.59</td>
<td>16.51</td>
<td>-1.24</td>
<td>-21.39</td>
<td>0.04</td>
</tr>
<tr>
<td>HT2</td>
<td>-0.54</td>
<td>-2.01</td>
<td>-2.59</td>
<td>-1.54</td>
<td>12.75</td>
<td>-14.53</td>
<td>0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m' = m - e</th>
<th>∆E[R]</th>
<th>±θ_s</th>
<th>∆CS</th>
<th>∆̃CS</th>
<th>∆TrC</th>
<th>∆W</th>
<th>∆̃W</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDLP1</td>
<td>-5.67</td>
<td>0.89</td>
<td>1.29</td>
<td>0.77</td>
<td>0.69</td>
<td>5.01</td>
<td>-0.07</td>
</tr>
<tr>
<td>EDLP2</td>
<td>1.90</td>
<td>-27.77</td>
<td>9.03</td>
<td>5.42</td>
<td>5.49</td>
<td>12.80</td>
<td>-0.10</td>
</tr>
<tr>
<td>HiLo</td>
<td>0.71</td>
<td>2.35</td>
<td>-22.59</td>
<td>5.79</td>
<td>1.83</td>
<td>21.43</td>
<td>-0.08</td>
</tr>
<tr>
<td>HT1</td>
<td>0.61</td>
<td>1.99</td>
<td>8.31</td>
<td>-24.45</td>
<td>1.57</td>
<td>21.39</td>
<td>-0.07</td>
</tr>
<tr>
<td>HT2</td>
<td>1.12</td>
<td>4.19</td>
<td>5.33</td>
<td>3.13</td>
<td>26.10</td>
<td>14.53</td>
<td>-0.06</td>
</tr>
</tbody>
</table>

\( \Delta W = \Delta E[R^1] + \sum_{q \neq s} \Delta E[R^q] + \Delta CS = (\Delta E[R^1] \pm \theta_s) + \sum_{q \neq s} \Delta E[R^q] + (\Delta CS - \Delta TrC) = \Delta \tilde{W} - \Delta TrC. \)

Unit: 1 dollar.
Table 9: Counterfactual 1 – Slight Change in Promotion Costs $\theta_s$

<table>
<thead>
<tr>
<th></th>
<th>Transportation Erosion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>After</td>
<td>Before</td>
</tr>
<tr>
<td>$\Delta CS$</td>
<td>1.27</td>
<td>217.72</td>
</tr>
<tr>
<td>$\Delta PS$</td>
<td>-171.91</td>
<td>-171.91</td>
</tr>
<tr>
<td>$\Delta W$</td>
<td>-170.64</td>
<td>45.81</td>
</tr>
<tr>
<td>$\Delta TrC$</td>
<td>216.46</td>
<td></td>
</tr>
<tr>
<td>$\Delta$ Shopping Prob</td>
<td>0.0001</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Promo Freq</th>
<th>Average Price Index</th>
<th>$\Delta PS_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDLP1</td>
<td>25.05</td>
<td>0.8260</td>
</tr>
<tr>
<td>EDLP2</td>
<td>28.14</td>
<td>0.8503</td>
</tr>
<tr>
<td>HiLo</td>
<td>33.03</td>
<td>0.8127</td>
</tr>
<tr>
<td>HT1</td>
<td>21.81</td>
<td>0.9210</td>
</tr>
<tr>
<td>HT2</td>
<td>23.75</td>
<td>0.9052</td>
</tr>
</tbody>
</table>

|                  | Transportation Erosion |                  |
|                  | After                  | Before           |
| $\Delta CS$      | -0.78                  | -130.54          |
| $\Delta PS$      | 100.10                 | 100.10           |
| $\Delta W$       | 99.32                  | -30.44           |
| $\Delta TrC$     | -129.77                |                  |
| $\Delta$ Shopping Prob | -0.0001             |                  |

<table>
<thead>
<tr>
<th>Promo Freq</th>
<th>Average Price Index</th>
<th>$\Delta PS_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDLP1</td>
<td>23.34</td>
<td>0.8838</td>
</tr>
<tr>
<td>EDLP2</td>
<td>26.25</td>
<td>0.9042</td>
</tr>
<tr>
<td>HiLo</td>
<td>28.07</td>
<td>0.8902</td>
</tr>
<tr>
<td>HT1</td>
<td>17.72</td>
<td>0.9608</td>
</tr>
<tr>
<td>HT2</td>
<td>21.47</td>
<td>0.9425</td>
</tr>
</tbody>
</table>
Table 10: Counterfactual – Online Grocery Shopping

<table>
<thead>
<tr>
<th>Counterfactual</th>
<th>Base Case</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>After Erosion</td>
<td>Before</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>3,681.77</td>
<td>1,702.28</td>
<td>2,672.79</td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td>27.15</td>
<td>20.17</td>
<td>20.17</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>3708.92</td>
<td>1,722.45</td>
<td>2,692.97</td>
<td></td>
</tr>
</tbody>
</table>

TrC
Shopping Prob

<table>
<thead>
<tr>
<th>Market Share</th>
<th>Promo Freq</th>
<th>Av. Price Index</th>
<th>Store Choice Prob</th>
<th>PSa</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDLP1</td>
<td>18.39%</td>
<td>27.62</td>
<td>0.7728</td>
<td>0.1543</td>
</tr>
<tr>
<td>EDLP2</td>
<td>21.86%</td>
<td>34.02</td>
<td>0.7937</td>
<td>0.1834</td>
</tr>
<tr>
<td>HiLo</td>
<td>29.10%</td>
<td>33.20</td>
<td>0.8067</td>
<td>0.2442</td>
</tr>
<tr>
<td>HT1</td>
<td>18.72%</td>
<td>23.15</td>
<td>0.8344</td>
<td>0.1571</td>
</tr>
<tr>
<td>HT2</td>
<td>11.94%</td>
<td>24.82</td>
<td>0.8221</td>
<td>0.1002</td>
</tr>
</tbody>
</table>

Unit: 1,000 dollars