

PRICING, COMPETITION, AND WELFARE IN THE  
SUPERMARKET RETAIL INDUSTRY: THEORY AND  
EMPIRICS

By

CIXIU GAO

A dissertation submitted to the  
Graduate School—New Brunswick  
Rutgers, The State University of New Jersey  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy  
Graduate Program in Economics  
written under the direction of  
Rosanne Altshuler  
and approved by

---

---

---

---

New Brunswick, New Jersey

May, 2014

## **ABSTRACT OF THE DISSERTATION**

# **Pricing, Competition, and Welfare in the Supermarket Retail Industry: Theory and Empirics**

**By CIXIU GAO**

**Dissertation Director:**

**Rosanne Altshuler**

The dissertation comprises three essays that investigate market performance and seller behavior in the supermarket retail industry. The first essay empirically examines welfare effects of the informative price advertising in the supermarket retail industry, using structural estimation approaches and individual scanner data. The simulation results numerically show that the private promotion intensities are socially excessive. The welfare implications of price advertising are determined by the two opposite effects of price advertising: (1) the informing and therefore welfare-improving effect, and (2) the welfare-harming effect of higher transportation costs incurred by consumers when promotions are used as a means of business stealing.

In the second essay, I provide an analytical model for the rationale behind supermarket pricing patterns characterized by long-term high prices and temporary price reductions. The model features oligopoly retailers selling a homogeneous storable good that can be consumed for multiple periods, with consumer heterogeneity with respect to search cost, inventory cost, and store loyalty. In the symmetric Markov-perfect equilibrium (MPE) found, retailers randomize prices, and consumer purchase decisions are characterized by a critical price. The Markov transition of states is non-absorbing: the probability of holding

a sale is low at high inventory levels, while at zero inventory retailers compete the hardest. The model is able to generate endogenous temporary price reductions and cyclical inventory variations.

In the third essay, I consider forward-looking purchase and pricing behavior. Consumers maximize the expected discounted future utility flows by balancing inventory cost and potential future savings, and a monopolistic retailer maximizes the present expected profit flows by making a pricing decision that accounts for consumer stockpiling behavior. I estimate the model with data from the laundry detergent market using a simulated minimum distance (SMD) estimator. The simulated market evolution implies that, when consumer inventory level is high and therefore the incentive of purchase is small, the retailer smooths its profit flow by lowering prices to induce purchase; when consumer inventory is low, the retailer expects a high demand driven by urgent consumption needs but tends to keep price high in order to preserve future demand.

## Acknowledgments

I would like to warmly thank my chair advisor Rosanne Altshuler for her great help. She agreed to be my advisor despite of her busy schedule being the Chairwoman of the Department of Economics. She gave generous help that considerably facilitated the final stage of this Ph.D thesis.

I would also like to thank my committee members Hilary Sigman, Douglas Blair and Sanjib Bhuyan. I am especially grateful to Professor Hilary Sigman for being an extraordinary committee member. She gave generously of her time and vast knowledge and experience in empirical economic methodologies. When asked for suggestions she always returned quickly to me with comprehensive and detailed comments. I also thank her for providing precious suggestions on academic job hunting and career development, and sharing her professional experience as an outstanding economist. I am so lucky to have such a serious and encouraging advisor.

Professor Douglas Blair has been extremely supportive throughout my whole dissertation period. Despite of his tight schedule he managed to give insightful comments on the papers. After each discussion with him, I felt more confident and clearer about how to improve my work.

Professor Sanjib Bhuyan, from the Department of Agricultural, Food & Resource Economics, gave many helpful suggestions, especially on the empirical analysis. He was always positive, trying his best to help solve both academic and personal problems of mine. I am also grateful for his suggestions on industrial job hunting and career development.

Last but not least, I would like to thank Tomas Sjöström who had been a great advisor and source of insights. His challenging questions prompted me to think deeper; his insightful ideas inspired and impelled me to complete my work. I am honoured to have him as my mentor and friend.

# Table of Contents

<b>Abstract</b> . . . . .	ii
<b>Acknowledgments</b> . . . . .	iv
<b>List of Tables</b> . . . . .	vii
<b>List of Figures</b> . . . . .	viii
<b>1. Introduction</b> . . . . .	1
<b>2. The Welfare Effects of Price Advertising with Basket Shopping: Structural Estimates from Supermarket Promotions</b> . . . . .	7
2.1. Introduction . . . . .	7
2.2. The Supermarket Retail Industry . . . . .	10
2.3. Model . . . . .	11
2.4. Data . . . . .	18
2.5. Estimation . . . . .	20
2.6. Results . . . . .	27
2.7. Welfare . . . . .	32
2.8. Conclusion . . . . .	41
<b>3. Dynamic Dispersion of Storable Good Prices</b> . . . . .	52
3.1. Introduction . . . . .	52
3.2. Literature . . . . .	54
3.3. Model . . . . .	56
3.4. Equilibrium . . . . .	64

3.5. Conclusion . . . . .	66
<b>4. An Empirical Analysis of Dynamic Supply and Demand of Storable Goods</b>	<b>68</b>
4.1. Introduction . . . . .	68
4.2. Literature Review . . . . .	70
4.3. Model . . . . .	72
4.4. Computation . . . . .	79
4.5. Empirical Application . . . . .	81
4.6. Results . . . . .	89
4.7. Conclusions . . . . .	93
<b>Appendix A.</b> . . . . .	<b>103</b>
<b>Appendix B.</b> . . . . .	<b>104</b>
<b>Bibliography</b> . . . . .	<b>105</b>

## List of Tables

2.1. Summary Statistics . . . . .	44
2.2. Product Preference . . . . .	45
2.3. Effects of Promoted Price Cuts – Percentage Increase* in Product Choice Prob.( $\rho$ ) . . . . .	46
2.4. Store Preference . . . . .	47
2.5. Effects of Promoted Price Cuts – Percentage Changes* in Store Choice Prob.( $\eta$ )	48
2.6. Estimated Bounds of Promotion Cost ( $\theta$ ) . . . . .	48
2.7. Current Equilibrium (Base Case) . . . . .	49
2.8. Market Efficiency . . . . .	49
2.9. Counterfactual – Overall Changes in Promotion Costs $\theta$ . . . . .	50
2.10. Counterfactual – Online Grocery Shopping . . . . .	51
4.1. Preliminary Analysis . . . . .	95
4.2. Store Value Function $W(D_t)$ . . . . .	96
4.3. Estimated Parameters Of The Base Model . . . . .	96
4.4. Simulated vs. Actual Moments I . . . . .	96
4.5. Simulated vs. Actual Moments II . . . . .	97

## List of Figures

2.1. Locations of Stores . . . . .	43
4.1. Simulated Histories Of Purchase and Consumption . . . . .	98
4.2. Simulated Price of a Randomly Selected Product . . . . .	98
4.3. Correlation Between Number Of Price-cuts and Quantity Sold . . . . .	99
4.4. Consumption and Purchase Policies . . . . .	99
4.5. Store Pricing Policy . . . . .	100
4.6. Consumer Value Function (The Number of Inventory Distributions = 2) . .	100
4.7. Consumer Value Function (The Number of Inventory Distributions = 7) . .	101
4.8. Simulated Market Evolution . . . . .	101
4.9. Patterns of Prices and Consumer Inventory . . . . .	102

# Chapter 1

## Introduction

This dissertation has focused on pricing strategy and consumer demand in the supermarket retail industry. The motivation is to understand consumer shopping behavior, the rigorous competition among retailers as they deploy an arsenal of marketing instruments, and how these instruments affect market performance. My research includes theoretical modeling, empirical estimations, and simulations. I have emphasized the use of structural estimation and other modern econometric approaches, and complex computational techniques. In this dissertation, I investigate three interesting questions of the supermarket retail market: (1) how does price advertising of competing supermarket retailers affect grocery shoppers' store choice and product choices and market efficiency; (2) when goods are storable, why supermarket retailers occasionally offer products at discounted prices, and how and why such price dispersion can exist as an equilibrium phenomenon; (3) how are consumer purchase and retailer pricing decisions strategically made for storable goods.

### **1. The Welfare Effects of Price Advertising with Basket Shopping: Structural Estimates from Supermarket Promotions**

This essay empirically examines welfare effects of the informative price advertising in the supermarket retail industry, using structural estimation approaches and individual scanner data.

Competing supermarkets use promotions (advertised temporary price cuts) to announce sales, informing potential customers about price offers via a variety of media forms. Such advertising affects consumer demand because it makes consumers aware of attractive price offers at a specific store location. During 2007-2010, the U.S. supermarkets spent about \$800 million on price advertising per year. Given the magnitude of dollars spent on promotions,

it is interesting to examine whether these promotions are socially efficient.

Theories of informative price advertising do not provide unambiguous predictions of welfare effects. Two effects of price advertising, the demand-creating effect and the business-stealing effect, have been recognized since Marshall (1919). On the one hand, because the firm cannot appropriate 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. Market efficiency with price advertising therefore depends on which of the two effects dominates.

In the supermarket retail industry, however, these two effects are complicated by shopping transportation cost. Since transportation erodes consumer surplus that would have been gained from purchase bundles, it dampens the demand-creating effect. Moreover, stealing one customer from a rival store means a longer shopping trip, if the shopper lives close to the rival store. In a competition-intensified market where shoppers are more likely to travel long distances, transportation costs would cause a worse surplus erosion. In this sense, the effect of price advertising (promotion) is two-fold: it improves market efficiency by reducing price uncertainties and expands quantity (the welfare-improving effect); but it creates inefficiency due to the higher transportation cost (the welfare-harming effect). If the latter effect is sufficiently large, the loss due to longer shopping distances would outweigh the surplus gain from quantity expansion, and price advertising may no longer be socially excessive.

I examine market performance of a supermarket oligopoly by comparing social welfare in the current equilibrium with its counterparts following small deviations in promotion intensity. If an extra unit of promotion improves (harms) social well-being, then the private promotion intensity is socially inadequate (excessive). To investigate this, I construct a model that accounts for both consumer shopping behavior and retail merchandising behavior, and structurally estimate demand and promotion cost parameters. In a Bayesian-Nash 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 estimate consumer preference following the discrete choice literature; the retailers' marginal costs of promotion are structurally estimated using the moment inequality approach (Pakes et al. 2011; Pakes 2010). These structural estimates will allow me to simulate equilibrium and counterfactual outcomes. In particular, the moment inequality approach allows me to circumvent dimensionality issue caused by the large number of products (identified by SKU, stock keeping unit) of a multi-product retailer. The estimation procedure 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 alternative decisions. The difference between the actual and the counterfactual profits provides the bounds for promotion costs.

The simulation results numerically show that, in the current equilibrium (the baseline case), the private promotion intensities are socially excessive; about 20 percent of consumer surplus that would have been gained from purchase bundles has been eroded by transportation. To further examine the effects of price advertising and transportation cost on welfare, I do counterfactual experiments where the promotion costs slightly deviate from their estimates. I find that the welfare implications of price advertising deviates from the usual conclusion that competition and information always improve welfare. The reason closely relates to the welfare-harming effect of price advertising. When promotion costs slightly decrease, for example, stores promote more and price lower. As a result of intensified competition, consumers shop at further-away stores with higher probabilities. However, since the transportation cost increases more than the gain in social surplus due to quantity expansion as numerically showed, social welfare is harmed even competition has been intensified. The reverse is found when promotion costs slightly increases.

Recognizing that transportation cost creates so large inefficiency that offsets the efficiency gain from price advertising, I simulate market outcome where this cost is artificially

removed. This experiment actually simulates the online shopping regime, in which, by assumption, store choice does not depend on home-store distance and shipping of the grocery bundle is free. The simulation shows that the new regime can improve social welfare by 31 percent. In theory, it is clear that the new regime avoids erosion and therefore improves consumer surplus; it is also straightforward that competition among stores will be intensified due to the removal of stores' local market power. In contrast, the effect on social welfare is theoretically vague: the stores could compete more aggressively by spending greater in price advertising. However, my numerical results show that on-line shopping regime only slightly increases advertising while considerably improves consumer surplus: in the new equilibrium stores promote more, price lower, making slightly small profits; compared to the base case, the increase in consumer surplus due to zero transportation costs more than offsets the decrease in producer surplus.

## **Chapter 2. Dynamic Price Dispersion of Durable Goods**

In the second essay, I provide an analytical model that seeks to explain the rationale behind supermarket pricing patterns characterized by long-term high prices and temporary price reductions.

Despite the ubiquitous nature of price promotions, there is little common ground among economists as to why supermarket retailers occasionally offer products at discounted prices, or even how and why such price dispersion can exist as an equilibrium phenomenon. A vast literature aims to generate price distributions, which characterize equilibrium, that are similar to empirically observations. Two classes of models have been constructed. Both examine the pricing decision of single product retailers, and show how consumer heterogeneity can lead to retail price variation over time. The first class assumes consumers differ in their knowledge. Since sellers face a tradeoff between selling to only non-searchers at high price and selling to both searcher and non-searchers at the lowest price among all sellers, the symmetric mixed-strategy equilibrium features a continuous distribution of price. The second class views sales as means of price discrimination. Consumers differ in their reservation prices, willingness to wait for sales, and/or inventory costs (analytically equivalent to willingness to wait).

To better understand retailers' strategic pricing behavior given consumer heterogeneities in store loyalty, willingness to wait or inventory cost, and knowledge in prices, I construct a model of an oligopoly retailers selling a homogeneous storable good, based on the understanding that temporary price reductions serve the role of price discrimination between consumers with different search costs, store loyalty, and willingness to wait (or equivalently, consumer inventory cost).

In this paper a symmetric Markov-perfect equilibrium is found. As in the classic search models, the competing stores face a trade-off between selling only to its own *loyals* at the regular price and to both *loyals* and *shoppers* at some sale price. Retailers randomize prices, and the *cdf* of the equilibrium price distributions have a mass point at the regular price. The equilibrium price distribution is a function of the shoppers' inventory. The mixed strategy equilibrium is characterized by a critical price depend upon which purchase decision is made in each period. The realized price evolution consists of several consecutive regular-price periods, where no sales are offered, and occasionally one-time price reductions. The endogenous price evolution exhibits non-absorbing Markov transition of states: when shoppers hold high inventory, the probability of holding a sale is low, which means inventory will more likely to drop down.

### **Chapter 3. An Empirical Analysis on Dynamic Supply and Demand of Storable Goods**

In this essay, I consider forward-looking purchase and pricing behavior. When goods are storable, consumers take future prices into consideration while making current purchase decision, and suppliers take into account future profit when making current pricing decision. A consumer has an incentive to make an unplanned purchase if she observes a price cut and believes that the price will return to the regular price in the near future. Since consumers would stock up, the unconsumed goods will be stored at an cost. Thus, consumers face a trade-off between storage costs and the attractive low price. On the other hand, the seller's pricing decision also faces a dynamic trade-off: if she offers a sale price, selling more today reduces demand tomorrow. Moreover, If adjusting prices incurs menu costs, it is optimal to cut prices only if the costs can be covered by the increase in total expected profit brought by

expanded sales. Similarly, changing prices from sales prices back to regular levels is optimal only if this menu cost can be covered by the gain in profits attributes to preserved future demand.

To investigate the rationale of such behavior, I propose a model to investigate the consumer stockpiling behavior and the dynamic pricing behavior of a monopolistic retail store. Consumers maximize the expected discounted future utility flows by balancing inventory cost and potential future savings, and the monopolistic retailer maximizes the present expected profit flows by making a pricing decision that accounts for consumer stockpiling behavior.

I estimate the model with data from the laundry detergent market using a simulated minimum distance estimator. In particular, the structural parameters that shape the dynamic behavior of agents, consumer inventory cost and price relabeling cost, are estimated. These structural estimates will allow me to simulate market evolution in order to draw patterns of the dynamic pricing and purchase behavior. I find that, when consumer inventory level is high and therefore the incentive of purchase is small, the retailer smooths its profit flow by lowering prices to induce purchase; when consumer inventory is low, the retailer expects a high demand driven by urgent consumption needs but tends to keep price high in order to preserve future demand.

## Chapter 2

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

### 2.1 Introduction

This paper empirically examines the effect of costly information on market outcome in the supermarket retail industry. Competing supermarkets use promotions (advertised temporary price cuts) to announce sales, informing potential customers about price offers via a variety of media forms. Such advertising affects consumer demand because it makes consumers aware of attractive price offers at a specific store location. During 2007-2010, the U.S. supermarkets spent about \$800 million on price advertising per year.<sup>1</sup> Given the magnitude of dollars spent on promotions, it is interesting to examine whether these promotions are socially efficient.

Theories of informative price advertising do not provide unambiguous predictions of welfare effects.<sup>2</sup> Two effects of price advertising, the demand-creating effect and the business-stealing effect, have been recognized since Marshall (1919). On the one hand, because the firm cannot appropriate 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. Market efficiency

---

<sup>1</sup>Data source: Kantar Media and <http://online.wsj.com>. See also Bolton et al. (2010) and Levy et al. (1997).

<sup>2</sup>For example, Butters (1977) and Roy (2000) predict that equilibrium advertising is socially optimal; in models by Stegeman (1991) and Stahl and Dale (1994) private advertising is socially inadequate; Grossman and Shapiro (1984) argue that it could be either socially inadequate or excessive.

with price advertising therefore depends on which of the two effects dominates.

In the supermarket retail industry, however, these two effects are complicated by shopping transportation cost. Since transportation erodes consumer surplus that would have been gained from purchase bundles, it dampens the demand-creating effect. Moreover, stealing one customer from a rival store means a longer shopping trip, if the shopper lives close to the rival store. In a competition-intensified market where shoppers are more likely to travel long distances, transportation costs would cause a worse surplus erosion. In this sense, the effect of price advertising (promotion) is two-fold: it improves market efficiency by reducing price uncertainties and expands quantity (the welfare-improving effect); but it creates inefficiency due to the higher transportation cost (the welfare-harming effect). If the latter effect is sufficiently large, the loss due to longer shopping distances would outweigh the surplus gain from quantity expansion, and price advertising may no longer be socially excessive.

I examine market performance of a supermarket oligopoly by comparing social welfare in the current equilibrium with its counterparts following small deviations in promotion intensity. If an extra unit of promotion improves (harms) social well-being, then the private promotion intensity is socially inadequate (excessive). To investigate this, I construct a model that accounts for both consumer shopping behavior and retail merchandising behavior, and structurally estimate demand and promotion cost parameters. In a Bayesian-Nash 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,<sup>3</sup> facing the tradeoff between attracting extra store visits and paying additional promotion costs.

Using scanner data of consumer shopping and store merchandising information, I estimate consumer preference following the discrete choice literature; the retailers' marginal costs of promotion are structurally estimated using the moment inequality approach (Pakes et al. 2011; Pakes 2010). These structural estimates will allow me to simulate equilibrium and counterfactual outcomes. In particular, the moment inequality approach allows me to

---

<sup>3</sup>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).

circumvent dimensionality issue caused by the large number of products (identified by SKU, stock keeping unit) of a multi-product retailer. The estimation procedure 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 alternative decisions. The difference between the actual and the counterfactual profits provides the bounds for promotion costs.

My estimation and simulation encounter three difficulties: (1) the wholesale prices that will be used to recover counterfactual profits are not observed; (2) the search for the optimal price vector in a large dimensional space using regular methods is extremely inefficient; and (3) the search for the optimal promotion decision in a large dimensional discrete space is practically impossible. I estimate wholesale prices using the firm's first-order condition at the observed pricing decisions. The second problem is solved using techniques of principal component analysis and factor analysis. The third difficulty is alleviated using a new algorithm that largely reduces computational complexity.

The simulation results numerically show that, in the current equilibrium (the baseline case), the private promotion intensities are socially excessive; about 20 percent of consumer surplus that would have been gained from purchase bundles has been eroded by transportation. To further examine the effects of price advertising and transportation cost on welfare, I do counterfactual experiments where the promotion costs slightly deviate from their estimates. I find that the welfare implications of price advertising deviates from the usual conclusion that competition and information always improve welfare. The reason closely relates to the welfare-harming effect of price advertising. When promotion costs slightly decrease, for example, stores promote more and price lower. As a result of intensified competition, consumers shop at further-away stores with higher probabilities. However, since the transportation cost increases more than the gain in social surplus due to quantity expansion as numerically showed, social welfare is harmed even competition has been intensified. The reverse is found when promotion costs slightly increases.

Recognizing that transportation cost creates so large inefficiency that offsets the efficiency gain from price advertising, I simulate market outcome where this cost is artificially

removed. This experiment actually simulates the online shopping regime, in which, by assumption, store choice does not depend on home-store distance and shipping of the grocery bundle is free. The simulation shows that the new regime can improve social welfare by 31 percent. In theory, it is clear that the new regime avoids erosion and therefore improves consumer surplus; it is also straightforward that competition among stores will be intensified due to the removal of stores' local market power. In contrast, the effect on social welfare is theoretically vague: the stores could compete more aggressively by spending greater in price advertising. However, my numerical results show that online shopping regime only slightly increases advertising while considerably improves consumer surplus: in the new equilibrium stores promote more, price lower, making slightly small profits; compared to the base case, the increase in consumer surplus due to zero transportation costs more than offsets the decrease in producer surplus.

## 2.2 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 accounts for 5.7% of disposable income in 2011. This industry has experienced considerable expansion over the years. Between 1997 to 2012, the average annual sales per store rose from \$2.5 million to \$16.3 million; the number of items stocked increased from 18,000 to more than 42,000 <sup>4</sup>. The rise of chainstores, the development of supermarkets, the introduction of food discounters, and the continual growth in the variety of products have affected the organization and

---

<sup>4</sup>Food Marketing Institute. <http://www.fmi.org/research-resources/supermarket-facts>

competitive behavior of grocery retailing.

Competition among supermarket retailers is fierce – the net profit margin (after tax) in 2012 is only 1.5%. Competing on the razor-thin profit margins, retailers use an arsenal of marketing instruments and sophisticated pricing strategies to attract customers and avoid being squeezed out of the market. Grocery retailers price each food 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. 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 where to shop: factors such as product assortment, geographical location of store, shopping experience and customer service, are also crucial.

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 reported that 54% of consumers shop for at least some categories in this emerging grocery channel.<sup>5</sup> Traditional grocers react to the threat by building their own online shopping sites, sometimes coordinating with e-grocers such as Peapods, and offering full-assortment products, downloadable price and promotion information, and home delivery services. It is believed that in the foreseeable future grocery retailers will have to respond, whether by bolstering in-store experience to defend their business or by building multichannel retailing capability that integrates in-store, online, and digital mobile offerings to meet shoppers' differing needs.

## 2.3 Model

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

---

<sup>5</sup>Four Forces Shaping Competition in Grocery Retailing, industry report, Booz & Company.

equilibrium shoppers choose the store that offers the greatest utility of shopping 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 how much they are priced) in order to compete over sales. The use of an estimated shopping utility function and an pricing condition will allow for counterfactual experiments in which shoppers reallocate themselves across stores and new promotional and pricing equilibria are computed.

### 2.3.1 Shopping Behavior

The model assumes that prior to a shopping trip, a shopper  $h$  receives promotion information from zero, one or more stores. Based on available price information, the shopper constructs an expected merchandising utility of each store. Along with that, the shopper also takes into account store valuation and transportation cost of the shopping trip, and chooses a single store to shop.<sup>6</sup> Once in the store when all prices are realized, for each product category the shopper chooses the optimal product that maximizes category utility. The model therefore 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. I first specify the in-store shopping behavior where shoppers make product choice within each category conditional on store choice, then describe store choice decision making.

#### Within-category Product Choice

Let  $J_c$  denote the set of product alternatives of category  $c$ . Once in the store and observe prices and all merchandising activities, for each category a shopper  $h$  chooses a product to maximize category utility. The product choices are independently made across categories.

At the time of purchase, the indirect utility that shopper  $h$  obtains from product  $j_c$  of

---

<sup>6</sup>The model does not take into account the "cherry-picker" behavior that a shopper choose multiple stores in a 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. See Fox and Hoch (2005), Gauri et al. (2008a), Smith and Thomassen (2012), and Talukdar et al. (2008).

category  $c$  in time  $t$  at store  $s$  takes the form

$$w_{hst,j_c} = \chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c} + \epsilon_{hst,j_c}, \quad (2.1)$$

where  $p_{st,j_c}$  is the price;  $m_{st,j_c}$  is the promotion dummy;  $n_{st,j_c}$  is the dummy of in-store display;<sup>7</sup>  $y_{j_c}$  contains dummies of brand and package size;  $\chi_c$  is the intrinsic utility of category  $c$  invariant over products within the category;  $\chi_c$ ,  $\alpha_c$ ,  $\beta_{c,1}$ ,  $\beta_{c,2}$ , and  $\gamma_c$  are parameters to be estimated;  $\epsilon_{hst,j_c}$  is an idiosyncratic shock assumed to follow type I extreme value distribution, i.i.d. across products, categories, stores, shoppers, and periods. Finally, the deterministic utility of the outside option, no purchase, is normalized to zero, thus  $w_{hst,0_c} = \epsilon_{hst,0_c}$ . The probability of choosing a particular product,  $\rho_{st,j_c}$ , is the probability that  $j_c \in \arg \max w_{hst,j_c}, j_c \in J_c$ . Following McFadden (1974), it is given by

$$\rho_{st,j_c} = \frac{\exp(\chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c})}{1 + \sum_{k_c \in J_c} \exp(\chi_c + \alpha_c p_{st,k_c} + \beta_{c,1} m_{st,k_c} + \beta_{c,2} n_{st,k_c} + \gamma_c y_{k_c})}. \quad (2.2)$$

Once the optimal product is chosen, the expected maximum utility of the category (hereafter abbreviated to category utility) is

$$v_{sct} = \log\left(1 + \sum_{j_c \in J_c} \exp(\chi_c + \alpha_c p_{st,j_c} + \beta_{c,1} m_{st,j_c} + \beta_{c,2} n_{st,j_c} + \gamma_c y_{j_c})\right). \quad (2.3)$$

There are three reasons to include brand and size dummies,  $y_{j_c}$ . First, it improves model fit. Second, the inclusion of brand and size dummies (their combination is sufficient to distinguish products from 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. Third, the brand-size combination captures unobserved product characteristics (e.g, quality). Therefore, the correlation between price and unobserved characteristics are accounted for and does not need instruments. 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. Another potential correlation between price and unobserved characteristics may result from

---

<sup>7</sup>In-store display, as a kind of merchandising activity, is included in estimating consumer preferences, while is not treated as a choice variable in firm's problem. The merchandising activities though out this paper refer to pricing and promotions only.

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).

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

$$u_{st}(\mathbf{x}_{st}) = \sum_{c \in C} v_{sct}, \quad (2.4)$$

where  $C$  denotes the set of product categories. The simple additive format of  $u_{st}$  makes the inclusion of category intrinsic utility clear:  $\chi_c$  accounts for different "weights" of categories in store choice decision making. A promoted price of an item from a category with higher intrinsic utility is more powerful in attracting customer.

### Store Choice

Prior to a shopping trip, a shopper evaluates the expected merchandising utility by forming an expected optimal purchase bundle at  $s$ , comprised of the optimal product of each category. The expected merchandising utility and the optimal bundle depends on the shopper's price knowledge. It is assumed that a shopper passively receives promotion ads from stores, and that the probability of receiving an ad from a specific store is independent across stores. Let  $\phi_s$  denote the time-invariant probability of receiving promotion ads from store  $s \in \{1, \dots, S\}$ . Let a dummy vector  $ad_{ht} = (ad_{h1t}, \dots, ad_{hst}, \dots, ad_{hSt})'$  denote the status of ad exposure of shopper  $h$  in  $t$ ,  $prob(ad_{hst} = 1) = \phi_s$ . The shopper has some prior distribution of merchandising decision,  $F_s(\mathbf{x}_s)$ . If she didn't receive promotion information prior to shopping from store  $s$  (uninformed,  $ad_{hst} = 0$ ), she maintains the prior price information and forms expectation on product choices according to  $F_s(\mathbf{x}_s)$ . If she received promotion information from  $s$  (informed,  $ad_{hst} = 1$ ), then she updates price knowledge about promoted products and perceives that the prices of the un-promoted items follow a distribution conditional on promotion information,  $F_{st}(\mathbf{x}_{st} | \mathbf{x}_{st}^{prom})$ , where the superscript *prom* denotes promoted items. The expected merchandising utilities at  $s$ ,  $\bar{u}_{st}(ad_{hst})$ , perceived

by uninformed and informed shoppers, respectively, are given by

$$\begin{aligned}\bar{u}_{hst}(ad_{hst} = 0) &= \int u_s(\mathbf{x}_s) dF_s(\mathbf{x}_s) \equiv \bar{u}_s, \\ \bar{u}_{hst}(ad_{hst} = 1) &= \int u_{st}(\mathbf{x}_{st}) dF_{st}(\mathbf{x}_{st} | \mathbf{x}_{st}^{prom}).\end{aligned}\tag{2.5}$$

For an uninformed shopper, as  $\mathbf{x}_s$  is integrated out, the expected merchandising utility is time-invariant; while for an informed shopper, the expected merchandising utility depends on information about promoted items because uncertainties about un-promoted items are integrated out.

Based on available price information, a shopper chooses a store to maximize shopping utility. Conditional on store characteristics, the indirect utility function of shopper  $h$  at store  $s$  in time  $t$  takes the form:

$$U_{hst}(ad_{hst}) = \lambda_s + \iota \bar{u}_{hst}(ad_{hst}) + \kappa dist_{hs} + \zeta_{hst},\tag{2.6}$$

where  $\lambda_s$  is the average store valuation that accounts for factors such as services and shopping environment;  $Eu_{hst}$  is the expected merchandising attractiveness at  $s$  that depends on ad exposure  $ad_{hst}$ ;  $dist_{hs}$  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 assumed type I extreme value distributed, i.i.d. across individuals, stores, and time;  $\iota$  and  $\kappa$  are parameters associated with expected store merchandising attractiveness and shopping distance, respectively. The deterministic utility of the outside option, no shopping, is normalized to zero. Following the discrete choice literature, shopper  $h$  will visit store  $s$  with probability

$$\eta_{hst}(\mathbf{x}_{st}, \mathbf{x}_{-st}, ad_{ht}, dist_{hs}) = \frac{\exp(\lambda_s + \iota \bar{u}_{hst}(ad_{hst}) + \kappa dist_{hs})}{1 + \sum_{q \in \{1, \dots, S\}} \exp(\lambda_q + \iota \bar{u}_{hqt}(ad_{hqt}) + \kappa dist_{hq})}.\tag{2.7}$$

Let  $\rho_{st}$  be the vector of product choice probabilities, and  $mc_{st}$  be the vector of wholesale prices, stacked across all products. 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

$$\bar{\eta}_{st}(\mathbf{x}_{st}, \mathbf{x}_{-st}) = \int \int \eta_{hst}(\mathbf{x}_{st}, \mathbf{x}_{-st}, ad_{ht}, dist_{hs}) d\Omega(ad_{ht}) dD(dist_{hs}).\tag{2.8}$$

Let  $MS$  be the market size. The sales revenue, that is, the revenue with wholesale costs

subtracted but excluding promotion costs and fixed cost, is the following:

$$R_{st}(\mathbf{x}_{st}, \mathbf{x}_{-st}) = MS \times \bar{\eta}_{st}(\mathbf{x}_{st}, \mathbf{x}_{-st}) \times \rho_{st}(\mathbf{x}_{st})'(\mathbf{p}_{st} - mc_{st}). \quad (2.9)$$

### 2.3.2 Store Behavior

Retail stores simultaneously make pricing and promotion decisions for all products carried to maximize the expected store-level profit given their expectations on rivals' decisions. The information set at the time of decision is denoted  $H_s$ , where  $H_s \in \mathcal{H}_s$ . The strategy played by  $s$  is a mapping  $\mathbf{x}_s = \sigma_s : \mathcal{H}_s \rightarrow \mathcal{X}_s$  where  $\mathcal{X}_s$  is the action set of  $s$ . The information set of the store agent include the time-varying wholesale prices and available products on shelf (items may be out of stock). As noted above, the action of  $s$  can be partitioned into a pricing decision and a promotion decision,  $\mathbf{x}_s = (\mathbf{p}'_s, \mathbf{m}'_s)'$ . Since this paper is focused on promotion activity, for practical reasons the decisions made by store agent include pricing and promotion only. My dataset also contains in-store display, another kind of merchandising decision made by store manager. Realizing that this activity will have an impact on demand, I assume that it is taken as given by the store agent. Thus the in-store display decision is also in the information set  $H_s$ .

I impose a restriction on promoted price for the coordination of pricing and promotion decisions, in the sense that only discounted prices are promoted. Suppose that for each item  $j_c$  there is an interval from which its price will be chosen,  $[\underline{p}_{s,j_c}, \bar{p}_{s,j_c}]$ , and an interval from which a promoted price will be chosen,  $[\underline{p}_{s,j_c}, \hat{p}_{s,j_c}]$ . The coordination of pricing and promotion implies  $\hat{p}_{s,j_c} < \bar{p}_{s,j_c}$ . This restriction is grounded on the consistent belief about prices given by the real shopping experience that promoted prices are expected to be discounted.

For the ease of notation I drop the subscript for time  $t$ . Let  $\pi_s(\mathbf{x}_s, \mathbf{x}_{-s})$  be the profit of  $s$ . The firm's problem is to maximize the expected profit  $E[\pi_s(\mathbf{x}_s, \mathbf{x}_{-s})|H_s]$ , by making pricing and promotion decisions subject to the coordination between them, given its belief

on rivals' actions. Formally,

$$\mathbf{x}_s \in \arg \max E [\pi_s(\mathbf{x}_s, \mathbf{x}_{-s}) | H_s]$$

subject to

$$m_{s,j_c} = 0 \text{ or } 1, \quad (2.10)$$

$$\underline{p}_{s,j_c} \leq p_{s,j_c} \leq \bar{p}_{s,j_c}, \text{ if } m_{s,j_c} = 0,$$

$$\underline{p}_{s,j_c} \leq p_{s,j_c} \leq \hat{p}_{s,j_c}, \text{ if } m_{s,j_c} = 1, \forall j_c \in J_c, c \in C.$$

I assume a unit promotional cost  $\theta_s$  will incur for each promoted product. The expected profit is the expected revenue minus the total wholesale costs, promotion costs, and fixed cost:

$$E [\pi_s(\mathbf{x}_s, \mathbf{x}_{-s}) | H_s] = E [R_s(\mathbf{x}_s, \mathbf{x}_{-s}) | H_s] - \theta_s \cdot (\mathbf{1}' \mathbf{m}_s) - FC_s, \quad (2.11)$$

where  $E [R_s(\mathbf{x}_s, \mathbf{x}_{-s}) | H_s]$  is the expected revenue with wholesale costs subtracted;  $R_s(\mathbf{x}_s, \mathbf{x}_{-s})$  is given by (2.9);  $\mathbf{1}' \mathbf{m}_s$  represents the total number of promotions; and  $FC_s$  is fixed cost.

There can be prediction error due to randomness in observed profits that is not known at the time decisions are made. For example, store  $s$ 's expectation on  $\mathbf{x}_{-s}$  would differ from the outcome. The expectation error is denoted by  $err_s = R_s - E[R_s | H_s]$ , and is mean zero conditional on the information set by construction, i.e.,  $E[err_s | H_s] = 0$ . Since agent's strategy  $\mathbf{x}_s = \sigma_s(H_s)$  is a function of  $H_s$  but  $err_s \notin H_s$ ,  $e_s$  is mean independent of  $\mathbf{x}_s$ . This means that agents are generally right about their decisions.

The firm's problem can be decomposed into a problem of discrete promotion decision, and a sub-problem of pricing conditional on promotion decision. In the sub-problem, assume the existence of an interior solution,  $\mathbf{p}_s^*(\mathbf{m}_s)$ . From (2.9) and (2.11), the first-order condition with respect to  $\mathbf{p}_s$  is

$$\begin{aligned} \frac{\partial}{\partial \mathbf{p}_s} E[\pi_s(\mathbf{p}_s, \mathbf{m}_s, \mathbf{x}_{-s}) | H_s] &= \frac{\partial}{\partial \mathbf{p}_s} E[R_s(\mathbf{p}_s, \mathbf{m}_s, \mathbf{x}_{-s}) | H_s] \\ &= 0 \\ &= E \left[ \frac{\partial \bar{\eta}_s}{\partial \mathbf{p}_s} \cdot \rho'_s(\mathbf{p}_s - mc_s) + \bar{\eta}_s \cdot \left[ \frac{\partial \rho_s}{\partial \mathbf{p}_s} \right] (\mathbf{p}_s - mc_s) + \bar{\eta}_s \cdot \rho_s \middle| H_s \right] \bigg|_{\mathbf{p}_s = \mathbf{p}_s^*(\mathbf{m}_s)}. \end{aligned} \quad (2.12)$$

Thus the original problem becomes a promotion decision making problem with an implicit price variable satisfying (2.12). A necessary equilibrium condition is that the strategy played

by the agent is at least as good as any alternative. That is, the optimal choice of  $\mathbf{m}_s$  satisfies (omitting the implicit price variable)

$$E[\pi_s(\mathbf{m}_s, \mathbf{x}_{-s})|H_s] \geq E[\pi_s(\mathbf{m}'_s, \mathbf{x}_{-s})|H_s], \quad (2.13)$$

where  $\mathbf{m}'_s \neq \mathbf{m}_s$ . From equation (2.9), this implies the following condition:

$$\begin{aligned} E[\Delta R_s(\mathbf{m}_s, \mathbf{m}'_s, \mathbf{x}_{-s})|H_s] &\equiv E[R_s(\mathbf{m}_s, \mathbf{x}_{-s})|H_s] - E[R_s(\mathbf{m}'_s, \mathbf{x}_{-s})|H_s] \\ &\geq \theta_s \cdot \mathbf{1}'(\mathbf{m}_s - \mathbf{m}'_s). \end{aligned} \quad (2.14)$$

The unit promotion cost,  $\theta_s$ , can be estimated by computing the difference between the current expected revenue and counterfactual expected revenues generated by alternative promotion decisions. To recover counterfactual expected revenue  $E[R_s(\mathbf{m}'_s, \cdot)|H_s]$ , a price vector associated with the alternative promotion decision will be re-optimized according to the first-order condition.

## 2.4 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 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 of whether the household shops, which store is visited if shop, 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 contains proxy measures for the distance to each store for each of the 548 households, using the households' and stores' five-digit zip codes. Since it is difficult to isolate the market of the five competing stores in the extent of geographical area or customer identity, I approximate market size,  $MS$ , by comparing the total quantities sold by the stores to the quantities purchased by the tracked households. 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.

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 2.1 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 2.1. Market share 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 average price index, computed as the ratio between period- $t$  price of a product and its regular price, weighted by market share. A deep price refers to a price reduction at least 15% lower than regular price. The statistics are consistent with stores' price positioning: EDLP stores have lower prices and HiLo store offers more (deep) price cuts and promotions, whereas 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 (6364), 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 tiny market share <sup>8</sup> and items that are not carried by all five stores. These restrictions reduce the number of items within each category from a range of 47 to 729 to a smaller

---

<sup>8</sup>Depending on category, I set the threshold of "tiny" market share from 0.5 to 2 percent, balancing between the efficiency of logit regression and product variety.

range of 16 to 38, and the total number of product from 6364 to 474.

Statistics related to shopping trips are shown in the bottom half of Table 2.1. On average, shoppers visit grocery stores 1.56 time a week, spending 37 dollars per visit. The standard deviation of basket spending is big, indicating that the basket spending is skewed largely to the right. The mean home-store distance is 2.7 miles, while the mean of the actual travel distance is only 1.47 miles, implying the tendency to choose a closer store. 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 removes others, in order to be compatible with logit model. Smaller transactions are treated as unplanned or urgent purchases.

## 2.5 Estimation

The goal of estimation is to find the promotion cost parameter,  $\theta_s, s \in \{1, \dots, S\}$ . This requires parameter estimates of product preference, store preference, and the wholesale prices that will be used to recover counterfactual profits. My Estimation of the behavioral model will implement three major methodologies. First, the demand system will be estimated using standard logit regressions and simulation method.<sup>9</sup> Second, the wholesale costs are estimated based on the firm's first-order condition with observed merchandising decisions. Third, the promotional cost parameters are estimated using moment inequality method.

### 2.5.1 Demand

In stage one, I estimate parameters related to product choice within each category ( $\Theta_1 = \{\chi_c, \alpha_c, \beta_{1,c}, \beta_{2,c}, \gamma_c, \text{ all } c \in C\}$ ) conditional on observed within-category purchase behavior using logit regressions, category by category.  $\alpha_c, \beta_{1,c}, \beta_{2,c}, \gamma_c$  can be identified from market shares. The parameter of category-intrinsic utility,  $\chi_c$ , is identified from purchase incidence, as a bigger  $\chi_c$  implies a higher probability of choosing any product from the category. Therefore, the outside choice here refers to the behavior that a shopper arrives at a store but makes no purchase from that category.

---

<sup>9</sup>Bell et al. (1998) and I use the same data set to estimate store choice. The primary difference is that their model does not account for price advertising thus consumers know no more than a prior distribution of prices.

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,  $\hat{F}_s(\mathbf{p}_s, \mathbf{m}_s)$ , is approximated by the empirical distribution; the updated price knowledge,  $\hat{F}_{st}(\mathbf{x}_{st} | \mathbf{x}_{st}^{prom})$ , is similar except  $\mathbf{x}_{st}^{prom}$  are equal to the promoted prices in that period. The distribution of ad exposure status  $\Omega(ad)$  remains to be empirically specified. Let  $\mathcal{AD}$  denote the set of all possible exposure statuses. There are  $2^S$  mutually different statuses in the set. Assuming shoppers are independently exposed to ads sent by different stores, the probability of being in status  $ad = (ad_1, \dots, ad_S)'$  is

$$prob(ad) = \prod_s ad_s \cdot \phi_s + (1 - ad_s)(1 - \phi_s). \quad (2.15)$$

The log-likelihood function of store choice is

$$l(\phi, \Theta_1, \Theta_2) = \sum_t \sum_h \sum_{ad \in \mathcal{AD}} prob(ad) \cdot \log \left( \sum_s \eta_{hst}(\mathbf{x}_t, ad, dist_{hs}; \Theta_1, \Theta_2) \cdot store_{hst} \right), \quad (2.16)$$

where  $store_{hst}$  equals 1 if store  $s$  is visited by  $h$  in  $t$ , and 0 otherwise. Given the parameter estimates of within-category choice preference,  $\hat{\Theta}_1$ , the identified  $\phi$  and  $\Theta_2$  are the parameters that jointly maximize the store-choice likelihood:

$$(\phi, \Theta_2) \in \arg \max l(\phi, \hat{\Theta}_1, \Theta_2). \quad (2.17)$$

The store choice likelihood needs to be constructed by integrating over  $F_s$ , as store choice probability depends on price knowledge (see equations (2.5) and (2.7)). Practically, I compute this likelihood function using simulation by randomly drawing prices from  $\hat{F}_s(\mathbf{x}_s)$  or  $\hat{F}_{st}(\mathbf{x}_{st} | \mathbf{x}_{st}^{prom})$ .

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.

### 2.5.2 Supply

#### 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 (Porter 1983; Bresnahan 1987; Nevo 2001), I estimate  $mc_s$  using the first-order condition in the firm's problem conditional on the observed merchandising decisions. Suppose the wholesale cost vector takes the form:

$$mc_{st} = mc_s + \tau_{st}, \quad (2.18)$$

where  $mc_s$  is the vector of mean wholesale cost vector to be estimated, and  $\tau_{st}$  is a vector of unobservable (to the econometrician) disturbances but is accounted for by the store, satisfying  $E[\tau_{st}] = 0$ .<sup>10</sup> Sources of this cost disturbance may include variations in manufacturer's price, delivery cost, and packing cost. The first-order condition in equation (2.12) implies

$$mc_s = \mathbf{p}_{st} + E \left[ \left[ \frac{\partial \bar{\eta}_{st}}{\partial \mathbf{p}_{st}} \cdot \rho_{st} + \bar{\eta}_{st} \cdot \left[ \frac{\partial \rho_{st}}{\partial \mathbf{p}_{st}} \right] \right]^{-1} (\bar{\eta}_{st} \cdot \rho_{st}) \middle| H_{st} \right] - \tau_{st}. \quad (2.19)$$

The mean wholesale cost vector  $mc_s$  is estimated by taking the average of (2.19) using the observed prices and product choice probabilities, and demand estimates. The numerical procedure includes integrating over the distributions of  $ad_{st}$ ,  $dist_{hs}$ , and  $F_{-s}(\mathbf{p}_{-s}, \mathbf{m}_{-s})$ .

#### Market Share

Using the discrete distribution of ad exposure status, the market share in (2.8) becomes

$$\bar{\eta}_{st}(\mathbf{x}_{st}, \mathbf{x}_{-st}, \phi) = \int \sum_{ad \in AD} prob(ad) \eta_{sht}(\mathbf{x}_{st}, \mathbf{x}_{-st}, ad, dist_{hs}) dD(dist_{hs}). \quad (2.20)$$

#### 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  $\mathbf{m}'_{st}$ . I

---

<sup>10</sup>Ishii (2011) uses this form to estimate the mean marginal cost.

use moment inequality method that allows me to circumvent the dimensionality issue and preserve the discrete nature of the variable. My estimation methods draws intensively from Pakes (2010) and Pakes et al. (2011) and are similar to applications including 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 profits from its observed choice are greater than its expected profits from alternative choices. Since in the counterfactual the manager makes different decisions, the counterfactual profits contains different promotional costs. A necessary condition for profit maximization is that each store's expected profit from choosing actual  $\mathbf{p}_s$  and  $\mathbf{m}_s$  is at least as good as its expected profit from 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. Thus, I am able to estimate the cost for each store without much efficiency loss.

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

$$\theta_{st} = \theta_s + \tilde{\theta}_{st}, \quad (2.21)$$

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  $\sum \tilde{\theta}_{st} = 0$ . 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 (2.14) implies that

$$E \left[ \Delta R_{st}(\mathbf{m}_{st}, \mathbf{m}'_{st}, \mathbf{x}_{-st}) | H_{st} \right] \geq (\theta_s + \tilde{\theta}_{st}) \cdot (\mathbf{1}'(\mathbf{m}_{st} - \mathbf{m}'_{st})). \quad (2.22)$$

I consider small deviations from the observed promotions as alternatives, that is, to alter the promotion decision of only one item, so that  $\mathbf{1}'(\mathbf{m}_{st} - \mathbf{m}'_{st}) = \pm 1$ . 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,  $\mathbf{m}'_{st} = \mathbf{m}_{st} - \mathbf{e}_{st,j_c}$  with  $m_{st,j_c} = 1$ , where  $\mathbf{e}_{st,j_c}$  is a vector of zeros of the same

length as  $\mathbf{m}$  except the  $j_c^{th}$  element equals one. Dropping off 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:

$$E \left[ \Delta R_{st}(\mathbf{m}_{st}, \mathbf{m}'_{st}, \mathbf{m}_{-st}) | H_{st} \right] \geq \theta_s + \tilde{\theta}_{st}. \quad (2.23)$$

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

$$E \left[ \Delta R_{st}(\mathbf{m}_{st}, \mathbf{m}''_{st}, \mathbf{m}_{-st}) | H_{st} \right] \geq -(\theta_s + \tilde{\theta}_{st}). \quad (2.24)$$

Suppose in observation  $t$ , the number of products on promotion is  $J_{st,1}$  and the number of products not promoted is  $J_{st,2}$ . The sample analogue of inequalities (2.23) and (2.24) are

$$\theta_s \leq \frac{1}{T} \sum_{t=1}^T \frac{1}{J_{st,1}} \sum_{\mathbf{m}'_s} \Delta R_{st}(\mathbf{m}_{st}, \mathbf{m}'_{st}, \mathbf{m}_{-st}) \equiv UB_s, \quad (2.25)$$

$$\theta_s \geq -\frac{1}{T} \sum_{t=1}^T \frac{1}{J_{st,2}} \sum_{\mathbf{m}''_s} \Delta R_{st}(\mathbf{m}_{st}, \mathbf{m}''_{st}, \mathbf{m}_{-st}) \equiv LB_s.$$

The cost disturbance  $\tilde{\theta}_{st}$  on the right hand side of (2.23) and (2.24) are averaged out: for example, the right hand side of (2.23) becomes

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T \frac{1}{J_{st,1}} \sum_{\mathbf{m}'_s} (\theta_s + \tilde{\theta}_{st}) \\ &= \frac{1}{T} \sum_{t=1}^T \frac{1}{J_{st,1}} J_{st,1} \times (\theta_s + \tilde{\theta}_{st}) \\ &= \frac{1}{T} \sum_{t=1}^T (\theta_s + \tilde{\theta}_{st}) \\ &= \theta_s + \frac{1}{T} \sum_{t=1}^T \tilde{\theta}_s \\ &= \theta_s. \end{aligned}$$

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

### Computational Issues

To recover counterfactual expected revenues and predict market outcome under alternative behaviors, I need to solve for the optimal promotion and pricing decisions. Unfortunately, realizing that the number of items supermarket carries is of thousands, the dimensionality issue as well as the discrete nature of promotion decision make it practically impossible to jointly solve for  $\mathbf{p}$  and  $\mathbf{m}$  using standard algorithms: <sup>11</sup> first, searching for the optimal  $\mathbf{p}$  in the subproblem of firm is itself time-exhausting and inefficient; second, searching for  $\mathbf{m}$  is of complexity  $\#J^2$  if the number of product items is  $\#J$ . I use principal component technique and factor analysis to deal with the first issue, and an "ordered" promotion decision rule for the second problem.

The number of product is greatly reduced using the method discussed in section 2.4. However, jointly solving for 474 prices in the continuous space is a big challenge. I use principal component technique compresses the large dimensional variable into a vector of much smaller space, solve the profit optimization problem in the reduced space, and transform the reduced variable back into the space of prices using factor analysis techniques. The principal component analysis on price variations shows that the first 12 components are able to account for 80 percent of the overall price variations in the data. I project the price vector into the reduced space using linear transformation consists 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

---

<sup>11</sup>Heuristic algorithms such as genetic algorithm are available in solving the mixed-integer optimization problem, but they tend to be time-consuming when the number of integer variables to solve is large.

among the 474 products considered, a considerable amount of them 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 times standard variance higher than the mean frequency of promotion). There are 52 products in this set.

For promotion decision making, I use an algorithm for searching for the optimal  $\mathbf{m}$ , aiming to effectively reduce time consumption. Suppose the number of items considered for promotion is  $\#J$ . Given the observed actions of rivals, choose the optimal one item for promotion such that the greatest profit is generated. Then, conditional on the first promoted item, choose the next optimal item for promotion, and record the increase (decrease) in profit. Iterate this procedure until the increase in profit difference between the  $n^{th}$  and the  $(n + 1)^{th}$  promotion is less than the unit promotion cost (assuming discrete convexity of profit function). The optimization of price vector, using the method described above, is nested in each iteration. The computational complexity is  $\#J$  in the first iteration,  $\#J - 1$  in the second iteration, and so on. Therefore, the algorithm largely reduces computational complexity from  $2^{\#J}$ , if search over all alternatives in the choice set, to at most  $\#J(\#J + 1)/2$ .

Practically, I set the bounds in the firm's problem using empirically observed measures. For un-promoted products, the bounds  $\underline{p}_{s,j_c}$  and  $\bar{p}_{s,j_c}$  are respectively set to be the observed minimum price, and the most frequently observed price (the regular price). The upper bound of a promoted price,  $\hat{p}_{s,j_c}$ , is set to be 90 percent of regular price.

### 2.5.3 A Summary of Estimation Procedures

To better 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,  $\Theta_1 = \{\chi_c, \alpha_c, \beta_{1,c}, \beta_{2,c}, \gamma_c, \text{ all } c \in C\}$ . 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  $R(\mathbf{m}, \cdot)$  and counterfactual revenue  $R(\mathbf{m}', \cdot)$  at the observed firm choices;
5. Estimate promotion cost  $\theta_s$  by finding the difference between  $R(\mathbf{m}, \cdot)$  and  $R(\mathbf{m}', \cdot)$ .

## 2.6 Results

### 2.6.1 Within-category Choice

I estimate demand in order to predict sales and profits generated by alternative pricing decisions. Table 2.2 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, once brands and package sizes are controlled, 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 results 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 attributed to the nature of the two categories: products are much less differentiated and the differences in unobserved characteristics are small.

I use intercept  $\chi_c$  to measure the intrinsic category utility. 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 it's common for all products of the same category. They are identified using both observed purchases and

outside-choice observations - shoppers who enter the store but didn't purchase the category. The estimates of  $\chi_c$  varies largely 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 to simultaneous promotion and a price cut, averaging across products. The price cut takes the value of 15 percent of its regular price (deep price cut). The percentage change of choice probability is computed as follows:

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

Table 2.3 provides the maximum, the mean, and the standard deviation of the percentage change in market share 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 increase in market shares of the promoted item; for Detergents, Hot Dogs, Tissue Papers, and Yogurt, they can cause increase by 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 market share is smaller, promoted price cut is more effective in categories with larger variety (Hot Dogs, Yogurt), while in categories with much fewer number of products and less product differentiations, like in Butter and Sugar, this effect is smaller. Another factor is the value of preference estimates that determine price and promotion sensitivities. Categories with higher  $\alpha_c$  and/or  $\beta_{c,a}$  (Detergent and Tissue Paper) may also have remarkable quantity effect.

### 2.6.2 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 2.5.1, 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,

$(\alpha_c, \beta_{1,c}, \beta_{2,c}, \chi_c)$ . The results are displayed in the first column of Table 2.4. 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, 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 2.4. 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 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 restriction of  $\phi = 0$  would imply a greater sensitivity to travel distance. The results indicate that the first scenario fits the subjects invested: realizing the disadvantage of its location, 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 2.4 shows, these two hypotheses 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 store choice probabilities, when one of the five stores offers a promoted 15 percent price cut. The percentage changes in store

choice probabilities are computed using simulation as

$$\begin{aligned}\frac{\Delta\eta_s}{\eta_s} &= \frac{1}{\eta_s} \left( \frac{\partial\eta_s}{\partial p_{j_c}} \times 0.15 + \frac{\Delta\eta_s}{\Delta m_{j_c}} \right) = \int \int \int \iota \times \rho_{j_c} (-\alpha_c \times 0.15 + \beta_{c,1}) (1 - \eta_s) \times dF d\Omega dD, \\ \frac{\Delta\eta_q}{\eta_q} &= -\frac{1}{\eta_s} \left( \frac{\partial\eta_s}{\partial p_{j_c}} \times 0.15 + \frac{\Delta\eta_s}{\Delta m_{j_c}} \right) = \int \int \int -\iota \times \rho_{j_c} (-\alpha_c \times 0.15 + \beta_{c,1}) \times \eta_s \times dF d\Omega dD,\end{aligned}$$

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 rival's market share when the promotion is offered by  $s$ . 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 2.5. They show that price promotion is able to drive 0.12 to 0.47 percent increase in self store visit probability, and cause 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 stole from each rival store, computed using the estimated market size.

The parameter of ad exposure  $\phi$  plays a crucial role in stores 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 is 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 self 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 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 the proportion of informed consumers ( $\phi_s$ ), as store choice probabilities of the uninformed consumers won't change.

### 2.6.3 Promotion Costs

The promotional costs are estimated by comparing the actual profits generated by the actual merchandising decisions, and the alternative profits led by small deviations in  $\mathbf{m}$ . For the store-side observations that spans 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 2.6 displays the estimates of promotion cost (per promotion per week) at each of the five stores. The bound estimates of this cost ranges from about ten to twenty dollars or higher, varying largely across stores. To check if the estimates are reasonable in size, I compute stores' average total weekly and yearly expenditure on promotion, given the observed frequency of promotions (Table 2.1), and compare them with reported data. The lower bound estimates imply that for the store investigated in this study, promotions would at least cost \$307 to \$722 per week, or \$15,596 to \$37,540 per year; the upper bound for that spending is \$837 to \$1,634 per week, or \$23,940 to \$85,005 per year (in 1993 dollars). The national average yearly ad spending per supermarket in 1993 is estimated to be \$13,324.<sup>12</sup> Thus my estimates are plausible in magnitude though higher than the national average. On the one hand, there are reasons to believe that the promotion cost in the metropolitan area where the data is collected from is higher than the national average. On the other hand, as equation (2.9) shows, the magnitudes of bound estimates are closely related to the approximated market size, which serves as a scalar in the process of estimation. Since market size is approximated by the average consumption rate of observed product categories of the tracked households, the bound estimates of promotion cost could be improved if better knowledge about this parameter is provided. For example, a larger household sample size and a broader range of categories.

The magnitude of estimated bounds imply a substantial dispersion of promotional cost across stores. The cost is around ten dollars at EDLP1 with the smallest lower bound, while it could be as high as twenty dollars or higher at EDLP2 and HiLo. 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 the dispersion in the efficiency of the marketing division at different stores.

---

<sup>12</sup>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.

## 2.7 Welfare

### 2.7.1 Market Efficiency of the Current Equilibrium

Although theories do not provide unambiguous predictions on welfare, two effects of additional price advertising in the neighbourhood of private optimum on social gain are well understood; The first is the demand-creating effect, as price advertising reduces price uncertainty and would therefore motivate purchase. But since the firm that provides additional advertising is unable to appropriate all of the resulting social surplus (if the ad motivates purchase), the private advertising level tends to be socially inadequate. The second is the business-stealing effect. The 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, which suggests that advertising may be excessive (Tirole 1988). Therefore, whether the private advertising level, or promotion intensity in the supermarket retail industry, is too much or too little depends on which of the two effects dominates.

Advertising efficiency in the supermarket industry is complicated by a number of facts. First, additional advertising is likely to create new demand if consumers are not aware of availability of products and ads announce both product availability and price as in models by Bagwell (2007), Butters (1977), Grossman and Shapiro (1984), and Stegeman (1991). In the supermarket industry, however, product availabilities are typically well known, which limits the demand-creating effect. New demand could be created if (1) the additional advertising reaches a consumer who wouldn't have done any shopping otherwise, but now decides to visit the store, or (2) it reaches a remote consumer who would have shopped at a rival store and who now purchase a bigger bundle. Second, social surplus created will be eroded by increased transportation cost, in either case above: in (1), it's the transportation cost of the whole shopping trip for the new shopper; in (2), it's the cost of the extra travel distance for the remote consumer. Third, 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 the advertising is motivated by the profit margin of a product bundle, not the margin of the promoted product only.

In sum, market efficiency in this industry is complex. It is determined by the magnitudes

of the two opposite effects that are complicated by transportation cost of shopping and basket shopping behavior. I examine this problem by simulating counterfactual outcomes and computing surpluses following small deviations to private promotion levels. If providing an additional promotion improves social welfare, then the private level is socially inadequate, otherwise it is excessive.

### Welfare Measures

Social welfare for the retail market is measured as the sum of total producer surplus and total consumer surplus, induced by store decisions  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_s, \dots, \mathbf{x}_S)$ :

$$W(\mathbf{x}) = \sum_s PS_s(\mathbf{x}) + \sum_h CS_h(\mathbf{x}) = PS(\mathbf{x}) + CS(\mathbf{x}), \quad (2.26)$$

where individual producer surplus  $PS_s$  is the expected profit excluding fixed cost, and individual consumer surplus  $CS_h$  is the expected net gain of a shopping trip, both measured in dollars. All surpluses are computed using the observed choices averaging across time.

The individual producer surplus  $PS_{st}$  is measured as the expected payoff (excluding fixed cost) of  $s$  in period  $t$  induced by strategy portfolio  $\mathbf{x}_t = (\mathbf{x}_{1t}, \dots, \mathbf{x}_{st}, \dots, \mathbf{x}_{St})$ .  $PS_{st}$  is computed using parameter estimates and optimal pricing and promotion decisions, as the following:

$$PS_{st}(\mathbf{x}) = \frac{MS}{H} \times \sum_h \sum_{ad_{ht} \in AD} \widehat{prob}(ad_{ht}) \cdot \hat{\eta}_{hst}(\mathbf{x}_t, ad_{ht}, dist_{hs}) \cdot \hat{\rho}'_{st}(\mathbf{p}_{st} - \widehat{mc}_s) - \hat{\theta}_s \cdot (\mathbf{1}' \mathbf{m}_{st}). \quad (2.27)$$

Consumer surplus is the expected gain from a shopping trip. The utility from a shopping trip must be rescaled in a way such that the gain is measured in dollars. First I compute the surplus generated from purchase,

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

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

surplus from purchase, averaging across time and stores:

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

Finally, consumer surplus is given by

$$CS_{ht}(\mathbf{x}_t) = \frac{1}{|\bar{\alpha}|} \frac{1}{|\bar{\ell}|} \log \left( 1 + \sum_s \exp(\hat{\lambda}_s + \hat{u}_{hst}(ad_{hst}) + \hat{\kappa} dist_{hs}) \right). \quad (2.28)$$

### Simulation Results of Surpluses

For the numerical simulation, first of all, I compute surpluses of the current equilibrium as the base case using observed store behavior and demand estimates. The results are displayed in Table 2.7. Other market outcome variables are the probability of non-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 regular price, weighted by within-category market share of each product.

To distinguish the two effects of price advertising, I compute surpluses including and excluding transportation cost. A surplus excluding transportation cost represents the surplus created by quantity, while a surplus including this cost represents the surplus after transportation erosion. The results in Table 2.7 show that in the current equilibrium, transportation erosion takes a considerable portion of consumer surplus that would have been gained from purchase bundles. Consumer surplus is \$898.58k excluding transportation cost and \$722.19k including the cost. This implies about that 20 percent of consumer surplus that would have been gained from purchase has been spent on travelling.

Next, I simulate the change in surpluses ( $\Delta CS$  and  $\Delta W_s$ ), the change in expected revenue of the deviating store  $\Delta R_s$ , and the total change in rival stores' revenue  $\sum_{q \in -s} \Delta R_q$ , when each of the five stores makes a hypothetical deviation by offering one more or less 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 2.8, 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 intervals of change in social surplus  $\Delta W$  following the deviation in promotion are computed based on the bounds of  $\theta_s$  in Table 2.6. I find that  $\Delta W < 0$  ( $> 0$ ) when one more (less) promotion is offered by any of the stores.<sup>13</sup> This means that the private promotion levels are inefficient and are socially excessive: an additional promotion won't expand quantity sufficiently to offset the extra promotion cost, while withdrawing one will save the society more than the loss from declined sales.

To see the result, let  $W^*$  denote the equilibrium welfare induced by the equilibrium decisions  $\mathbf{x}^*$ , and let  $W'_s(\mathbf{x}')$  denote the welfare induced by a deviation, where  $s$  denotes the deviating store. Now consider the case where the deviating store offers one more promotion. Store  $s$  chooses promotion variable  $\mathbf{m}'_s = \mathbf{m}^*_s + e_{j_c}$ . The newly promoted product is now priced at the sale price, and prices are re-optimized at the deviated promotion decision. This deviation will result in presumably an increase in  $E[R_s]$ , a unit cost of promotion  $\theta_m$ , a decrease in rivals' expected revenues  $\sum_{q \in -s} E[R_q]$ , and some change in consumer surplus. Notice that the extra promotion will affect the behavior of three kinds of consumers. (1) the staying consumers, who still shop at  $s$  but enjoy a greater surplus provided by the extra promotion; (2) the new consumers who were non-shoppers in equilibrium but now decides to visit  $s$ ; <sup>14</sup> and (3) the switched consumers who shopped at rival stores. To illustrate the point I decompose  $\Delta R_s$  into three terms: the change in the deviating store's revenue from the staying shoppers  $\Delta E[R_s]^{stay}$ , the revenue from new shoppers  $R_s^{new}$ , and the revenue from switched shoppers  $E[R_s]^{switch}$ . Similarly,  $\Delta CS$  is decomposed into the change in consumer surplus of the staying shoppers  $\Delta CS^{stay}$ , and surplus of the new shoppers  $CS^{new}$ , and the

<sup>13</sup>The lower bound of  $\Delta W$  is negative (but close to zero) when EDLP1 or HiLo drops one promotion.

<sup>14</sup>Besides those who didn't shop, the new consumers may also include those who switched from unobserved stores (treated as outside choice) to  $s$ . However, this won't cause any computation bias because given that the shopping utility of outside choice has been normalized to zero,  $\Delta CS^{new}$  measures the absolute change in consumer surplus for either new shoppers or shoppers switched from unobserved stores. I thank Tom Prusa for pointing this out.

change in surplus of the switched shoppers  $\Delta CS^{switch}$ .

$$\begin{aligned}
\Delta W_s &= W(\mathbf{x}'_s) - W(\mathbf{x}^*) \\
&= \Delta PS_s + \sum_{q \in -s} \Delta PS_q + \Delta CS_s \\
&= \Delta E[R_s] - \theta_{m,s} + \sum_{q \in -s} \Delta E[R_q] + \Delta CS_s \\
&= (\Delta E[R_s]^{stay} + R_s^{new} + E[R_s]^{switch}) - \theta_{m,s} + \sum_{q \in -s} \Delta E[R_q] \\
&\quad + (\Delta CS_s^{stay} + CS^{new} + \Delta CS_s^{switch}).
\end{aligned} \tag{2.29}$$

The simulation results in Table 2.8 show that social welfare is harmed by the extra promotion ( $\Delta W < 0$ ). To explain this, first notice that in the neighbourhood of optimum, the change in the deviating store's revenue ( $\Delta E[R_s]^{stay} + R_s^{new} + E[R_s]^{switch}$ ) should be very close to  $\theta_s$ . The remaining components are the increase in consumer surplus  $\Delta CS_s$ , and the decrease in rivals' profit,  $\sum_{q \in -s} \Delta E[R_q]$ , which measures the business-stealing effect. Thus, the change in  $W$  followed by the deviation depends on the absolute value of the two components. Table 2.8 suggests that the new consumer surplus created by the extra promotion is not sufficient to offset the decrease in rivals' profit.

Why the extra consumer surplus is small while the business-stealing effect is large? To explain this, let's examine the three components of  $\Delta CS$  more closely.  $\Delta CS_s^{stay}$ , presumably positive, results from the increase of purchase incidence and product switching within the category of the newly promoted product (the staying shoppers' purchase behavior in other categories will stay the same).  $CS^{new}$  is the consumer surplus of those who would have not shopped but now shop at the deviating store.  $\Delta CS_s^{switch}$  results from the difference of surpluses between shopping at the deviating store  $s$  and a rival stores that would have otherwise been visited.

The first reason for small extra consumer surplus is due to the limited demand-creating effect in the supermarket setting where product existence is typically well known. The extra promotion conveys price information only and does not announce product availability. Therefore, the quantity expansion is small. In contrast, the demand-creating effect studied in the literature would be much larger because price advertisings are assumed to announce product existence. The second reason has to do with transportation cost. The

extra consumer surplus generated by the extra promotion (and lower price) is eroded by the increased transportation cost. Since a transportation cost must be paid for shopping, a consumer won't shop unless the transportation cost can be at least offset by the surplus generated from the purchase bundle. This erosion applies to the new shoppers and the switched shoppers. For a new shopper, the surplus generated from the purchase bundle at  $s$  will be eroded by the transportation cost that must be paid for travelling to  $s$ ; for a switched shopper, if she obtains extra surplus from the bundle purchased at  $s$  compared to the bundle that would have been purchased at a rival store, the extra surplus is eroded by the increased transportation cost. These two factors, the limited demand-creating effect and transportation cost erosion, together result in the small extra consumer surplus. On the contrary, the incentive of business stealing is big. By offering on extra promotion, the deviating store is able to appropriate not only the profit marginal of the promoted products but also the margins of other high-priced items in the same purchase bundle. This business-stealing conduct implies a re-distribution of margins from one firm to another and contributes nothing positive to social surplus.

In sum, the demand-creating effect of promotions are too small to offset the business-stealing effect; thus private promotion intensities are socially excessive. Though extra demand is created by promotion, the surplus gain from quantity expansion has been eroded by transportation cost, implying a significant welfare-harming effect of price advertising.

### **2.7.2 Counterfactual Experiments**

#### **Changes in Promotion Costs**

To better understand the role transportation cost plays in market efficiency, I simulate market outcome in two counterfactual experiments, with a small overall increase and decrease in promotion costs, respectively. Due to the discrete nature of promotion decision, I allow for a 5 percent deviation in promotion cost 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.

As shown in the last section, the inefficiency due to shopping transportation cost dampens the demand-creating effect, because it erodes consumer surplus that would have been gained from purchase bundles. Furthermore, stealing one customer from a rival store means a longer shopping trip if the shopper is attracted by a distant store. In a competition-intensified market where shoppers are more likely to travel long distances, transportation cost would cause a worse surplus erosion. In this sense, the effect of price advertising (promotion) is two-fold: it improves market efficiency by reducing price uncertainties and expands quantity; in the meantime, it creates inefficiency due to the higher transportation cost. If the latter effect is sufficiently large, the loss from longer shopping distance would outweigh the surplus gain from quantity expansion, and price advertising may no longer be welfare improving.

Table 2.9 reports the market outcome variables of the two experiments. When price advertising increase by 5%, in the new optimum retailers choose to promote less but price higher. Producer surplus is slightly lower than its counterpart in the baseline case, because the promotion cost saved is not sufficient to offset the decline in revenue. The higher price levels and less price information negatively affect shopping probabilities and purchase incidences. As a result, consumer bundle surplus (consumers' gain from what they buy, or *CS* excluding transportation cost) has decreased from \$898.58k to \$878.53k. Shoppers are less likely to travel to distant stores and therefore the cost spent on travelling has decreased from \$176.39k to \$152.87k, since stores now compete less intensively and promote less, compared to the baseline case. The reduction in transportation cost (\$23.52k) is so large that it offsets the decrease in bundle surplus (\$20.06k). Despite of less promotions and higher prices, consumers are better off, since now the transportation cost that can be saved is bigger than the decreased in bundle surplus.

When promotion costs decreases by 5%, we see the reverse. Competition among stores has been intensified; stores promote more and price lower, and shoppers visit distant stores with higher probabilities compared to the base case. This outcome is in line with the 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. However, shoppers are worse off even if they now enjoy lower prices and more promotions. The

reason is that, similar to the first experiment, the increase in transportation cost (\$17.22k) outweighs the increase in bundle surplus (\$8.85k).

The comparison between experimental social welfare in Table 2.9 and its baseline counterpart suggests that, instead of the usual conclusion that social surplus can always be improved by information and intensified competition, we see the opposite: social welfare has been improved (\$1,403.62k  $\bar{}$  \$ 1,401.17k) with higher prices and less price information, and worsened (\$1,398.32k  $\bar{}$  \$1,401.17k) when competition among stores is intensified with lower prices and more price information. This is a striking finding but can be explained by the two-fold effect of price advertising in this market. First notice that social welfare excluding transportation cost measures the pure effect of price advertising (and pricing) on quantity. This variable has decreased by \$21.07k (from \$1,577.56k to \$1,556.49k) when promotions become expensive, and increased by \$14.35k to \$1591.91k when promotions are cheap, paralleling with quantity reduction and expansion in the two experiments, respectively. However, these changes are respectively dominated by the change in transportation cost. As a result, the welfare implication departs from the usual pattern; the welfare-harming effect of price advertising plays a crucial role here.

### Online Grocery Shopping

The counterfactual experiment in this section aims to simulate the emerging shopping regime where a growing proportion of consumers choose online ordering and home delivery for grocery shopping. Since product bundles are delivered by the store, shopping travel distance no longer affects store choice.<sup>15</sup> Another feature of home-delivered grocery shopping, compared to traditional shopping, is that the search cost for prices is much lower: information on weekly specials are available online and browsing them is easy.

To simplify online shopping behavior, I make three assumptions. First, shipping is free and thus home-store distance does not affect store choice. Second, all shoppers read promotion circulars posted by all stores and evaluate the expected shopping utilities before making store choice; once the store choice is made, the bundle of product choices is suboptimal conditional on store choice. Third, unobserved store characteristics that affect online

---

<sup>15</sup>In reality the shipping fee is related to home-store distance. I assume shipping is free for simplicity.

grocery shopping regime, e.g., shipping cost and service, are of the same quality as those in the traditional shopping regime. This means that the coefficient associated with transportation cost  $\kappa$  is set to zero, all shoppers are informed by promotions ( $\phi_s = 1$ ), and store dummies  $\lambda_s$  stay constant. By setting  $\kappa$  to zero, the stores' local market power due to the spatial factor is eliminated.

How would the market perform under this shopping regime? Firstly, it is clear that the new regime avoids erosion and therefore improves consumer surplus. Secondly, spatial models of store 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 locations give firms market power. Therefore, we expect that competition among stores will be intensified due to the removal of stores' local market power. However, the effect on social welfare is vague: the stores could compete more aggressively by spending greater in price advertising. The welfare implication depends on the magnitudes of the increase in consumer surplus due to the avoidance of transportation, the social gain due to quantity expansion, and the increase in the total promotion cost.

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 2.10. The top section lists consumer surplus, producer surplus, social welfare, and the probability of non-shopping. As expected, comparing to the base case, the new shopping regime intensifies competition. Price indices

have been driven down by 5 to 12 percent, and promotion intensity has increased by 9 to 24 percent. Stores' total profit has fell slightly from \$678.98k to \$665.46k. The difference in the changes of stores' market share and profits could be due to the alternative setting that the probability of ad exposure,  $\phi_s$ , equals one for all stores, whereas in the base case these probabilities differ widely across stores. Intensified competition results in higher probability of shopping incidences and purchase incidences: the pure effect of quantity expansion on consumer surplus ( $CS$  excluding transportation costs) has increased from the baseline \$898.58k to \$1171.22k. If include, consumer surplus has increased by \$449.03k or 62 percent. Finally, the numerical results show that the new shopping regime is welfare improving: the increase in consumer surplus outweighs the decline in producer surplus; social welfare has been improved by 31 percent.<sup>16</sup>

## 2.8 Conclusion

Theories of informative price advertising do not provide unambiguous predictions on its welfare implication. Moreover, a few facts about the supermarket industry deviate from the common assumptions made in standard theories. These assumptions include single-product oligopoly firms and single-product shopping behavior,<sup>17</sup> consumers' unawareness of product availability unless informed by price advertising<sup>18</sup>, and firms' optimization by choosing the market "reach" of advertising. In reality, however, 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. These facts determine the way and magnitude in which welfare is affected by price advertising.<sup>19</sup>

---

<sup>16</sup>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 is strictly positive. By economies of scale the delivery cost of a grocery bundle can be small, if orders are delivered using efficient logistic system, such as van trucks

<sup>17</sup>One exception is the Hotelling model of two-product duopoly constructed by Lal and Matutes (1994).

<sup>18</sup>See, for example, Bagwell (2007), Butters (1977), Grossman and Shapiro (1984), and Stegeman (1991).

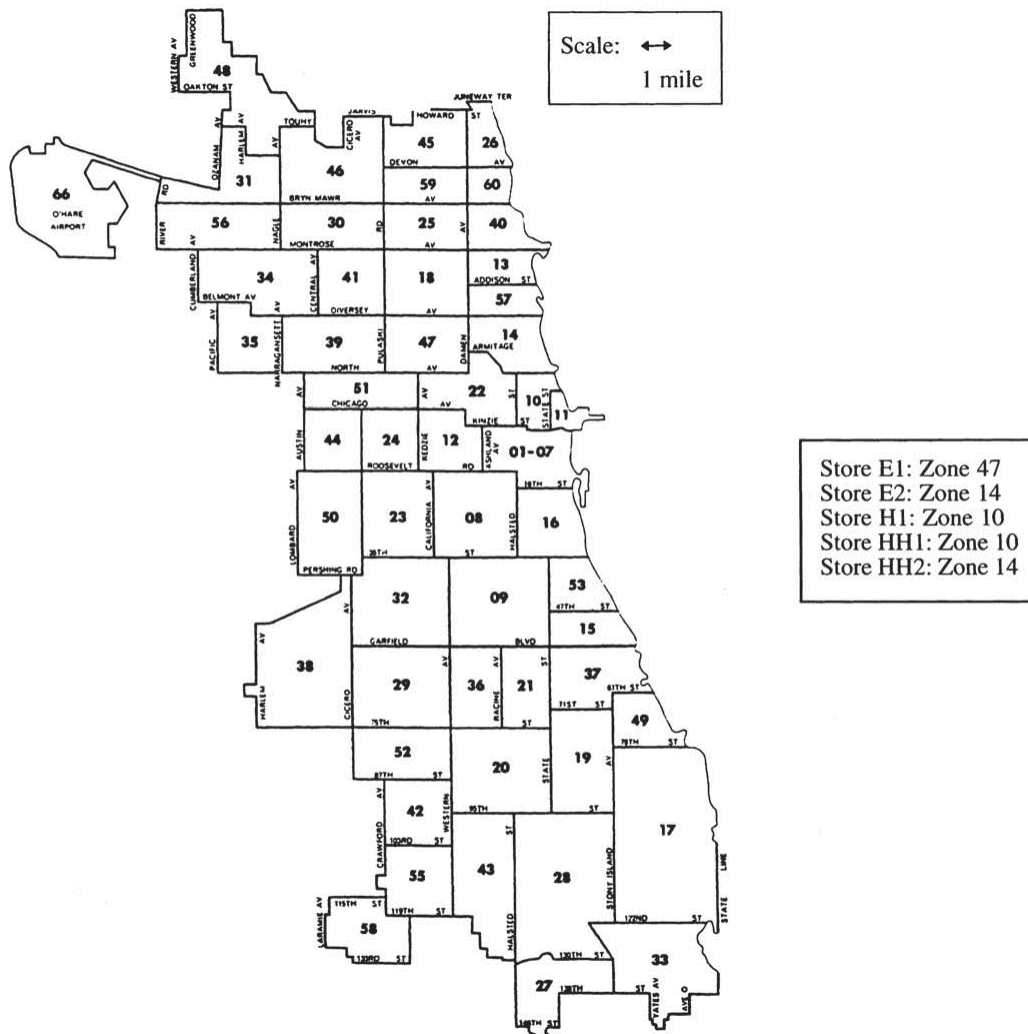
<sup>19</sup>This paper focuses on the effect of informative advertising which conveys price information only and distinguishes with the literature that examines welfare implication of 'persuasive' advertising Dixit and Norman (1978); Stigler and Becker (1977); Nichols (1985), where advertisements shift consumer preference.

In this study, the demand-creating effect and the business-stealing effect of price advertising are complicated by shopping transportation cost. The effect of price advertising (promotion) is two-fold: it improves market efficiency by reducing price uncertainties and expands quantity consumed; but it creates inefficiency due to the higher transportation costs as consumers are attracted by promotions at distant stores. If the latter effect is sufficiently large, the loss from longer shopping distance would outweigh the surplus gain from quantity expansion, and price advertising may no longer be welfare improving.

I empirically examine the efficiency of such markets, using simulation methods, by studying changes in equilibrium social surplus following small deviations in promotion intensity. To do this, a spatial model that accounts for consumer shopping and retailer pricing behavior is built. Using scanner data of consumer shopping and store merchandising information, consumer preference is estimated following the discrete choice literature; the retailers' marginal costs of promotion are structurally estimated using the moment inequality approach. These structural estimates allow me to simulate equilibrium and counterfactual outcomes.

The simulation results numerically show that the equilibrium promotion levels are socially excessive, because the demand-creating effect of advertised price cuts, after transportation erosion, is too small to outweigh the business-stealing effect. Motivated by the welfare implications of shopping transportation costs, I artificially remove this cost and simulate the counterfactual market outcome, as an experiment of online grocery shopping. I found this new shopping regime is welfare improving, for two reasons. First, obviously, social surplus is not eroded by transportation. Second, the removal of transportation cost implies that the firms no longer possess local market power, and therefore intensifies competition (lower prices and more promotions). Thus, despite the higher probability of shopping at distant stores, no higher transportation cost will be paid. In sum, if intensified competition leads to worse surplus erosion by transportation, the welfare effects of price advertising may deviate from the usual conclusion that information and competition always improve social well-being. But this erosion can be avoided by, for example, online shopping regime.

Figure 2.1: 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 2.1: Summary Statistics

<i>Store Pricing and Promotions</i>					
Store	Market Share <sup>1</sup>	Average Price Index <sup>2</sup>	Weekly Frequency		
			Promotions	Price Cuts	Deep Price Cuts <sup>3</sup>
EDLP1	0.2147	0.8334	41.53	54.03	24.71
EDLP2	0.2678	0.8327	46.46	59.64	26.34
HILO	0.3254	0.8602	58.38	73.37	48.66
HT1	0.0904	0.9124	32.06	44.91	17.97
HT2	0.1013	0.9253	39.37	34.15	18.02

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

<sup>1</sup>Market share refers to the proportion of store visits at a specific store.

<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 2.2: Product Preference

Category	i				ii			
	$\alpha_c$	$\beta_{1,c}$	$\beta_{2,c}$	$\chi_c$	$\alpha_c$	$\beta_{1,c}$	$\beta_{2,c}$	$\chi_c$
Bacon	-0.5551 (0.0235)	1.1236 (0.0540)	0.9911 (0.0390)	1.6394 (0.0480)	-0.8038 (0.0424)	1.0573 (0.0588)	0.6185 (0.0469)	0.4638 (0.0346)
Butter	-0.5240 (0.0223)	1.1684 (0.0298)	1.2192 (0.0284)	0.0893 (0.0147)	-0.4879 (0.0285)	1.1435 (0.0301)	1.1582 (0.0291)	1.1447 (0.0098)
Cereal	-0.0769 (0.0174)	1.3594 (0.0556)	0.9480 (0.0688)	2.3350 (0.0320)	-0.2272 (0.0227)	1.3092 (0.0465)	0.8724 (0.0586)	0.2462 (0.0277)
Toothpaste	-0.4913 (0.0334)	0.9557 (0.0514)	0.8493 (0.0805)	2.1228 (0.0223)	-0.9273 (0.0678)	0.9013 (0.0551)	0.7265 (0.0227)	0.2291 (0.0201)
Coffee	-0.0478 (0.0088)	1.5691 (0.0091)	1.4485 (0.0397)	1.6968 (0.0017)	-0.0730 (0.0142)	1.3580 (0.0152)	0.9752 (0.0416)	2.7209 (0.0128)
Crackers	-0.2953 (0.0277)	1.6941 (0.0388)	0.8704 (0.0493)	1.8944 (0.0359)	-0.8026 (0.0432)	1.5336 (0.0329)	0.7612 (0.0691)	0.0902 (0.0280)
Detergent	-0.1384 (0.0053)	1.6808 (0.0348)	1.0894 (0.0434)	1.8578 (0.0117)	-0.3069 (0.0149)	1.4024 (0.0364)	0.8965 (0.0865)	0.6072 (0.0627)
Eggs	-1.0955 (0.0412)	2.0032 (0.1689)	1.5630 (0.0585)	0.9806 (0.0206)	-1.0732 (0.0738)	1.8571 (0.1678)	1.5075 (0.0613)	1.3176 (0.0326)
Hot Dogs	-0.3778 (0.0112)	1.1609 (0.0399)	0.8919 (0.0286)	0.4280 (0.0117)	-0.6395 (0.0215)	0.9984 (0.0411)	0.6394 (0.0306)	0.0949 (0.0217)
Ice Cream	-0.6180 (0.0210)	0.9821 (0.0643)	0.6960 (0.0424)	1.7651 (0.0101)	-0.9029 (0.0264)	0.9349 (0.0647)	0.3529 (0.0453)	1.9362 (0.0195)
Peanuts	-0.3537 (0.0234)	1.4422 (0.0741)	0.9202 (0.8818)	0.0796 (0.0380)	-1.0187 (0.0297)	1.2167 (0.0793)	0.7213 (0.9312)	2.0246 (0.0289)
Frozen Pizza	-0.1347 (0.0119)	1.0583 (0.0401)	1.0286 (0.0359)	0.6923 (0.0260)	-0.6724 (0.0368)	0.7734 (0.0407)	0.7552 (0.0428)	0.3428 (0.0688)
Potato Chips	-0.2372 (0.0255)	1.4522 (0.0458)	1.1792 (0.0512)	0.1154 (0.0145)	-0.5471 (0.0311)	1.6276 (0.0416)	1.3291 (0.0313)	1.1007 (0.0079)
Soap	-0.2783 (0.0173)	1.1905 (0.0503)	0.9849 (0.0725)	0.2428 (0.0208)	-0.6982 (0.0199)	1.1275 (0.0524)	0.8904 (0.0471)	0.6679 (0.0527)
Sugar	-0.1537 (0.0174)	1.3432 (0.0552)	1.9931 (0.0545)	2.0586 (0.0134)	-0.9092 (0.0909)	0.7830 (0.0622)	0.6784 (0.0663)	1.6425 (0.0390)
Tissue paper	-0.5665 (0.0131)	0.8383 (0.0253)	1.2240 (0.0257)	1.7371 (0.0105)	-0.8071 (0.0376)	1.3745 (0.0268)	1.0506 (0.0284)	0.6147 (0.0200)
Paper Towel	-0.2751 (0.0148)	1.1172 (0.0272)	1.1819 (0.0287)	0.7927 (0.0284)	-0.9016 (0.0241)	1.2179 (0.0285)	1.0552 (0.0320)	1.5670 (0.0297)
Yogurt	-0.1100 (0.0193)	0.8034 (0.0564)	1.1706 (0.0378)	2.3756 (0.0656)	-0.2283 (0.0197)	0.7361 (0.0572)	1.2024 (0.0426)	1.3485 (0.0234)

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 2.3: Effects of Promoted Price Cuts – Percentage Increase\* in Product Choice Prob.( $\rho$ )

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

\*The percentage increase is

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

Table 2.4: Store Preference

	joint	$H_1$		
		$\phi = 0$	$\phi = 1$	$\phi_s = \phi$
$\kappa(\text{distance})$	-0.4659 (0.1072)	-0.3789 (0.0833)	-0.8977 (0.0892)	-0.4328 (0.1377)
$\iota(\text{merchandising utility})$	0.0397 (0.0029)	0.0986 (0.0021)	0.0008 (0.0001)	0.0432 (0.0238)
$\lambda_s(\text{store dummy})$	0.5583 (0.0477)	3.7267 (1.4742)	0.4791 (0.0835)	2.7982 (1.1721)
	0.2593 (0.0769)	3.9117 (1.2008)	0.6344 (0.1742)	4.8903 (1.1905)
	0.1148 (0.0496)	3.6297 (0.9073)	0.6730 (0.2746)	4.2877 (2.0033)
	-0.7954 (0.1724)	2.7240 (1.1355)	-0.3411 (0.7829)	3.5786 (0.9982)
	-1.0659 (0.3233)	2.6685 (1.1084)	-0.4491 (0.0672)	3.2090 (1.7953)
$\phi_s(\text{prob. of ad exposure})$	0.0350 (0.0102)			0.7230 (0.2447)
	0.1958 (0.0473)			
	0.1784 (0.0509)			
	0.0277 (0.0076)			
	0.0389 (0.0122)			
Log-likelihood	$-4.3941 \times 10^4$	$-4.4118 \times 10^5$	$-8.0716 \times 10^5$	$-4.2946 \times 10^4$
Likelihood ratio		$7.9448 \times 10^5^*$	$1.5264 \times 10^6^*$	1,990*

\*significant at the 1% level.

$$U_{hst} = \lambda_s + \iota \bar{u}_{hs} + \kappa \text{dist}_{hs} + \zeta_{hst}.$$

Table 2.5: Effects of Promoted Price Cuts – Percentage Changes\* in Store Choice Prob.( $\eta$ )

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

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

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

$$\frac{\Delta\eta_q}{\eta_q} = -\frac{1}{\eta_s} \left( \frac{\partial\eta_s}{\partial p_{j_c}} \times 0.15 + \frac{\Delta\eta_s}{\Delta m_{j_c}} \right) = \int \int \int -\iota \times \rho_{j_c}(-\alpha_c \times 0.15 + \beta_{c,1}) \times \eta_s \times dF d\Omega dD, \text{ where } q \neq s.$$

Table 2.6: Estimated Bounds of Promotion Cost ( $\theta$ )

	$\theta_s$	95% Confidence Interval	
		LB	UB
EDLP1	[7.40, 11.09]	[6.33, 10.12]	[10.36, 12.68]
EDLP2	[7.88, 22.60]	[5.64, 12.48]	[19.00, 27.09]
HiLo	[12.36, 28.00]	[9.81, 14.63]	[26.41, 31.68]
HT1	[10.99, 14.35]	[5.79, 12.20]	[9.04, 15.86]
HT2	[8.75, 21.28]	[6.46, 10.69]	[16.96, 22.30]

Table 2.7: Current Equilibrium (Base Case)

	transportation costs	
	included	excluded
CS ( $\times 10^3$ )	722.19	898.58
PS ( $\times 10^3$ )	678.98	
W ( $\times 10^3$ )	1,401.17	1,577.56
transportation cost( $\times 10^3$ )	176.39	
non-shopping prob	0.3121	

	# of Promo.	Ave. Price Ind.	Store Choice Prob.	$\pi (\times 10^3)$
EDLP1	24.23	0.8450	0.1042	102.09
EDLP2	27.44	0.8943	0.1515	148.02
HiLo	30.46	0.8474	0.2165	212.92
HT1	20.43	0.9396	0.1381	135.38
HT2	22.75	0.9142	0.0766	76.26

Table 2.8: Market Efficiency

<i>one more promotion</i>							
	$\Delta E[R]$					$\Delta CS$	$\Delta W$
	EDLP1	EDLP2	HiLo	HT1	HT2		
EDLP1	9.45	-2.88	-6.36	-2.48	-2.38	8.95	[-6.81, -3.35]
EDLP2	-11.04	38.11	-12.04	-10.74	-10.74	8.36	[-20.70, -5.59]
HiLo	-7.56	-12.34	23.38	-3.48	-4.17	11.63	[-20.54, -4.90]
HT1	-2.88	-3.08	-3.78	10.15	-3.58	7.74	[-9.78, -6.43]
HT2	-3.48	-3.68	-9.05	-1.39	9.55	7.94	[-21.39, -8.86]

<i>one less promotion</i>							
	$\Delta E[R]$					$\Delta CS$	$\Delta W$
	EDLP1	EDLP2	HiLo	HT1	HT2		
EDLP1	-12.63	2.98	5.87	2.98	2.68	-9.35	[-0.05, 3.63]
EDLP2	11.24	-19.85	12.53	10.94	11.24	-11.18	[22.82, 37.53]
HiLo	6.76	12.83	-26.32	3.08	4.67	-13.63	[-0.22, 15.41]
HT1	3.18	3.38	4.67	-10.52	2.78	-8.47	[6.03, 9.38]
HT2	2.88	3.78	9.25	1.69	-12.42	-9.14	[4.80, 17.33]

$\Delta W = \Delta R_s + \sum_{q \in -s} \Delta R_q + \Delta CS = (\Delta R_s \pm \theta_s) + \sum_{q \in -s} \Delta R_q + \Delta CS$ .  
The ranges of  $\Delta W$  are computed based on the bounds of  $\theta_s$ .

Table 2.9: Counterfactual – Overall Changes in Promotion Costs  $\theta$ 

<i>an overall 5% increase</i>				
	transportation costs			
	included		excluded	
CS ( $\times 10^3$ )	725.65		878.52	
PS ( $\times 10^3$ )	677.97			
W ( $\times 10^3$ )	1,403.62		1,556.49	
transportation cost( $\times 10^3$ )	152.87			
non-shopping prob	0.3533			
	# of Promo.	Ave. Price Ind.	Store Choice Prob.	$\pi (\times 10^3)$
EDLP1	21.34	0.8838	0.0957	102.09
EDLP2	26.25	0.9042	0.1345	148.02
HiLo	27.07	0.8902	0.2083	212.92
HT1	18.72	0.9608	0.1333	135.38
HT2	20.47	0.9425	0.0749	76.26
<i>an overall 5% decrease</i>				
	transportation costs			
	included		excluded	
CS ( $\times 10^3$ )	713.82		907.43	
PS ( $\times 10^3$ )	684.50			
W ( $\times 10^3$ )	1,398.32		1,591.91	
transportation cost( $\times 10^3$ )	193.61			
non-shopping prob	0.2513			
	# of Promo.	Ave. Price Ind.	Store Choice Prob.	$\pi (\times 10^3)$
EDLP1	26.05	0.8260	0.1229	101.19
EDLP2	28.14	0.8503	0.1526	141.13
HiLo	33.03	0.8127	0.1903	224.13
HT1	22.81	0.9210	0.1214	138.96
HT2	23.75	0.9052	0.0782	79.09

Table 2.10: Counterfactual – Online Grocery Shopping

CS ( $\times 10^3$ )	1,171.22			
PS ( $\times 10^3$ )	665.46			
W ( $\times 10^3$ )	1,836.69			
transportation cost( $\times 10^3$ )	0			
non-shopping prob	0.1114			
	# of Promo.	Ave. Price Ind.	Store Choice Prob.	$\pi$ ( $\times 10^3$ )
EDLP1	27.62	0.7728	0.1641	97.96
EDLP2	34.02	0.7937	0.2038	137.15
HiLo	33.20	0.8067	0.2542	213.57
HT1	23.15	0.8344	0.1621	136.46
HT2	24.82	0.8221	0.1045	76.87

## Chapter 3

### Dynamic Dispersion of Storable Good Prices

#### 3.1 Introduction

Supermarket retailers adjust retail prices on a weekly basis, and change the set of promoted products time to time. In each product category, a shopper can almost always find some products on promotion, marked with eye-grabbing price labels or piled at the entrance of aisles. Supermarket managers clearly find it more profitable to put different items on sales time to time, in spite of the high menu cost and administrative cost (Levy et al. 1997), than keeping prices constant at some high level over time (like what convenience stores usually do). Nowadays, not only can consumers download on-line next week's promotion brochures of local super markets, but also smart phone applications designed by specific supermarket chains to advertise their in-store sales items are available.

When goods are storable (such as ground beef, laundry detergent, potato chips, etc), consumers purchase more than they will consume in that period and store the rest as inventory. Consumer heterogeneity with respect to "willingness to wait" is well acknowledged by store managers. While the regular price is offered to impatient consumers, the manager would reduce the price in order to "clear out" the demand of those who have been patiently waiting for the next sale to occur. Price promotion serves as a means of intertemporal price discrimination. On the other hand, because the current sales induce stockpiling, future demand would be jeopardised. Therefore, store managers tend to offer medium-size price reductions in order to preserve future sales. Temporary price reductions also serve the role of price discrimination between informed and uninformed consumers. Informed consumers with small search costs purchase at the low price occasionally offered, while uninformed consumers with high search cost purchase at the regular price, unless they encounter the low price by chance.

Empirical studies show that several price variation patterns are widely observed. Using retail price data over 20 categories for 5 years from 30 US Metropolitan areas, Hosken and Reiffen (2004) find that most items carried can be characterized as having a high regular price, and most deviations from that price are below that level. Price variations mostly results from store promotion, with insignificant impact from wholesale price changes. Therefore, price reductions are likely to be temporary, and will go back to the regular level soon.

Despite the ubiquitous nature of price promotions, there is little common ground among economists as to why supermarket retailers occasionally offer products at discounted prices, or even how and why such price dispersion can exist as an equilibrium phenomenon. To better understand retailers' strategic pricing behavior given consumer heterogeneities in store loyalty, willingness to wait or inventory cost, and knowledge in prices, I construct a model of an oligopoly retailers selling a homogeneous storable good. The good in the model are assumed to be consumed for multiple periods, and therefore does not need purchased frequently. Stores can sell the good at a *regular* price, or hold sales, selling the good at lower prices. Under the infinite horizon setting, the *High* type consumers are assumed to be loyal to a specific store, have infinite inventory cost and therefore never stockpile, and do not search for price reductions; they purchase only when the good is ran out, at the store they are loyal to. The *Low* type consumers search for price at zero search cost, purchase if the lowest price offered is below some critical price, and store it at some inventory cost.

In this paper a symmetric Markov-perfect equilibrium (MPE) is found. As in the classic search models, the competing stores face a trade-off between selling only to its own loyals at the regular price and to both loyals and shoppers at some sale price. Retailers randomize prices, and the *cdf* of the equilibrium price distributions have a mass point at the regular price  $p^R$ . Moreover, the equilibrium price distribution is a function of the shoppers' inventory. The mixed strategy equilibrium is characterized by a critical price depend upon which purchase decision is made in each period. The realized price evolution consists of several consecutive regular-price periods, where no sales are offered, and occasionally one-time price reductions. The endogenous price evolution exhibits non-absorbing Markov transition of states: when shoppers hold high inventory, the probability of holding a sale is low, which

means inventory will more likely to drop down.

### 3.2 Literature

Models in the literature aim to explain the strategic pricing interactions among competing retailers and consumer purchase behaviors. Specifically, models are developed to generate price distributions, which characterize equilibrium, that have similar patterns to empirical observations. Two classes of models have been constructed. Both examine the pricing decision of single product retailers, and show how consumer heterogeneity can lead to retail price variation over time. Their basic setups in consumer heterogeneity consider the following factors: whether the two types (High and Low) differ in reservation price, willingness to wait, inventory cost, store loyalty, and search cost. The market structure considered is typically an oligopoly. The equilibria of the models are characterized by a price distribution, continuous or not, implying price is drawn from that distribution and therefore varies every period.

The first class assumes consumers differ in their knowledge. Since sellers face a trade-off between selling to only non-searchers at high price and selling to both searcher and non-searchers at the lowest price among all sellers, the symmetric mixed-strategy equilibrium features a continuous distribution of price. Varian (1980) model is the seminal contribution of this class. It describes a monopolistically competitive equilibrium in which sales are the outcome of mixed strategy equilibrium among retailers who compete over cohorts of informed and uninformed consumers. Narasimhan and Wilcox (1998) characterize competitive promotional strategies by their depth and frequency within a mixed strategy equilibrium similar to Varian (1980).

The second class views sales as means of price discrimination. Conlisk et al. (1984); Sobel (1984); Pesendorfer (2002a) fall into this type. Consumers differ in their reservation prices, willingness to wait for sales, and/or inventory costs (analytically equivalent to willingness to wait). Since the high types with high reservation price are not willing to wait for sales, while the low types only purchase at low prices, the equilibrium is characterized by purchases of high types in all periods, and periodically reduced price that is to "clear out" the low types.

Hong et al. (2002) is a combination of the two classes. In this model, not only do consumers differ in store loyalty and searching costs, but also only the low-type consumer store for inventory. Under this setup, oligopoly stores have an incentive to reduce price, both to sell to searchers and to consumer inventory. Observing consumer inventory, each store draw his own prices from a common distribution, and the low type consumers make purchase decision given the lowest price offered, taking into account the next-period state and payoffs predicted by her action and the transition rule. Because of the existence of store loyals, the Markov equilibrium price distribution has a mass point. Moreover, the equilibrium price distribution varies over time.

Though Varian's model can explain price variations of both perishable and nonperishable goods, it fails to predict the fact that most goods have a regular price as the price distribution has no mass point. The random price behavior that emerges from a mixed strategy equilibrium is fundamentally inconsistent with observed prices that tend to stay fixed for a long period of time and then fall temporarily, returning to the previous level after one period or two. This static model also fails to provide intuition of purchase at sales for inventory. The price discrimination type models (Conlisk et al. 1984; Sobel 1984; Pesendorfer 2002a) succeed in predicting a mass point in price distribution, yet they do not assume price searching behavior. Moreover, although these models are of infinite horizon, many have restrictions on the number of packages purchased, consumption amount in each period, or the maximum capacity of storage. For example, to fit into a Logit estimation, Hendel and Nevo (2006) assume that a consumer can purchase at most one package of laundry detergent. Hong et al. (2002) assumes that a consumer can store at most one package, and must consume one package every period.

Besides consumer heterogeneities, another dynamic explanation for incentive of price promotion is provided by Yang et al. (2005) in which price reductions are necessary to restore brand loyalty given its tendency to degrade over time. However, this loyalty restoring could be more likely actions of manufactures rather than retailers. If retail markets do have the incentive to restore loyalty of some brands, then it means that elasticities must vary across brands, and that the promoted brand probably has great impact on store revenue.

Finally, many authors argue that price promotions on a limited set of products may

serve as a tool of "luring" consumers come into store and shop other products within store at regular prices (Nevo and Wolfram 2002; Chintagunta et al. 2002; McAfee 1995). Since supermarkets virtually offer a diversity of products (20,000+ items), consumers with a purchase bundle face a trade-off between being loyal to one store and switching to another one for promotions at a transportation cost. Effective price promotions must increase store traffic. Furthermore, price promotion of one product is likely to cannibalize sales from other products within store or attracts customers from a rival retailer, if the degree of heterogeneity among store is low, but heterogeneity among products is high (Richards 2007).

### 3.3 Model

There are  $N$  identical retailers selling a homogeneous good in each time period  $t = 0, 1, 2, \dots$  in an isolated city of population 1. The good is assumed to be storable, and the package size is  $L$ . In each period, each consumer must consume 1 unit of the good, thus one package will last for  $L$  periods. A proportion  $\gamma$  of the population have zero search cost and will be referred to as *shoppers*, who will buy from the lowest price retailer, if they buy at all. A proportion  $1 - \gamma$  do not search, and each of them is loyal to a fixed and specific retailer. They will be referred to as *loyals*. I assume that the retailers have equal share of the non-searchers, and that the number of loyal customers who buy a package lasting  $L$  periods is the same in every period. In each period, a consumer, shopper or loyal, can purchase at most one package of the good. Shoppers can store the new package as inventory at a unit cost  $c$ . Loyals never purchase if they still hold inventory, and only purchase when they run out. Both shoppers and loyals value the per period consumption of the good at  $v$ . All agents share a common discount factor  $\delta$ . The good has a regular price  $p^R$  that satisfies  $p^R = \sum_{t=0}^{L-1} \delta^t v - \sum_{t=0}^{L-1} \delta^t (L - t)c$ , which is the total consumption utility net off the total inventory costs. Define any price strictly lower than  $p^R$  as a *sales price*. A retailer *holds a sale* in period  $t$  if she offers a sales price. Therefore, in each period, each retailer can sell to at least  $(1 - \gamma)/NL$  consumers at any price no greater than  $p^R$ . A loyal's surplus is zero if she buys at  $p^R$ , and positive if she buys at a sale.

I make three assumptions on shoppers' purchase behavior.

*Assumption 1.* The inventories of all shoppers are the same at the starting point.

*Assumption 2.* If a shopper runs out of stock, she must purchase at any price no greater than  $p^R$ .

*Assumption 3.* A shopper does not purchase if her inventory is greater than a positive integer  $K$ , where  $K < L$ . It implies that a shopper stores no more than one package of the good.

In each period, retailers simultaneously make pricing decisions  $p_{it}$ , and the set of prices offered in period  $t$  is denoted by  $P_t = \{p_{it} | i = 1, \dots, N\}$ . The marginal cost for retailers (wholesale price) is assumed zero for simplicity. Observing  $P_t$ , a shopper  $j$  maximizes the total expected discounted utility by making purchase decision, taking into account future states and prices. Denote the purchase decision of shopper  $j$  by  $d_{jt} = (d_{jt}^i)'$  satisfying  $d_{jt}^i \in \{0, 1\}$  and  $\sum_i d_{jt}^i \leq 1$ , and her inventory in period  $t$  by  $I_{jt}$ . If the shopper makes a purchase at retailer  $i$ , then  $d_{jt}^i = 1$ . The good not consumed in the current period will be stored at a unit cost  $c$ . Denote the state variable of this infinite-horizon problem by  $\iota_t$ , and in equilibrium it equals the synchronized shopper inventory, because as shown below the purchase decisions of all shoppers will be synchronized in equilibrium,  $d_t = d_{jt}^i, \forall j$ . It follows that the state transition rule is  $\iota_{t+1} = \iota_t - 1 + d_t L$ .

The shopper's problem is represented by

$$\begin{aligned}
 & \max_{\{\{d_{jt}^i\}_{t=0}^\infty\}} E \left[ \sum_{t=0}^{\infty} \delta^t u_{jt}(P_t, I_{jt}, d_{jt}^i) | \iota_1 \right] \\
 & \text{subject to } d_{jt}^i \in \{0, 1\}, \forall i, \\
 & I_{jt} \geq 0, \\
 & \sum_i d_{jt}^i = 1 \text{ if } I_{jt} = 0, \\
 & \sum_i d_{jt}^i = 0 \text{ if } I_{jt} > K, \\
 & I_{j,t+1} = I_{jt} - 1 + \sum_i d_{jt}^i L.
 \end{aligned} \tag{3.1}$$

The expectation is taken with respect to future states. The flow utility  $u_{jt}$  is given by

$$u_{jt}(P_t, I_{j,t+1}, \iota_t, d_{jt}) = -cI_{j,t+1} - d_{jt}p_{t,min} + v, \tag{3.2}$$

where  $p_{t,min} = \min P_t$ , the lowest price offered in period  $t$ ;  $I_{j,t+1} = I_{jt} - 1 + \sum_i d_{jt}^i L$  is the

low of motion. Define the value function  $V(\cdot)$  that represents a shopper's maximized payoff under her action rule:

$$V(P_t, I_{jt}, \iota_t) = \max_{\{d_{jt}\}} \left\{ E \sum_{t=0}^{\infty} \delta^t u_{jt}(P_t, I_{j,t+1}, d_{jt}, \iota_t) | d_{jt}, d_{-jt} \right\}. \quad (3.3)$$

Define shopper  $j$ 's choice-specific value function  $\tilde{V}(\cdot)$  as

$$\tilde{V}(P_t, I_{jt}, \iota_t, d_{jt}) = u_{jt}(P_t, I_{j,t+1}, \iota_t, d_{jt}) + \delta E[V(P_{t+1}, I_{j,t+1}, \iota_{t+1}) | d_{jt}, d_{-jt}]. \quad (3.4)$$

Retailers simultaneously make pricing decisions  $\{\{p_{it}\}_{t=0}^{\infty}\}$  to maximize the expected total payoff:

$$\max_{\{\{p_{it}\}_{t=0}^{\infty}\}} \sum_{t=0}^{\infty} \delta^t \pi_{it}, \quad (3.5)$$

where  $\pi_{it}$  is the per-period profit and  $\delta$  is the discount factor. In period  $t$ , a retailer  $i$  can always sell to its royals who ran out of stock at any price no greater than  $p^R$ . And it sells to shoppers if it offers the lowest price and if the shoppers buy at all. Retailer  $i$ 's per-period payoff is given by

$$\pi_{it} = p_{it} \left( \frac{1-\gamma}{NL} + \gamma d_t^i(p_{it}, p_{-it}, \iota_t) \right), \quad (3.6)$$

where  $p_{it}$  is the price;  $\frac{1-\gamma}{NL}$  is the quantity sold to its royals;  $d_t^i(p_{it}, p_{-it}, \iota_t)$  is the synchronized purchase decision of all shoppers, as a function of its price  $p_{it}$ , rivals' prices  $p_{-it}$ , and consumer inventory  $\iota_t$ ; and  $\gamma d_t^i$  is the quantity demanded of retailer  $i$  (assuming all shoppers' inventories are the same at the starting point, their equilibrium purchase decisions will synchronize as shown later). It is straightforward that  $d_t^i = 1$  only if  $p_{it} = p_{t,min}$ , given that shoppers buy at all.

I seek for a subgame-perfect MPE that satisfies several conditions. First, I assume that the current inventory levels, shoppers' and royals', are sufficient for the decision making of all agents. Second, the current decision making does not depend on states or actions taken in previous periods. Said differently, retailers' behavior is predicted by the current shopper inventory only.

**Bounding the state space.** Since the number of states in a MPE must be finite, we would like to have an upper bound of the inventory level upon which stockpiling behavior is endogenously irrational (*Assumption 3*). For simplicity, such upper bound is assumed to

be less than the size of one package,  $K < L$ . Thus, the highest possible inventory level is  $K - 1 + L$ , and the number of states is  $K + L$ .<sup>1</sup>

**Critical prices.** The MPE is characterized by a series of critical prices,  $\{p_k\}_{k=0}^K$ , where  $p_k \leq p^R$ . By *Assumption 2*, the critical price at  $k = 0$  is simply  $p^R$ . In each state  $k > 0$ , shoppers will purchase, if the lowest price offered is no greater than  $p_k$ , and not purchase otherwise, i.e.,  $\sum_i d_t^i(P_t, \iota_t = k) = 1$  if  $p_t \leq p_k$ .

### 3.3.1 Pricing Strategy

The retailers in the model will randomize prices, because it always pays to break ties, provided shoppers will purchase. Denote the distribution of prices in state  $k$  by  $F_k$ . Since the pricing strategy is state-dependent, for ease of notation I drop the subscript  $t$ . First notice that retailers will not choose any price greater than  $p^R$  because such prices produces zero sales. In the event that shoppers will purchase, it would be profitable to slightly undercut other retailers; such undercutting will not result in Bertrand consequence because it is more profitable to charge simply  $p^R$ .

Next, consider retailer  $i$ 's pricing strategy at state  $k$ . If  $p_i$  is greater than the critical price, then retailer  $i$  sells to its loyal consumers only in that period; the next period state depends on other retailer's prices: if there exists at least one retailer that offers a price no greater than  $p_k$ , which occurs with probability  $1 - (1 - F_k(p_k))^{N-1}$ , then the state transfers to  $k - 1 + L$ ; if no retailers offer such price, then the state transfers to  $k - 1$  with probability  $(1 - F_k(p_k))^{N-1}$ . If  $p_i$  is no greater than  $p_k$  and happens to be the lowest price among all retailers with probability  $(1 - F_k(p_i))^{N-1}$ , then it sells to both loyals and shoppers, and next period state transfers to  $k - 1 + L$  with certainty. If  $p_i$  is not the lowest price though lower than  $p^k$ , then it profits from loyals only.

From the retail's problem in equation (3.5), the value function of a retailer  $i$  is given by

$$W^k \equiv \max_{\{p_{it}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \delta^t \pi_i(p_{it}, P_{-it}, \iota_t = k). \quad (3.7)$$

Retailer  $i$ 's choice-specific value function in state  $k \neq 0$  can be written as

---

<sup>1</sup>The upper bound could also be, for example,  $L < K < 2L$ , but it implicitly allows purchase with storage, which complicates the transition paths of states and the calculation of continuation value.

$$W_i(p, p_{-i}, \iota = k) = \begin{cases} p \frac{(1-\gamma)}{NL} + \delta W^{k-1+L} \times (1 - (1 - F_k(p_k))^{N-1}) \\ + \delta W^{k-1} \times (1 - F_k(p_k))^{N-1}, \text{ if } p > p_k, \\ \\ p \frac{(1-\gamma)}{NL} + p\gamma \times (1 - F_k(p))^{N-1} \\ + \delta W^{k-1+L}, \text{ if } p \leq p_k. \end{cases} \quad (3.8)$$

The mixed-strategy equilibrium requires that a retailer makes equal profits at any price drawn from  $F_k$ . This means that, when shoppers hold inventory, no retailer would charge a price in the interval of  $(p_k, p^R)$ : a price slightly less than  $p^R$  will induce no greater sales, therefore there will be no loss from raising price to  $p^R$ . In other words, the probability of drawing  $p^R$  at state  $k > 0$  is strictly positive.

The pricing strategy in state  $k = 0$  is slightly different. The reason is that, according to *Assumption 2*, when inventory drops to zero all shoppers must purchase at any price no greater than  $p^R$ , and the subsequent state will be  $L-1$  with certainty. The price distribution from which retailers randomize their prices will be continuous, since any price lower than  $p^R$  would attract shoppers with positive probability. The choice-specific value function of retailer  $i$  is

$$W_i(p, p_{-i}, \iota = 0) = p \frac{(1-\gamma)}{NL} + p\gamma(1 - F_0(p))^{N-1} + \delta W^{L-1}. \quad (3.9)$$

**Lemma 1**

For all  $k \geq 0$ ,

$$W^k = \frac{p^R}{1-\delta} \frac{(1-\gamma)}{NL}. \quad (3.10)$$

*Proof.*

Because a retailer randomizes prices, then undercutting other retailers must bring equal profit as charging  $p^R$ . Thus,

$$\begin{aligned} W(k, p^R) &= p^R \frac{(1-\gamma)}{NL} + \delta W^{k-1+L} (1 - (1 - F^k(p^k))^{N-1}) \\ &\quad + \delta W^{k-1} (1 - F^k(p^k))^{N-1} \\ &= W(I_t = k, p_{it} \leq p_k) \\ &= p \frac{(1-\gamma)}{NL} + p\gamma(1 - F_k(p))^{N-1} + \delta W^{k-1+L}. \end{aligned} \quad (3.11)$$

Denote  $(1 - F^k(p^k))^{N-1}$  by  $prob_k$ . Then,

$$\begin{aligned}
W^0(0, p < p^R) &= W_0(0, p^R) = p^R \frac{(1-\gamma)}{NL} + \delta W_{-1+L} \\
W^1 &= p^R \frac{(1-\gamma)}{NL} + \delta W_L(1 - prob_1) + \delta W_0 prob_1 \\
&\vdots \\
W_{L-1} &= p^R \frac{(1-\gamma)}{NL} + \delta W_{2L-2}(1 - prob_{L-1}) + \delta W_{L-2} prob_{L-1}.
\end{aligned} \tag{3.12}$$

So,

$$\begin{aligned}
W^0 &= p^R \frac{(1-\gamma)}{NL} + \delta p^R \frac{(1-\gamma)}{NL} + \delta(\delta W^{2L-2}(1 - prob_{L-1}) + \delta W_{L-1} prob_{L-1}) \\
&= p^R \frac{(1-\gamma)}{NL} + \delta p^R \frac{(1-\gamma)}{NL} + \delta^2 p^R \frac{(1-\gamma)}{NL} + \dots \\
&= p^R \frac{1}{1-\delta} \frac{(1-\gamma)}{NL}.
\end{aligned} \tag{3.13}$$

Similarly, it can be sequentially shown that the total payoff of a retailer in all states is equal to the payoff as if she charged one single price  $p^R$  in all periods.

$$\begin{aligned}
W^k &= p^R \frac{(1-\gamma)}{NL} + \delta W^{k-1+L}(1 - (1 - F^k(p^k))^{N-1}) + \delta W^{k-1}(1 - F^k(p^k))^{N-1} \\
&= p^R \frac{(1-\gamma)}{NL} + \delta(p^R \frac{(1-\gamma)}{NL} + \delta W^{k-2+2}(1 - (1 - F^k(p^k))^{N-1}) \\
&\quad + \delta W^{k-2+L}(1 - F^k(p^k))^{N-1}) \\
&= \dots \\
&= p^R \frac{1}{1-\delta} \frac{(1-\gamma)}{NL}.
\end{aligned} \tag{3.14}$$

**Lemma 1** permits the direct calculation of price distribution in state  $k$ :

$$(1 - F_0(p))^{N-1} p \gamma = \frac{1-\gamma}{NL} (p^R - p), \text{ for } \underline{P}_0 \leq p \leq p^R, \tag{3.15}$$

and

$$\begin{cases} (1 - F_k(p))^{N-1} p \gamma = \frac{1-\gamma}{NL} (p^R - p), \text{ for } \underline{P}_k \leq p \leq p_k. \\ F_k(p) = 1 - F_k(p_k), \text{ for } p = p^R \end{cases} \tag{3.16}$$

where  $\underline{P}_0$  and  $\underline{P}_k$  are the lower bounds of the support of price distribution, satisfying  $F_0(\underline{P}_0) = 0$  and  $F_k(\underline{P}_k) = 0$ , respectively. It is clear that the supports of price distribution for different  $k$  have the same lower bound:

$$\underline{P} = \underline{P}_k = p^R \frac{1-\gamma}{1-\gamma + NL\gamma}, \forall 0 \leq k \leq K. \tag{3.17}$$

While  $F_0$  is continuous on its support,  $F_k$  has a mass point on  $p^R$ . When  $k > 0$ , the probability of choosing  $p^R$  is strictly positive and is given by  $1 - F_k(p_k)$ . This difference is due to the increment in sales when slightly reducing price from  $p^R$ : since shoppers must purchase at  $k = 0$ , a price slightly lower than  $p^R$  results in a positive probability of sales. In contrast, at  $k > 0$ , there is no gain in sales if the reduced price is not as low as  $p_k$ .

A retailer's equilibrium pricing strategy is that, when shoppers' inventory is zero,  $p \sim F_0$ ; when shoppers hold inventory,  $0 < k \leq K$ ,  $p \sim F_k$ ; when  $k > K$ , according to *Assumption 3*,  $p = p^R$  with probability one.

### 3.3.2 Purchase Strategy

A rational shopper takes into account the current lowest price  $p_{t,min}$ , her own next period inventory  $I_{jt+1}$ , and future state  $\iota_{t+1}$ . Buying today at a low price means the inventory cost will immediately occur, while postponing purchase will risk losing the good deal but avoiding this cost.

The purchase strategy is characterized by a critical price  $p_k$  at each state. Given retailers' symmetric pricing strategy, suppose in period  $t$  at state  $k$  the lowest price available is  $p_{t,min} \leq p_k$ . Shoppers will purchase a new package for inventory, and the state will transit into  $\iota_{t+1} = k - 1 + L$ . By the assumption that  $K < L$  is the highest state where sales prices could occur, the succeeding states will be  $k - 2 + L, \dots, K, \dots$ , and prices will stay at  $p^R$  for  $k + L - K$  periods until  $\iota = K$  where the next sale would occur.

Now consider a potential deviator who does not purchase at  $p_{t,min} \leq p_k$ . According to *Assumption 2* and *3*, she is certain about the states and price distribution in the subsequent  $K + L - k$  periods. Because all prices will be staying at  $p^R$  for at least  $k + L - K$  periods, and the deviator will run out in  $k < k + L - K$  periods, she will have to pay  $p^R$  for the next package. Thus, she would rather postpone the next purchase as late as possible. Said differently, if postpone purchase, then it will be postponed for  $k$  periods. After this very late purchase, her inventory will again synchronize with the rest.

In order to prevent deviation,  $p_{t,min}$  must be small enough such that buying today is no worse than postponing it. Recall that  $V(P_t, I_{j,t}, \iota_{t+1})$  is the value function which can denote the continuation value of a potential deviator  $j$ , whose inventory is  $I_{j,t}$  at the beginning

of period  $t$  at state  $\iota_{t+1}$ . Expressing the deviator's tradeoff using the choice-specific value functions, we obtain the upper bound of  $p_{t,min}$ .

Suppose  $p_{t,min} \leq p_k$ , then  $p_k$  satisfies

$$\tilde{V}(P_t, k, k, 1) \geq \tilde{V}(P_t, k, k, 0) \quad (3.18)$$

$\Downarrow$

$$p_{t,min} \leq -cL + \delta \left( V(P_{t+1}, k-1+L, k-1+L) - V(P_{t+1}, k-1, k-1+L) \right).$$

$V(P_{t+1}, k-1+L, k-1+L)$  is the expected continuation value of a non-deviator with inventory  $k-1+L$ , where the expectation is taken with respect to  $P_{t+1}$  in which all  $p_{i,t+1} \sim F_{k-1+L}$ . Her inventory will be  $k-1+L$  after this purchase and gradually declined to  $K$  along with consumption. She needs not consider her purchase decision until the next sale, which occurs after  $k-1+L-K$  periods. Thus, it can be written in the sum of the total inventory costs and total consumption utilities in the next  $k-1+L-K$  periods, and the expected value at state  $K$ .

$$\begin{aligned} & V(P_{t+1}, k-1+L, k-1+L) \\ &= \sum_{i=1}^{k-1+L-K} -c(k-1+L-i)\delta^{i-1} + v \left( \sum_{i=1}^{k-1+L-K} \delta^{i-1} \right) \\ & \quad + \delta^{k-1+L-K} V(P, K, K). \end{aligned} \quad (3.19)$$

Similarly,  $V(P_{t+1}, k-1, k-1+L)$ , the deviator's continuation value, can be written as the sum of cumulative inventory costs and consumption utilities in the next  $k-1$  periods, and the continuation value when she runs out of stock and is forced to purchase at  $p^R$ . The latter continuation value can be decomposed into three parts, consumption utilities and inventory costs before inventory declines to  $K$ , and the discounted continuation value at state  $K$  with inventory  $K$  (see **Appendix A** for details):

$$\begin{aligned} & V(P_{t+1}, k-1, k-1+L) \\ &= \sum_{i=1}^{k-1} -c(k-1-i)\delta^{i-1} + \sum_{i=0}^{L-1+K} (-c(L-1-i))\delta^{k-1+i} \\ & \quad + v \left( \sum_{i=1}^{k-1+L-K} \delta^{i-1} \right) + \delta^{k-1+L-K} V(P, K, K) \end{aligned} \quad (3.20)$$

Not surprisingly, the two terms in the parenthesis of inequality (3.18), the continuation values of a non-deviator and a deviator, contain a common component,  $\delta^{k-1+L-K}V(P, K, K)$ . This is consistent with the fact that the deviator's inventory will synchronize with the rest after  $k - 1 + L - K$  periods. The infinite-horizon optimizing problem boils down to a  $k + L - K$ -period problem. Plugging the above two expressions back, we obtain

$$p_{t,min} \leq -cL \left( \sum_{i=0}^{k-1} \delta^i \right) + \delta^k p^R = -cL \frac{1 - \delta^k}{1 - \delta} + \delta^k p^R \quad (3.21)$$

The deviator postpones her purchase till  $k$  periods later and pays  $p^R$ , while the rest of shoppers purchase now at  $p_{t,min}$ , but have to store the new package for  $k$  periods. Therefore, there would be no incentive to deviate, if the cumulated inventory costs of the new package do not offset the difference between  $\delta^k p^R$  and  $p_{t,min}$ . To rationalize shoppers' purchase behavior, the critical price is

$$p_k = -cL \frac{1 - \delta^k}{1 - \delta} + \delta^k p^R. \quad (3.22)$$

There are two things worth noticing about the critical price. First, the value of the right hand side of the above equation is the payoff of purchasing a new package at the regular price when the current package is ran out ( $k$  periods later), a reservation value that a shopper can always get by postponing purchase. Second, the critical price monotonically decreases on  $k$ ; when shoppers possess higher inventory, the price that induce purchase has to be lower.

If  $p_{t,min} > p_k$ , or  $p_{it} = p^R$ , for all  $i$ , the rational decision is "not buy". This is a trivial case. The deviator cannot gain by making a purchase. First notice that there would be sales held in the subsequent periods as long as no retailer offers a sale price in any of the previous periods, during which the inventory of all shoppers declines by one in each period. If she buy in  $t$  at  $p^R$ , the new package will incur inventory cost, and, if there would be a sale in the subsequent periods, she would have to miss it.

### 3.4 Equilibrium

First of all, an equilibrium series of critical prices  $\{p_k\}_{k=0}^K$  ( $p_0$  is trivially  $p^R$ ) satisfies

- $F_k(p_k) > 0$ , which is equivalent to  $\underline{P} < p_k, \forall 1 \leq k \leq K$ .

Recall that  $p_k$  monotonically decreases on  $k$  and the lower bounds of price supports are  $\underline{P}$ , the above condition will be violated when  $k$  is sufficiently big. This idea is used to rationalize *Assumption 3*. If the highest possible inventory level at which the probability of sales is strictly positive is  $K$ , the above condition implies

$$\begin{aligned} p^R \frac{1-\gamma}{1-\gamma+NL\gamma} &< -cL \frac{1-\delta^k}{1-\delta} + \delta^k p^R, \text{ for } 0 \leq k \leq K, \\ p^R \frac{1-\gamma}{1-\gamma+NL\gamma} &> -cL \frac{1-\delta^k}{1-\delta} + \delta^k p^R, \text{ for } K+1 \leq k. \end{aligned} \quad (3.23)$$

Since the series of critical prices monotonically decreases on  $k$  and the starting value  $p_0 = p^R > \underline{P}$  and the limit of the series  $\lim_{k \rightarrow \infty} p_k = \frac{cL}{1-\delta} < \underline{P}$ , there exists an integer  $K > 0$  that satisfies the above inequalities. To rationalize *Assumption 3*, we would like the parameters satisfy  $K < L$ . A sufficient and necessary parameter condition is

$$p^R \frac{1-\gamma}{1-\gamma+NL\gamma} > -cL \frac{1-\delta^L}{1-\delta} + \delta^L p^R, \quad (3.24)$$

which can be satisfied as long as  $\delta$  is sufficiently small.

In the states where  $k > K$ , the critical price is lower than  $\underline{P}$ , all retailers will charge  $p^R$  with probability one at those states. A number of observations immediately follow.

1. Because  $p_k$  decreases on  $k$  and each  $F_k$  is identical to  $F_0$  within the interval  $[\underline{P}, p_k]$ ,  $F_k(p_k)$  also decreases on  $k$ . This means that the probability of holding a sale is low at high inventories. Though retailers have to offer lower prices in order to induce purchase at high inventories, the prices are more likely to be at the regular level. The oligopoly competition is the most rigorous at state zero, where all retailers offer sales prices.

2. Given observation 1, the transition of states is non-absorbing, in a sense that high states are more likely to decrease than low states, and low states ( $\leq K$ ) are more likely to jump up by  $L-1$  than high states.

The MPE requires that the decision making depends on the current state only. One would argue that the history affects agents' expectation, because, for example, a shopper anticipates a sale occurs with high probability if there has been no sales in the past long period of

time. This intuitive hazard rate of holding a sale is actually consistent with our equilibrium outcome: at a low state, which implicates that no sales have occurred in the past several periods and therefore no purchase has been made by shoppers, the probability of holding a sale is high.

3. The realized price evolution consists of several consecutive periods where price stays at  $p^R$  followed by temporary one-time price reductions. During the periods where inventory is low enough for retailers to offer sales, either one or more retailers simultaneously hold sales, or none of them do so. If any one of the stores offers a sale price, the Low type consumers will stock up, and this sale will be followed by another several consecutive non-sale periods.

Suppose the system starts from a "low" state with  $k < K$  and the lowest price  $p_t$  is less than the critical level, purchase by shoppers will occur and the state transits to  $k - 1 + L$  where all prices will be  $p^R$ . The price will stay at  $p^R$  for at least  $k - 1 + L - K$  periods until the state falls back to  $K$ . The next sale will occur within  $K$  periods, at a state  $0 \leq k \leq K$ . This sale is again followed by several consecutive periods in which all prices stay at  $p^R$ . Since an equilibrium sale must induce purchase resulting in inventory levels higher than  $K$ , the sales are one-time and temporary.

### 3.5 Conclusion

I construct a model of oligopoly retailers selling a homogeneous storable good. The good in the model are assumed to be consumed for multiple periods, and therefore does not need purchased frequently. Under the infinite horizon setup, the High type consumers are assumed to be loyal to a specific store, never purchase for inventory cost, and do not search for price; they purchase only when run out at the store she is loyal to. The Low type consumers do search for price at zero cost, purchase if the best deal price offered among all stores is below some cut-off level, and store it as inventory at some cost.

One equilibrium among the continuum of equilibria is characterized by a critical price that decreases on inventory. At each state, retailers simultaneously draw prices from  $F_k$ ,

and shoppers purchase if the lowest price is no greater than the critical price. The critical price and the probability of holding a sale are low at high inventories. The pattern of state transition is non-absorbing: high states are more likely to decrease by one, and low states ( $\leq K$ ) are likely to jump up by  $L - 1$ .

This model assumes consumer heterogeneities with respect to store loyalty, search cost, and inventory cost. The predicted patterns of price variation is consistent with empirical observation that there is a strictly positive probability of charging the regular price. The model also predicts that the probability of holding a sale is low when shoppers hold a high inventory; and no shopper would store more than one packages, apart from the package that is currently in use. The predicted price panel consists of several consecutive periods where all prices stay at the regular level and occasional one-period price reductions.

The model relies on the assumption that the inventories and therefore purchase behaviors of all shoppers are synchronized. Relaxing this assumption would result in a complex of equilibria in which only some shoppers whose inventories below some cut-off level would purchase.

## Chapter 4

### An Empirical Analysis of Dynamic Supply and Demand of Storable Goods

#### 4.1 Introduction

When goods are storable, consumers take future prices into consideration while making current purchase decision, and suppliers take into account future profit when making current pricing decision. A consumer has an incentive to make an unplanned purchase if she observes a price-cut and believes that it will return to the regular price level in the near future. The unexpectedly purchased product goes to the storage at an inventory cost. Thus, the consumer faces a trade-off between storage cost and attractive low price. On the other hand, the seller's pricing decision also faces a dynamic trade-off: selling more today reduces demand tomorrow. Suppliers tend to keep the prices high to preserve future demand. Moreover, if adjusting prices incurs menu costs, it is optimal to cut prices only if the costs can be covered by the increase in total expected profit brought by expanded sales. Similarly, changing prices from sales prices back to regular levels is optimal only if this menu cost can be covered by the gain in profits attributes to preserved future demand.

Dynamic markets that are frequently modelled are of durable goods. Similar to product replacement in a durable-good market that is driven by innovation and obsolescence, repeated purchases in a storable-good market are forced by continuous consumption. Despite the importance of dynamic demand and pricing decision in differentiated goods market, the equilibrium implications of stores' and consumers' strategies remain unclear. I therefore construct a model in the framework of Markov-perfect equilibrium (MPE) of dynamic demand and supply of storable goods with endogenous optimal strategies and consistent beliefs .

In this model, products are differentiated, and are sold by a store. Consumers maximize

the expected discounted future utility flows, accounting for the store's strategies lead to future prices. In each period a consumer decides whether to buy, how much to buy, which product to buy, and how much to consume. Quantities not consumed are stored at a cost, and can be consumed in the future. Observing prices, consumers balance inventory cost and potential future savings. Store maximizes the present expected future profit flows by choosing prices of all products it carries while taking into account the dynamic stockpiling behavior of consumers. In each period, a menu cost incurs if the store chooses a price different from the last period. A store balances the current menu cost, current profit, and future demand. Since each consumer's demand depends on how much inventory she holds, the distribution of currently owned products affects aggregate demand in each period.

I estimate the model with data of laundry detergent market. This industry is well-suited for the analysis as there are only five typical package sizes that largely decrease the computational burden. Despite of this, several adjustments need to be made: first, neither inventory of a specific consumer nor the inventory distribution, the state variables of consumers and store, is observed. To address this problem, I generate an initial distribution, and updated it by the optimal endogenous purchase and consumption policies. To reduce the dimensionality (there are 29 brands and 5 package sizes, though not all brands have all 5 sizes), I assume that the product characteristics enter consumer's value function only in the current period and will not affect future utility, and that the consumption decision is not product specific. Thus, product differentiation is irrelevant to the dynamic tradeoffs, and all related parameters can be estimated using a static model. The structural parameters related to consumption and inventory cost are estimated in a much smaller space by solving the consumer's dynamic problem, considering only the quantity decisions.

The dimensionality problem is potentially remarkable for the supply side. With continuous bounded action sets of prices of a large number of products, finding the optimal prices would be extremely computational burdensome. The price series of each product from the data tells me the pattern of pricing decision: each product has a finite discrete set of price, and the price flow is jumping over the several candidate prices period to period. I therefore assume that the price sets are exogenous, in a sense that the number of candidates and values in each set are not determined by the dynamic optimizing problem.

The structural parameters of the model are estimated with SMD estimator. The optimal estimates are found by minimizing the distance between a set of moments of the data and their counterparts simulated by the model. I choose the price elasticities of non-purchase duration and quantity sold, average consumption rate, average market shares and prices, and average price adjustment intervals as matching moments. Each simulation consists of 104 periods, the same length as the observed data. In each period, the optimal strategies of store and consumers are solved numerically at given states, and the state variables are updated under the optimal policies.

The structural estimates allow me to simulate the market evolution to examine optimal behavior and belief. Store decides to hold big sales by cutting prices of a large number of products when the average inventory of consumers is high. After the big sales, store keeps prices at regular levels, along with the gradual decline in average inventory level that is primarily due to consumption and small purchases. Store does not hold the next big sale until the average inventory recovers to high level. Store does not cut prices when the average inventory level is low, because a rise in demand will be driven by consumption needs instead of low prices. At extremely low average inventory, a purchase spike occurs, even the prices are mostly at regular levels. Inventory level recovers to high level after the purchase spike. Furthermore, the agents' beliefs on inventory distribution are consistent with evolution of inventory distribution that is governed by the optimal strategies under those beliefs.

## 4.2 Literature Review

The empirical framework of Markov-perfect dynamic industry (Maskin and Tirole 1988) was first developed by Ericson and Pakes (1995) for oligopoly evolution. Pakes and McGuire (1992) propose an algorithm for computing the MPE. At each tempted parameter vector, MPE needs to be solved by analytically solving value functions using polynomial approximation. Bajari et al. (2007) describe a two-step algorithm for estimating dynamic games under the assumption that observed actions are consistent with MPE. The advantage is that in the first step of this method, the agents' policy functions can be fully recovered nonparametrically from observed actions and states, as well as the Markov transition kernel

that determine the evolution of relevant state variables. This essentially involves regressing observed actions to observed state variables. The second step is to find the structural parameters that rationalize the observed actions as a set of optimal response. Aguirregabiria and Mira (2007) point out that because the two-step method requires consistent nonparametric estimators of players' choice probabilities in the first step, it is sometimes unavailable, biased in small samples, and asymptotically inefficient. They therefore propose nested pseudo likelihood estimator for unavailable nonparametric estimation of choice probabilities and unobserved heterogeneity.

The empirical framework of Ericson and Pakes (1995) has been applied to a variety of industries with entry/exit and/or R&D. Pakes et al. (2007) use the two-step method to estimate industry dynamics with entry and exit. The study of learning-by-doing of Benkard (2004) allows the market size for airplanes to change stochastically based on current sales to mimic the dynamic implications of forward-looking consumers. Another application of the empirical framework lies in durable goods market, where forward-looking consumers decide when to update a product and firms choose optimal innovation investments and prices. Carranza (2010) estimates product innovation and adoption of digital cameras, using a reduced-form solution of dynamic demand that dramatically facilitates estimation. In their empirical study on PC microprocessor industry, Goettler and Gordon (2008) propose a model of dynamic oligopoly with durable goods taking into account the dynamic behavior of consumers who rationally anticipate future product improvements and price declines.

Consumer dynamic behaviors in durable and storable goods markets are very similar: a consumer make purchase decision in each period taking into account the prices and product attributes (qualities, improvements, etc.) of available products in the market, predictions on future prices and attributes, and the current product owned by her. For durable goods market the purchase decision depend on the attributes of product she owns, while for storable goods market this decision is determined by the amount of products in storage. Hendel and Nevo (2006) examine consumer stockpiling behavior that results from contemporary price reduction of storable goods. They show that demand estimation based on temporary prices are mismeasured when long-run response to prices are ignored. Nair (2007) studied inter-temporal price discrimination with forward-looking consumers in the market of

console video-games. His model simulation reveals that ignoring the dynamic behavior of consumers results in large profit losses.

This paper nests the dynamic demand model developed by Hendel and Nevo (2006). For computation algorithm, the popular two-step method famous for its simplicity cannot be used here because the state variables and some of the optimal responses are unobserved by the econometrician. The state variables, individual inventory of the storable good and inventory distribution are not recorded. Consumption, one of the actions chosen by consumers, is also unobservable. Therefore, I propose a hybrid algorithm that combines 'traditional' equilibrium solving and SMD estimator for parameter search.

### 4.3 Model

Time is discrete with an infinite horizon, indexed by  $t$ . A monopolistic store sells products  $jx$ , where  $j \in \{1, \dots, J\}$  denotes brand,  $x \in \{0, x_1, \dots, x_X\}$  denotes the index of the package size (0 for non-purchase). In each period, the store make price decision  $p_t = (\dots, p_{jxt}, \dots)'$ . Price is a dynamic control, since lowering price in period  $t$  increases current profits but reduces future demand. Moreover, adjusting the price of a specific product will incur a menu cost. Observing the prices, consumers decide in each period whether to buy or to continue using the product stored, and how much to consume. Goods that is not consumed will be stored at a cost. Because consumer purchase decisions depend on the inventory hold by each consumer, the aggregate demand is determined by the distribution of inventory. Denote the distribution of consumer inventory by  $D_t$ . The store and consumers are forward-looking and take into account the optimal dynamic behavior of other agents when making their respective decisions. I assume that the distribution of inventory is observable for all agents.

#### 4.3.1 Consumer Purchase Behavior

Each consumer  $h \in \{1, \dots, H\}$  in every period  $t$  decides whether to buy, which brand and which package size to buy, and how much to consume. The good that is not consumed in

period  $t$  is stored as inventory at a cost, and can be consumed in the future. Denote a consumer's decision  $\sigma_t = \{j, x, c\} \in \{1, \dots, J\} \times \{0, \dots, x_X\} \times R^+$ , where  $c$  denotes consumption,  $x = 0$  stands for no purchase. Consumer  $h$  obtains per-period utility consists of the utility from consumption, the disutility from inventory, and a one-time payoff from purchasing the product:

$$U_{ht} = u_C(c_{ht}; \Theta_C) - u_I(i_{h,t+1}; \Theta_I) + \sum_{j,x} d_{h j x t} u_{h j x t} \quad (4.1)$$

where

$$u_{h j x t} = \alpha p_{h j x t} + \xi_{h j x} + \epsilon_{h j x t} \quad (4.2)$$

$$i_{h,t+1} = i_{ht} + x_{ht} - c_{ht} \quad (4.3)$$

where  $d_{h j x t}$  is the purchase indicator, satisfying

$$d_{h j x t} \in \{0, 1\}$$

$$\sum_{j,x} d_{j x t} = 1$$

$\xi_{h j x}$  is the time-invariant idiosyncratic taste that could be a function of characteristics of product  $jx$ , and  $\epsilon_{h j x t}$  is a individual and product-specific random shock to consumer choice. The subscript  $h$  is omitted hereafter.

Each consumer maximizes her expected discounted utility, which can be formulated using Bellman's equation:

$$V(D_t, i_t, \epsilon_t) = \max_{c_t, d_t} \left\{ U_t + \delta \sum_{D_{t+1}|D_t} \int_{\epsilon} V(D_{t+1}, i_{t+1}, \epsilon_{t+1}) d\epsilon_{t+1} g(D_{t+1}|D_t, \epsilon_t, c_t, d_t) \right\} \quad (4.4)$$

subject to  $c_t \leq i_t + x_t$ ,

$$\sum_{j,x} d_{j x t} = 1,$$

where  $g(D_{t+1}|D_t, \epsilon_t, c_t, d_t)$  is the distribution of inventory of the next period conditional on current state variables and choices. Following Rust (1987) and assuming  $\epsilon_{j x t}$  is i.i.d. extreme value type-1 across all consumers and over time, the computational burden can be significantly reduced. The standard multinomial logit formula for the demand side can be obtained by integrating over all future  $\epsilon_{j x s}, s > t$ . in particular, the product-specific value

function is

$$\begin{aligned} \widehat{V}_{jx}(D_t, i_t) = & \max_{c_t} u_C(c_t; \Theta_C) - u_I(i_{t+1}; \Theta_I) + \alpha p_{jxt} + \xi_{jx} \\ & + \delta \sum_{D_{t+1}|D_t} \log \left( \sum_{j', x'} (\exp(V(D_{t+1}, i_{t+1}))) g(D_{t+1}|D_t, \epsilon_t) \right) \end{aligned} \quad (4.5)$$

The conditional choice probabilities for a consumer with inventory  $i_t$  are therefore

$$s_{jxt|i_t} = \frac{\exp(\widehat{V}_{jx}(D_t, i_t))}{\sum_{k,y} \exp(\widehat{V}_{ky}(D_t, i_t))} \quad (4.6)$$

The market share of product  $jx$  is

$$s_{jt} = \int s_{jxt|i_t} dD_t(i_t) \quad (4.7)$$

Once a consumer purchases a product of size  $x$ , the brand or any characteristic of the product no longer matters. The consumer receives a one-time utility payoff of  $\xi_{jx}$  from purchasing product  $jx$ . This payoff dose not occur in future periods because the consumption and purchase decisions depend only on inventory. With products of different brands serving as perfect substitutes in her inventory, neither  $c_t$  nor  $i_t$  is product specific. That is, prices, merchandising activities, and product characteristics affect purchase decision at store, while none of them affect consumption decision conditional on a purchased size. Relaxing this assumption would require  $D_t$  to be product specific, which substantially increase the state space. Hendel and Nevo (2006) show that the parameters related to product differentiation are independent from the dynamic quantity choice, and the probability of choosing a brand conditional on quantity dose not depend on dynamic considerations. Therefore,  $\alpha, \xi_{jx}$  can be estimated using a static logit model.

Each consumer is small relative to the size of the market so that individual actions do not affect the evolution of inventory distribution. In other words, when a consumer is making her own decision, she dose not consider the trade-off between shifts in  $D_{t+1}|D_t$  and her quantity choice.

#### 4.3.2 Store Pricing Behavior

Each period the store makes pricing and promotion decisions. The store can keep the price of each product unchanged as in the last period, or raise/cut price at a menu cost. The

flow profit function can be written as:

$$\begin{aligned} \pi_t(p_t, p_{t-1}, D_t) = & M \sum_{j,x} s_{jxt}(p_t, D_t)(p_{jxt} - mc_{jx}) \\ & - \gamma \sum_{j,x} I(p_{jxt} \neq p_{jx,t-1}), \end{aligned} \quad (4.8)$$

where  $M$  is the market size,  $s_{jxt}$  is market share as a function of pricing and promotion decisions and current inventory distribution,  $mc_{jx}$  is the input price, and  $\gamma \sum_{j,x} I(p_{jxt} \neq p_{jx,t-1})$  represents the total menu cost, in which  $I(\cdot)$  is an indicator function. A unit menu cost of  $\gamma$  incurs if a price is adjusted from the last period. The store maximizes the expected discounted profits, which can be expressed by Bellman equation

$$W(D_t, p_{t-1}) = \max_{p_t} \left\{ \pi_t(p_t, p_{t-1}, D_t) + \delta \sum_{D_{t+1}|D_t, p_t} W(D_{t+1}, p_t) f(D_{t+1}|D_t, p_t) \right\} \quad (4.9)$$

The market shares determined by consumer optimization translate into the law of motion for the distribution of inventory. The share of consumers with inventory  $i$  at the beginning of period  $t + 1$  is

$$D_{t+1}(i)|D_t, p_t, p_{t-1} = s_{0t} \cdot D_t(i) + \sum_{j,x \neq 0} s_{jxt}(p_t, p_{t-1}) \cdot D_t(i = i_t + x - c_t) \quad (4.10)$$

The first part of  $D_{t+1}(i)$  accounts for the probability of non-purchase of a consumer with inventory  $i$ . If a consumer with inventory  $i_t$  in the current period purchase product  $jx$  at probability  $s_{jxt}$ , updating her next-period inventory to be  $i$ , then  $D_{t+1}$  shifts. The continuation inventory distribution is a deterministic function of prices and merchandising activities.

### 4.3.3 Store's Dynamic Tradeoffs

Although the action sets of price and merchandising activities are not continuous, differentiating value function can gives a feeling about store's dynamic trade-offs. Given a last-period price  $p_{t-1}$  and inventory distribution  $D_t$ , consider the optimal price of product  $ky$ , at which

$$\frac{\partial W(\cdot)}{\partial p_{kyt}} = 0:$$

$$\begin{aligned} \frac{\partial W(\cdot)}{\partial p_{kyt}} &= \frac{\partial \pi_t(p_t, p_{t-1}, D_t)}{\partial p_{kyt}} + \delta \sum_{D_{t+1}|D_t, p_t} W(D_{t+1}) \frac{\partial f(D_{t+1}|D_t, p_t)}{\partial p_{kyt}} \\ &= M \cdot \sum_{jx \neq ky} \frac{\partial s_{jxt}}{\partial p_{kyt}} (p_{jxt} - mc_{jx}) + Ms_{kyt} - \gamma \frac{\partial I(\cdot)}{\partial p_{kyt}} \\ &\quad + \delta \sum_{D_{t+1}|D_t, p_t} W(D_{t+1}) \frac{\partial f(D_{t+1}, |D_t, p_t)}{\partial p_{jxt}} \\ &= 0 \end{aligned} \tag{4.11}$$

Similarly, a small decrease in price will bring a small current profit increase of this product, decrease in the sum of all other products, a menu cost (if different from the last period), and a small decrease in future total profit through a shift in next-period inventory distribution. Suppose the current change from a small deviation in  $p_{kyt}$  causes a loss in current profit:

$$M \cdot \sum_{jx \neq ky} \frac{\partial s_{jxt}}{\partial p_{kyt}} (p_{jxt} - mc_{jx}) + Ms_{kyt} - \gamma \frac{\partial I(\cdot)}{\partial p_{kyt}} < 0$$

then

$$\delta \sum_{D_{t+1}|D_t, p_t} W(D_{t+1}) \frac{\partial f(D_{t+1}, |D_t, p_t)}{\partial p_{jxt}}$$

is the gain in future profits. By contrast, a myopic seller who ignores the shift in future inventory distribution will choose a lower price that satisfies

$$\frac{\partial \pi_t(p_t, p_{t-1}, D_t)}{\partial p_{kyt}} = 0$$

for all  $ky$ . In sum, the dynamic trade-off of the seller means a higher price is chosen so that more consumer will likely to purchase in the next period. Because the action set of store is finite, there exists at least one  $p_t^* \in \Pi_{j,x} P_{jx}$  as a solution for store's problem (may not be unique).

#### 4.3.4 Equilibrium

I consider a mixed-strategy MPE of this dynamic game, based on the framework developed by Ericson and Pakes (1995). The inventory distribution is observable for both consumers and store. They also have consistent beliefs about how future distribution changes. The equilibrium is of rational expectation, in a sense that the evolution of the state variable

$(D_{t+1}|D_t$  in this application) governed by the optimal actions of all agents is consistent with their beliefs on future states. Both store and consumers choose their optimal policies based on expectations that accurately reflect the policies. Specifically,

1. The strategies of all agents,  $\{p_t, d_t, c_t\}_{t=0}^{\infty}$ , depend only on state variables;
2. Consumers possess rational expectations about store's policy functions and the evolution of inventory distribution;
3. Store possesses rational expectation about the evolution of inventory distribution.

Formally, I define an equilibrium for this dynamic game as the set

$$\{W(D), p^*, \{V(D, i), c^*, d^*\}_{h=1}^{h=H}, g(\cdot|\cdot), f(\cdot|\cdot)\}$$

which contains the value functions of the store and consumers, their optimal policies, and beliefs about future inventory distributions. Given their beliefs and at any  $(D_t, i_t, \epsilon_t) \in \mathfrak{D} \times \mathfrak{I} \times \mathfrak{E}$ , the policy functions maximize all agents' payoffs:

$$\begin{aligned} (p^*|D_t, p_{t-1}) &= \underset{(p)}{\operatorname{argmax}} \{ \pi_t + \delta E[W(D_{t+1}|D_t, p_t)] \} \\ (c^*, d^*|i_t, D_t) &= \underset{(c, d)}{\operatorname{argmax}} \{ U_t + \delta E[V(D_{t+1}|D_t, \epsilon_t)] \} \end{aligned} \quad (4.12)$$

The expectations are rational in that the expected distributions match the distribution from which realizations are drawn when consumers and store behave according to their policy functions. That is,

$$\begin{aligned} D_{t+1}|D_t, p_t^* &= s_{0t} \cdot D_t(i) + \sum_{j, x \neq 0} s_{jxt} \cdot D_t(i = i_t + x^* - c_t^*) \\ &= \int g(D_{t+1}|D_t, \epsilon_t, c_t^*, d_t^*) \operatorname{Prob}(c_t^*, d_t^*|\epsilon_t, D_t) dF_\epsilon \\ &= f(D_{t+1}|D_t, p_t^*) \operatorname{Prob}(p_t^*|D_t) \end{aligned} \quad (4.13)$$

I use the following assumption for the existence of MPE:

1. Each consider is small so that  $i_t$  is ignorable for  $D_t$ ;
2. There exists an initial inventory distribution,  $D_0$ , that is observed by all agents; all consumers know her own initial inventory,  $i_{ht}$ ;

3. Consumers are identical;
4.  $p_{jxt}$  and  $i_{ht}$  are positive and bounded:  $0 \leq i_{ht} \leq \bar{i}$ , for all  $h, t$ , and  $0 \leq p_{jxt} \leq \bar{p}$  for all  $j, x$ . Thus,  $D_t$  has an positive and bounded support.  $c_t$  is also bounded since  $c_t \leq i_t + x_t$ .

Note that for any set of strategies  $\sigma$ , in equilibrium or not, the value functions  $V(\cdot)$  and  $W(\cdot)$  and inventory distribution depend on players' strategies only through the choice probabilities  $P$  associated with  $\sigma$ . Let  $\sigma^* = (\sigma_c^*, \sigma_s^*)$  be equilibrium strategies of consumers and store, and  $P^*$  be the associated probabilities, then

$$\begin{aligned} P_h^*(c_t, d_t | i_t, D_t) &= \int I(\{c_t, d_t\} = \sigma_c^*(i_t, D_t, \epsilon_t)) dF_\epsilon \\ P_s^*(p_t | D_t, p_{t-1}) &= I(\{p_t\} = \sigma_s^*(D_t, p_{t-1})) \end{aligned} \quad (4.14)$$

where  $I$  is the indicator. Following the definition of MPE in the game,

$$\begin{aligned} P_h^*(c_t, d_t | i_t, D_t) &= \int I(\{c_t, d_t\} = \underset{(c,d)}{\operatorname{argmax}} \{U_t \\ &\quad + \delta \sum_{D_{t+1} | D_t, p_t} V(i_{t+1}, D_{t+1}) I(D_{t+1} | D_t, p_t) P^*(p_t | D_t, p_{t-1})\} dF_\epsilon \\ P_s^*(p_t | D_t, p_{t-1}) &= I(p_t = \underset{(p)}{\operatorname{argmax}} \{\pi_t \\ &\quad + \delta \sum_{D_{t+1} | D_t, p_t} W(i_{t+1}, D_{t+1}) I(D_{t+1} | D_t, \{c_{ht}, d_{ht}\}_{h=1}^H) \\ &\quad \prod_{h=1}^H P_h^*(c_t, d_t | i_t, D_t)\}) \end{aligned} \quad (4.15)$$

where  $P_h^*(c_t, d_t | i_t, D_t)$  is the marginal probability at  $i_t$ , and  $P^*(p_t | D_t, p_{t-1})$  can be written by  $\int P_h^*(c_t, d_t | i_t, D_t) dD(\cdot)$ . We finally have

$$P^*(\{c_{ht}, d_{ht}\}_{h=1}^H, p_t) = \Lambda(P^*(\{c_{ht}, d_{ht}\}_{h=1}^H, p_t)) \quad (4.16)$$

Therefore, the equilibrium probabilities are a fixed point. Given the assumptions on the distribution of  $\epsilon$ , best response probability function  $\Lambda$  is well-defined and continuous on the compact sets of players' choice probabilities. By Brower's Theorem, there exist at least one mixed-strategy MPE. In the next few sections, I assume there exist a pure-strategy MPE. Model simulation and parameter estimation are fulfilled for a pure-strategy MPE.

## 4.4 Computation

### 4.4.1 Approximation of Inventory Distribution

The challenge due to the inventory distribution as a state variable is that  $D_t$  is a high dimensional simplex. I approximate this continuous state variable with a finite set  $\{D_m\}_{m=1}^M$  that restrict  $D_t$ . Let  $\hat{D}_{t+1}$  be the unapproximated transition implied by equation(4.13) , and let  $\rho_m$  denote the distance between  $\hat{D}_{t+1}$  and  $D_m$ . Among the several candidates of distance metrics, the mean is the most relevant moment for logit-based demand systems (Goettler and Gordon 2008). Since the mean inventory is the main factor that drives purchase, consumption, and pricing behaviour. Thus, the distance metrics is written as

$$\rho_m = \left| \sum_i i \hat{D}(i)_{t+1} - \sum_i i D_m(i) \right| \quad (4.17)$$

where the summation is over the discrete grids from 0 to  $\bar{i}$ . With multiple dimensions, the probabilities of transforming to  $D_m$  from  $D_t$  can be defined as some function negatively related to  $\rho_m$ . To further simplify computation, I choose the two closest  $D_m$ , denoted by subscript  $m$  and  $m+1$ , to approximate  $\hat{D}_{t+1}$ :

$$\begin{aligned} Prob(D_{t+1} = D_m) &= \frac{\rho_{m+1}}{\rho_m + \rho_{m+1}} \\ Prob(D_{t+1} = D_{m+1}) &= \frac{\rho_m}{\rho_m + \rho_{m+1}} \end{aligned} \quad (4.18)$$

I generate  $\{D_m\}_{m=1}^M$  from a single family of distribution (log-normal) parameterized by a scalar. The interval of this scalar is chosen so that the mean of  $D_m$  are evenly spreading over  $(0, \bar{i})$ . This approximation suggests the key feature of dynamic tradeoffs: selling more today means a high-mean inventory distribution for tomorrow, which lowers future sales. The effect of this approximation in computing equilibrium depends on the coarseness of the discretization and the choose of  $\bar{i}$ .

### 4.4.2 Solving the model

Discretizing the action sets of prices can significantly mitigate the computational burden, considering the large number of products. By contrast, looking for optimal price vector would be almost impossible if action sets are continuous or even unbounded. The price of

each product has to be chosen from a price set  $P_{jx} = \{p_{jx,1}, \dots, p_{jx,n_{jx}}\}$  with  $n_{jx}$  elements. Restricting prices to be chosen from several fixed values is inspired by typical price series given by the data, in which price seems to be jumping around a few numbers. It is unclear why stores usually price at numbers like \$ 5.99 and \$ 9.99, but it's almost certain that these are not the global optimal prices, if the action sets of prices are continuous. I assume the numbers of candidate prices of each product, as well as their values, are exogenous. Relaxing this assumption, one has to look for the optimal action sets of prices, which would be even more burdensome and meaningless than just stick with continuous actions sets.

A challenge of pricing policy searching is that the number of products is large. Though each has a discrete set of price candidates, there is still an astronomical number of price combinations:  $\prod_{j,x} n_{jx}$ . I search the optimal price vector as follows:

1. Randomly choose a price vector from the set  $\prod P_{jx}$ .
2. Starting from the first element of the two vectors, update each element by choosing the optimal  $p_{jx} \in P_{jx}$  that gives the highest  $W(\cdot)$  while keeping all other elements unchanged. Record the new vector and  $W_1$  when this process reaches the last element.
3. Repeat step 2 and record  $W_k$  at the end of each  $k_{th}$ , until  $|W_k - W_{k-1}| < \phi$ .

It is a bit concerned that the return of this approach is probably the local maximizer (if  $W(\cdot)$  converges), not the global one. To look for the *true* maximizer, one has to either increase computer speed when using a Jacobian-based software command, or try different starting vectors that returns different local maximizers.

The equilibrium is computed as follows: starting at iteration  $k = 0$ , initialize the value functions  $V^0$  and  $W^0$  to zero, pricing and merchandise policy functions randomly chosen from  $\prod P_{jx}$ , consumption policy to be half of each consumer's inventory, and purchase policy of each consumer to be a randomly-chosen element from  $\{0, \dots, X\}$ . Next, for iterations  $k = 1, 2, \dots$ , follow the steps

1. For each  $D_m$ , update the consumer's value function  $V^k$  given the store's pricing and merchandise activity policies  $\{P^{k-1}\}$  by looking for the optimal consumption and purchase decision  $\{c^k, d^k\}$ . Calculate the market share  $\{s_{jx}^k\}$ .

2. Update inventory distribution and approximate it with  $\{D_m^k, \rho_m^k\}_{m=1}^M$ . Calculate beliefs about future distribution  $(\frac{\rho_{m+1}}{\rho_m + \rho_{m+1}}, \frac{\rho_m}{\rho_m + \rho_{m+1}})^k$ . For each  $D_m$ , update  $\{P^k\}$  given the market share  $\{s_{jx}^k\}$  obtained in step 1. Evaluate the store's value function  $W^k$  with  $\{D_m^k, \rho_m^k\}_{m=1}^M$ .
3. Repeat step 1 and 2 until converge.

To simulate a history of the dynamic game, I first generate an initial distribution of inventory, then for each period  $t = 0, \dots, T$ , update the next-period distribution and transition using the law of motion as described in equation (4.13). For the details of the Matlab program, see **Appendix B**.

## 4.5 Empirical Application

### 4.5.1 Data

I use the same scanner dataset of laundry detergent as in Hendel and Nevo (2006). The dataset records the detergent purchase history of 376 households in supermarkets and prices of each week in a time period of 104 weeks. Briefly, the dataset contains two components, store-level and household-level data. The store-level data is collected from nine supermarkets using scanning devices, from which the price of each product at each week is known. The household-level data contains the purchase history of each tracked consumer. It tells me when the consumer purchases a detergent product, how much she paid, and prices and merchandising activities of all other products on-shelf observed but not purchased by the consumer.

There are in total 29 brands and 5 typical package sizes (32, 64, 96, 128, 256 oz.) of laundry detergent. Some brands produce 256 oz. packages, but these purchases are omitted in this analysis to reduce computational dimensions, since the total market share of this size is merely 3 percent. Note that not all brands have all sizes, and usually not all products are available at each store.

The dataset also contains demographics of tracked households. Hendel and Nevo (2006) divide the households into six groups by demographics to explain different price sensitivities

and inventory costs. This extension is worth pursuing in future study, while all consumers are assumed to be identical in this analysis.

#### 4.5.2 Preliminary Analysis

Table 4.1 displays summary statistics of household purchase and consumption behaviors, as well as price adjustments of store. Households make a purchase of laundry detergent about every 13 weeks on average, and consume 3.5144 oz. every week. The stores adjust prices every two weeks, and are able to sell 2 packages per week on average to the observed population. The most popular package size is 64 oz. with a market share of 51 percent. Among the 29 brands recorded, 21 of them provide 64-oz. packages, comparing to 8, 11, and 17 brands that provide 32, 96, 128-oz. packages, respectively. Besides their size advantage, 64-oz products are mostly preferred because it is the most available package on shelf.

#### Duration Analysis of Non-purchase

The inventory model developed by Boizot et al. (2001) asserts that the expectation of non-purchase interval is a decreasing function of price of the last period, and an increasing function of current price. A duration analysis helps to get a feeling of data. Since households do not observe prices of goods every period, price of the last period would have no explanatory power. I Thus regress durations between two adjacent purchases on average price of the purchased good, current price, and dummies of package size:

$$\begin{aligned}\lambda(dur_{it}|p_{it}, \bar{p}) &= \alpha \phi_{it}^{\alpha} dur_{it}^{\alpha-1} \nu_i \\ \phi_{it} &= \exp(\theta_1 p_{it} + \theta_2 \bar{p} + \sum \theta_{3x} D_x)\end{aligned}\tag{4.19}$$

where subscript  $t$  denotes time,  $i$  denotes observation  $i$ ,  $\lambda$  is the probability of duration,  $p_{it}$  is the current price of the purchased product,  $\bar{p}$  is the average price over all observed period of the product,  $\nu_i$  represents random effect of each purchase that captures all unobserved characteristics of that purchase,  $D_x$  are dummies for package sizes. One can also consider include dummies for all brand-size combinations to capture brand-specific preferences, but since in the long run purchase durations depend on consumption rates, brand dummies would have very small explanatory power. Note that the parameters do not depend on

$t$ . This hazard function belongs to the Weibull family. It can be written as  $\ln dur_{it} = -\ln \phi_t + \ln u_t$ , in which  $u_t$  follows a type I extreme value distribution.

Since  $\bar{p}$  does not vary over time, the above regression is equivalent to regressing durations on a constant, the difference between current price and average price, and size dummies. The coefficient  $\theta_2$  can be interpreted as the price elasticity of duration. I expect  $\theta_2 < 0$  because of a positive correlation between current price and purchase duration. If include four dummies for the four packages sizes, the regression has no constant term. The final form of the regression is:

$$\log(dur_{ht} | d_{jxht} = 1) = \beta_1 p_{jxt} + \sum_x \beta_{1x} D_x + \epsilon_{jxht} \quad (4.20)$$

The result is reported in Table 4.1. The coefficient associate with price elasticity is 0.0215, implying at the mean duration (12.58 weeks), a price increase of 1 dollar will postpone purchase by 0.27 week.

If the households were truly homogeneous and consumption rate is constant over time, the coefficients associate with size dummies should decrease on size. however, the regression shows that those coefficients have no explicit relationship with package size. One possible reason is that household heterogeneity is ignored in the regression. If, for instance, family size is included, then the coefficients would be helpful explain package effects.

### Quantities sold

The store expects larger quantities sold during price-cuts, so the quantity of items sold is supposed to be negatively correlated with its promotion price. Following the empirical work by Boizot et al. (2001), I regress the quantity sold of each product during one week on its current price, average price, and size dummies:

$$Q_{jxt} = \beta_2 p_{jxt} + \sum_x \beta_{2x} D_x + \epsilon_{jxt} \quad (4.21)$$

where  $Q_{it}$  denotes total quantity sold of a specific product in week  $t$ . Given the large variances of the estimated coefficients, the null hypothesis of zeros price elasticity cannot be rejected. The sign of  $\beta_2$  is expected to be negative. I don't use the  $\log(Q)$  because the observed quantity sold in each week contains a lot of ones. The regression has no

intersection, as full dummies are included. The estimate of  $\beta_2 = -0.1480$  implies that a price-cut of 1 dollar will result in a 13% increase in quantity.

#### 4.5.3 Estimation

The parameters of the model are estimated by minimizing the distance between a set of moments of the data and their counterparts simulated by the model. This type of estimator, falls in the class of GMM estimator, is referred as simulated minimum distance (SMD) estimator by Hall and Rust (2003) because it minimizes the (weighted) distance between simulated and actual moments. By searching over the space of structural parameters, the SMD estimator finds a model with a stationary distribution that yields moments matching the actual ones. The efficient weight matrix is the inverse of the covariance matrix of the actual moments.

Because in general the equilibrium is not unique (Aguirregabiria and Mira, 2007), maximum-likelihood estimators cannot be used as the sum of probabilities of all possible outcomes is greater than 1. If the stochastic process smoothly jumps from one equilibrium to another, the probabilities conditional on previous evolution cannot be obtained and MLE becomes meaningless. We have to believe that the parameters are identified if the moments suggested by the simulation are similar to those actually observed. The idea is that at a given state, if the simulated policies and outcomes under the estimated parameters are identical to the observed actions and outcomes, then the estimates are true parameters. For example, the average prices in the actual and simulated data ought to be similar; at a simulated inventory distribution with a mean similar to the real one, the price elasticities should be similar to their counterparts implied by the data.

Formally, the Markov process is written by

$$\begin{aligned}
 & f(D_{t+1}, s_{t+1}, p_{t+1}, \{c_{h,t+1}, d_{h,t+1}\}_{h=1}^H | D_t, s_t, p_t, \{c_{h,t}, d_{h,t}\}_{h=1}^H, \epsilon_t; \Theta) \\
 & = f(D_{t+1} | D_t, s_t, p_t, \{c_{h,t}, d_{h,t}\}_{h=1}^H; \Theta) \\
 & \quad \times Prob(p_{t+1} = \sigma^*(D_t); \Theta) \\
 & \quad \times \prod_h Prob(c_{h,t+1}, d_{h,t+1} | D_t, i_{ht}, \epsilon_{jxht}; \Theta) \\
 & \quad \times I(s_{t+1} | D_t, p_t, \{c_{h,t}, d_{h,t}\}_{h=1}^H, \epsilon_t; \Theta)
 \end{aligned} \tag{4.22}$$

where  $\Theta$  denotes the structural parameter vector needs to be estimated. In a pure-strategy MPE, the two probabilities in the above equation are simply one, the market share and policies are deterministic functions of the current state,  $D_t$  and  $\epsilon_t$ . For each candidate  $\Theta$  encountered, starting at the initial state  $D_0$ , I solve for equilibrium policies and outcomes and simulate the model for  $T$  periods, assuming the process is governed by a pure-strategy equilibrium with all policies deterministic. By simulating the model  $N$  times, I can obtain a  $N \times T$  panel,  $\{\{D_t, s_t, p_t, a_t, \{c_{h,t}, d_{h,t}\}_{h=1}^H, \epsilon_t; D_0, \Theta\}_{t=1}^T\}_{n=1}^N$ , in which each observation is a model simulation of equilibrium evolution. Notice that the model distribution is of i.i.d and the randomness stems from the demand shocks,  $\epsilon_t$ .

Let  $\varphi = (s_t, p_t, a_t, \{c_{h,t}, d_{h,t}\}_{h=1}^H)$ , of which the moments are used for distance minimization. The moments implied by the actual data is denoted by the vector

$$m_a = m(\{\varphi^a\}_{t=1}^T) \quad (4.23)$$

and the simulated moment vector is the average over  $N$  simulations:

$$m_s(\Theta, D_0, p_0) = \frac{1}{N} m(\{\{\varphi^s; D_0, \Theta\}_{t=1}^T\}_{n=1}^N) \quad (4.24)$$

The SMD estimator is therefore the minimizer of the weighted quadratic form of distance between  $m_a$  and  $m_s$ :

$$\hat{\Theta}(D_0, p_0) = \underset{(\Theta)}{\operatorname{argmin}} (m_a - m_s(\Theta, D_0))' (\Omega^a)^{-1} (m_a - m_s(\Theta, D_0)) \quad (4.25)$$

where  $\Omega^a$  is the covariance matrix of the real moments. Suppose the parameter vector  $\Theta$  is of length  $K$  and the data implies  $L$  moments, then identification requires  $L > K$ . A problem with the SMD estimator in this application is that neither the inventory nor its distribution are observed by a econometrician. The estimates in the above equation is thus a function of the initial distribution and prices.

#### 4.5.4 Auxiliary Models

The policies of prices, promotions, consumption, purchase, and market share are deterministic functions of the state variables,  $D_t$  and  $\epsilon_t$ . A natural choice of variables whose moments used for SMD would be these policy functions at a given state. However, the

inventory distributions in this application are not observed. Even if they are, there is no good measure of  $D_t$  available.  $D_t$  is thus excluded from the choice set of moment variables.

The data reveals the nonpurchase duration of each consumer, as well as duration since the last promotion of each product. Boizot et al. (2001) shows that the duration since the last purchase is positively related to the current price and negatively to the last price, and quantity purchased increases in the last price. The empirics on ketchup conducted by Pesendorfer (2002b) shows that the aggregate quantity sold during a promotion is positively related to the duration since the last promotion. With these information, I choose the following moments:

1. Coefficients from regressing nonpurchase duration  $dur_h$  on the current price of the purchased product and size dummies:

$$\log(dur_h | d_{jxht} = 1) = \omega_{10} p_{jxt} + \sum_{s=1}^4 \omega_{1s} D_s + e_{jxht}.$$

2. Coefficients from regressing quantity sold (number of product items) on current price and size dummies, for all  $j, x, t$ :

$$Q_{jxt} = \omega_{20} p_{jxt} + \sum_{s=1}^4 \omega_{2s} D_s + e_{jxt}.$$

3. The average consumption rate. Since consumption is not observed by the econometrician, the actual consumption rate is calculated using the total amount of detergent (in oz.) that is purchased over the recorded period divided by period length, averaging over all households. The simulated consumption rate can be obtained from the consumption policy,  $\{c_t\}_{t=1}^T$ .

4. The average duration of nonpurchase. The simulated moment is calculated from the series of purchase policy,  $\{d_t\}_{t=1}^T$ .

5. The average price of each product, obtained from the time series of each brand-size specific product,  $\{p_{jxt}\}_{t=1}^T$ .

6. The average intervals of price adjustment, which is also calculated from store's pricing policy.

7. The average market share of the four package sizes, as an outcome under the optimal actions chosen by consumers and store, calculated by equation (4.7).

#### 4.5.5 Parameterizations

I estimate parameters  $\alpha, \xi$  using a static logit model, and structural parameters in  $\Theta = (\Theta_C, \Theta_I, \gamma)$ , where  $\Theta_C$  and  $\Theta_I$  are coefficients ahead of consumption and inventory cost, respectively:

$$\begin{aligned} u_C &= \theta_C * \log(c) \\ x_I &= \theta_{I1} * i + \theta_{I2} * i^2 \end{aligned} \tag{4.26}$$

The function forms above, following Hendel and Nevo (2006), capture the diminishing marginal utility of consumption and convexity of inventory cost ( $\theta_{I2} > 0$ ).

For computational simplicity, I first set the number of distributions to be 2, i.e., consider the simplest case in which there are only two states, a low inventory distribution and a high one. The grid of inventory distributions can be set more sophisticatedly by finer distance between adjacent distributions and/or more controlled density function family. I use uniform distribution to generate the initial individual inventories and distributions as state variables:

$$\begin{aligned} D_1 &\sim U(0, 20) \\ D_2 &\sim U(0, 60) \end{aligned}$$

The grid for policy functions are set as follows: The action set of consumption contains integers from 0 to 100 with an interval of 1; The action set for pricing of each product contains five numbers that are randomly selected from observed price series. The discount factor  $\delta$  is set to 0.95. The market size  $M$  is normalized to 1.

One has to be cautious when choosing the starting values for parameter search. Hendel and Nevo (2006) use maximum likelihood estimators to estimate a partial equilibrium for consumer stockpiling behavior. I adopt their estimates for the associate parameters in my model. For the parameters of the supply side, I choose the starting value of  $\gamma$  to be 0.001.

#### 4.5.6 Estimates and Model Fit

I choose  $T = 104$  and simulate the model  $N = 5$  times for each parameter estimation, and bootstrap the standard error of the estimates. For each simulated market evolution

at a tentative  $\Theta_k$ , where  $k$  denotes the index of simulation, three random variables are generated: the initial distributions of inventory are generated from the same family with parameterization; the initial prices are randomly picked from the price set of each product; and  $\epsilon_{jxht}$  are drawn from type I extreme value distribution with a variance of 50.

The parameters to be estimated are  $\Theta = (\theta_C, \theta_{I1}, \theta_{I2}, \gamma)$ . The model is over-identified, since the number of moments is greater than the number of parameters. I estimated the parameters using a model using moments and the auxiliary models with a base setting: the number of inventory distribution, one of the state variable, equals 2. In this case there exists a low-average inventory distribution,  $D_L$ , and a high-average distribution,  $D_H$ . More sophisticated estimates can be obtained by increasing the number of inventory distributions and changing distribution family from which the initial distributions are drawn. The estimates and variances estimated from a base model are reported in Table 4.3. The simulated moments and their pseudo t-values using the estimated parameters are reported in Table 4.4 and Table 4.5. The pseudo t-values are calculated by dividing the difference between the simulated and actual moments by the standard error of the actual moment. The pseudo-t for market shares are omitted, because the actual market shares of the four package sizes are calculated using the total quantities sold during the entire observed time period. The purchases of the tracked households in each period usually cover only a few kinds of products.

Moments that are below 3 are generally considered to be well-matched. Average interval of price change is two large comparing to the actual counterpart, though pseudo-t is less than 3. This means the estimate of menu cost is too high, inducing a low frequency of price adjustments. An alternative explanation would be that the menu cost of price adjustments are not included in variable cost. The estimated consumption rate of each household per week, 7.7480 oz., is very high comparing to common sense. The actual consumption rate of laundry detergent would largely depend on the amount of laundry. According to the model, an increase in inventory after a purchase will result in acceleration of consumption in order to increase the utility from consumption and to avoid a high inventory cost. However, in reality, an unplanned purchase of detergent due to price-cut would probably not cause consumption acceleration, as a large amount of detergent will simply be wasted if overdose.

Therefore, the choice of function form for consumer utility could be altered to some threshold function, in which consumption rate is more rigidly depend on the amount of laundry, and inventory cost dose not increase as long as the inventory is not too high. For the same reason, the continuous function form of consumer utility induces a shorter simulated non-purchase duration.

The simulated market shares of the four sizes are quite well-matched. The simulated time series of purchase behavior reveals that 64 oz. is the most popular package, with market share of 45 percent (actual 50). The smallest package has 25 percent of the market share, comparing to actual value 4 percent. The utility function of consumer may indicate a high sensitivity to inventory cost than actual households, thus the smallest package is more preferred.

The coefficient associate with price elasticity of quantities sold is negative for both actual and simulated regressions, implying a negative relationship between current price of a specific product and its quantity sold in that period. An price-cut of 1 dollar will induce an increase in sale by 0.272 for the given set of households. The price elasticities of non-purchase duration are both positive for both simulated and actual regressions, as expected. It implies that a consumers would postpone purchase if she observe a high price, or an unplanned purchase is made if a price-cut is observed. The simulated elasticities are less positive for all package sizes than their actual counterparts, because the model indicates some sensitivity to inventory cost and consumers are more cautious when making purchase decisions. At the average duration, the simulation implies a purchase would be brought advance by 0.12 week at a price-cut of 1 dollar, while for actual consumers it's 0.27 week.

## 4.6 Results

### 4.6.1 The Realized Series of Purchase and Pricing Decisions

To get a feeling of the histories of optimal behaviors, I first present the simulated time series of consumer and store's actions. Figure 4.1 depicts the time series of consumption, inventory, and purchase behavior of a randomly selected consumer in 104 weeks. Inventory increase immediately after a purchase, and decreases gradually if no purchases are made.

It's also worth noticing that after purchase, consumption level increases, in order to avoid the high inventory cost caused by the newly bought product.

Figure 4.2 shows a typical price fluctuation of a randomly picked product. Since the action set of store consists of candidate prices of all products that are drawn from actual prices, the optimal pricing series of a specific product indicates when the store decides to hold a price-cut. The regular price of this product is \$ 5.79, while it might be reduced to \$ 4.85 or \$ 4.99 during sales.

It is expected that consumers would make unplanned purchases if they observe price-cuts. Thus the total quantity sold during sales is supposed to be higher than regular periods. Figure 4.3 shows the time series of total quantity sold of the four sizes and the number of products with price-cuts. There is a positive correlation between the two variables: the more the products are on sale, the more the total quantities are expected to be sold. One can also study the price elasticity by regressing quantity sold of a specific product on its current price, as shown in Table 4.4.

Figure 4.4 shows consumer policies of consumption and purchase as functions of current inventory  $i_{ht}$  and inventory distribution,  $D_t$ . Again I consider the simplest case where there exists two inventory distributions, low and high. At both distributions, first note that in each period, most consumers (77 percent for low distribution and 94 percent for high inventory) don't buy. At high distribution state, consumers are less likely to make purchases because a high inventory cost would incur. Moreover, the consumption policy as a function of  $i_{ht}$  increases almost linearly with inventory level, but with negative drops at certain points. For example, at  $i_t = 20$  and  $22$ ,  $c_t(i_t = 20|D_t = D_H)$  is higher than  $c_t(i_t = 22|D_t = D_H)$ . The negative step can be explained by the optimal consumption chosen conditional on purchase a 64 oz. product. Thus, a high consumption is chosen to obtain high utility from utility and avoid inventory cost.

To induce consumers to buy, the store's pricing policy is to cut prices of a set of products. Figure 4.5 shows the average prices over all brands of the four sizes at the two distributions. At the state of high distribution, store sets lower average prices for 96 oz. and 128 oz. packages. Note the behavior of price-cuts at high-distribution state would cause a lower probability of high distribution in the next period. Thus the store's pricing behavior can be

interpreted as smoothing inventory distributions over periods, and thus smoothing its profit over periods. The probability of purchase would be even lower if store didn't cut prices, and the probability of high distribution in the next period conditional on high distribution would be even higher.

Consumer's  $V(i_t, D)$  is generally lower at high  $i_t$ , as a higher inventory cost in the current period reduce the value of  $V(i_t, D)$ , as showed in figure 4.6. Moreover, at high-distribution state, the average value of  $V(\cdot)$  is lower than at low-distribution state. Figure 4.7 shows consumer value function when the number of inventory distribution is 7.

#### 4.6.2 The Optimal Store Pricing

Markov-perfect equilibrium requires that the evolution of the inventory distribution that is governed by the optimal actions of all agents is consistent with their beliefs on future states. To Study the dynamics of the evolution of the entire market, I plot the time series of average inventory level of all consumers, total quantity of purchase, beliefs in next-period probability of  $D_H$  ( $\rho$ ), and average price of all products in Figure 4.8. This helps explain the dynamics of the overall inventory, pricing, and sales quantity.

Store's optimal pricing behavior is revealed by the correlation between the average price of all products(four sizes, all brands) and store's belief in inventory distribution in the next period.  $\rho$  stands for the probability of transiting to high inventory distribution. Although averaging over all products is a very loose measure of average market price, it is able to reflect the connection between store's belief and its sales promotion strategy. In the simulated time period of 104 weeks, eleven 'big' sales can be observed, each marked by a drop in the average price. First, note that the average price do not drop if store believes that the next period inventory distribution will be low. The reason is that a forward-looking store decides to preserve sales for future by setting current prices high. Since a low-inventory distribution in the next period means a high probability of purchase, thus a high sales quantity, the store's optimal action is to set current price high so that profit in the next period will not be too low. Therefore, comparing to a myopic store that choose optimal prices to maximize profit of a single period, the forward-looking store's pricing strategy is to induce a smoother series of sales quantities over periods.

In fact, keeping prices at regular levels guarantees a low purchase probability in the current period and a high probability of low inventory distribution in the next period. The store's belief in next-period distribution is consistent with the evolution of distribution states that is governed by its optimal pricing decision.

Second, most 'big sales' are hold during the periods after which inventory distribution will be high. Because a high distribution in the next period implies low sales quantities, the store chooses to cut prices to increase quantities sold in the current period and therefore make a high current-period profit. Similarly, it is the store's price-cut behavior that induces a high inventory distribution in the next period.

Figure 4.2 shows store's value function, in which  $M$  denotes the number inventory distributions. The total expected profit is generally decreasing on average inventory. This implies that it's difficult to make profits when consumers are holding large amount of inventory in the current period. At a low inventory distribution, a forward-looking store tends to set prices higher than at a high one, because the store knows that given the high probability of purchase of the forward-looking consumers, it is anyway easy to make a profit in the current period. Thus, the store would rather set high prices to preserve future demand. A low inventory distribution is always preferred by a store than a high one. In the extreme case of zero inventories, a store could sell at very high prices, obtaining a high current profit, and is still able to preserve a fairly large amount of demand for tomorrow.

#### 4.6.3 The Optimal Consumer Purchase

Figure 4.8 also shows the time series of average price of all products and the total purchase quantity(four sizes, all brands) of all consumers. First, it is clear that the 'big sales' indeed induce purchase spikes. Rational forward-looking consumers are more likely to make purchase if observe price-cuts, because consumers' belief about next-period high inventory distribution is of high probability, which means high future prices. Again, consumers' belief about  $Prob(D_{t+1} = D_H)$  is consistent with the evolution of states that is induced by the optimal action under this belief. The remarkable purchases rushes mostly occur during regular prices periods. These purchases primarily result from extremely low inventories and urgent consumption needs. The store knows that the big purchase spike will cause a high

probability in  $D_H$  in  $t + 1$ , thus decides to keep prices at regular levels to maintain a high profit in the current period.

Having established the optimal dynamic strategies of the store and consumers, I now present cyclical features of equilibrium behavior and outcomes when there exists only two distributions, high and low. The cyclical patterns of the optimal behavior and evolution of states are depicted in figure 4.9.

**Observation 1.** Store decides to hold big sales by cutting prices of a large number of products when the average inventory of consumers is high (Figure fraction 1, 4.9). The expected quantity sold is high enough so that the next-period inventory level is very likely to be high. The store's belief is  $Prob(D_{t+1} = D_H | D_t = D_H)$  is close to one.

**Observation 2.** After the big sales, Store keeps prices at regular levels, along with the gradual decline in average inventory level that is primarily due to consumption and small purchases. Store dose not hold the next big sale until the average inventory recovers to high level. The consistent belief is that at some medium average inventory in the current period,  $Prob(D_{t+1} = D_L | D_t)$  is greater than  $Prob(D_{t+1} = D_H | D_t)$ . (fraction 2, Figure 4.9)

**Observation 3.** Store dose not cut prices when the average inventory level is low, because a rise in demand will be driven by consumption needs instead of low prices. The consistent belief is: next-period average inventory level will be higher than current level, and  $Prob(D_{t+1} = D_H | D_t = D_L)$  is small.

**Observation 4.** At extremely low average inventory, a purchase spike occurs, even the prices are mostly at regular levels. Inventory level recovers to high level after the purchase spike (fraction 3, Figure 4.9),  $Prob(D_{t+1} = D_H | D_t = D_L)$  is small.

## 4.7 Conclusions

This paper presents a dynamic model of storable goods with endogenous consumption, purchase, and pricing. The model entails two methodological contributions. First, I bind

the action set of pricing to be a set of finite candidate prices. Under this specification, the optimal pricing behavior of store can be solved without tedious dimensionality problem. Second, because inventory as a state variable and consumption as a control variable are not observed, I propose an alternative approach to parametrically estimate the model, using SMD method with moments of price elasticities of non-purchase duration and total quantity sold.

The model is estimated using data that consists of both store- and household-level observations. I numerically show that the optimal pricing policy of store is to lower prices at high overall inventory with weak incentives to purchase new product, and keep prices at regular level at low average inventory with a high demand driven by urgent consumption needs, in order to preserve future demand and profits. purchase spike of consumers occurs when average inventory is low, primarily driven by consumption needs. Consumers respond to price-cuts by some medium purchase spike. Furthermore, the agents' beliefs on inventory distribution are consistent with evolution of inventory distribution that is governed by the optimal strategies under that beliefs.

The estimates of the model may depend on the initial conditions, which are generated from the same set of distributions for each estimation. I generate the initial distribution of inventory from the family of Uniform distribution. The choice of distribution family and parameterization will surely affect policy functions and thus model estimates.

One possible extension of the interest is to take into account heterogeneity of the households for better parameter estimates. Another possibility is to change the function form of the utility functions to get different value function and numerical results. The threshold functions is worth considering, because it implies a weaker aversion to inventory cost.

Table 4.1: Preliminary Analysis

	mean	std
Summary Statistics of Household and Store		
non-purchase duration (weeks)	12.5864	12.8844
consumption rate (oz./week)	3.6144	4.3003
interval between price changes (weeks)	2.0122	2.5552
quantity sold	2.0636	1.7158
Market Share		
32 oz.	0.0450	
64 oz.	0.5087	
96 oz.	0.0822	
128 oz.	0.3462	
Preliminary Regressions		
non-purchase duration		
$p_t$	0.0215	0.7347
$D_1$	2.3768	5.4969
$D_2$	1.9239	2.9168
$D_3$	1.8853	5.0488
$D_4$	2.1074	4.7756
quantity sold		
$p_t$	-0.1480	1.2387
$D_1$	1.8329	9.0676
$D_2$	2.4796	4.9728
$D_3$	2.0338	8.5542
$D_4$	3.4061	8.1578

Table 4.2: Store Value Function  $W(D_t)$ 

	$D_1$	$D_2$	$D_3$	$D_4$	$D_5$	$D_6$	$D_7$
average inventory	10	30	50	70	90	110	130
M = 2	3.1454	2.7234					
M = 5	4.0638	4.0461	4.0778	3.5389	2.7158		
M = 7	4.0748	3.4387	3.0162	3.0368	3.0762	2.0973	2.0963

Table 4.3: Estimated Parameters Of The Base Model

Parameter	Estimates	Variance
$\theta_C$	4.2125	0.4712
$\theta_{I1}$	1.9860	0.0175
$\theta_{I2}$	$9.6702 \times 10^{-4}$	$1.1871 \times 10^{-4}$
$\gamma$	1.1375	0.0001

Table 4.4: Simulated vs. Actual Moments I

Auxiliary Moment	Simulated	Actual	pseudo-t
Average non-purchase duration	4.6956	12.5864	0.6124
Average consumption rate	7.6480	3.5144	1.7785
Average price change interval	9.2444	2.0122	2.8344
Market Share			
32 oz.	0.2586	0.0450	
64 oz.	0.4546	0.5087	
96 oz.	0.0632	0.0822	
128 oz.	0.2236	0.3642	
Coefficients From Regression			
$\log(dur_h)$			
$p_t$	0.0259	0.0215	0.0023
$D_1$	0.3930	2.3768	0.0043
$D_2$	0.7281	1.9239	0.0343
$D_3$	0.8316	1.8853	0.0044
$D_4$	1.4829	2.1074	0.0280
$Q$			
$p_t$	-0.1238	-0.1480	0.0158
$D_1$	1.3038	1.8329	0.0158
$D_2$	1.1421	2.4796	0.0541
$D_3$	0.8292	2.0338	0.0165
$D_4$	1.0448	3.4061	0.0355

Table 4.5: Simulated vs. Actual Moments II

Brand	Simulated Average Price of 32 fl. oz.	Actual	pseudo-t
AJAX	3.18	2.87	3.4087
ALL	3.28	3.15	1.4680
ALL FREE	3.34	3.37	0.0639
ARM & HAMMER	4.99	3.86	0.4902
BOLD 3	3.89	3.45	4.2089
CHEER	4.58	4.48	1.4143
DASH	2.98	2.84	6.2216
DREFT	5.58	5.39	1.5871
DYNAMO	5.23	3.51	5.0954
ERA PLUS	4.59	3.89	3.3661
FAB	4.57	4.29	1.1335
IVORY SNOW	4.79	4.79	0
TIDE	5.17	4.92	2.1076
PRIVATE LABEL	2.98	2.47	4.6548
PUREX	2.98	2.61	14.0468
SOLO	3.98	3.70	3.1322
SURF	4.47	4.11	1.4420
ULTRA TIDE	4.58	4.40	1.4696
WISK	4.23	3.44	2.1859
YES	3.08	2.90	3.8176

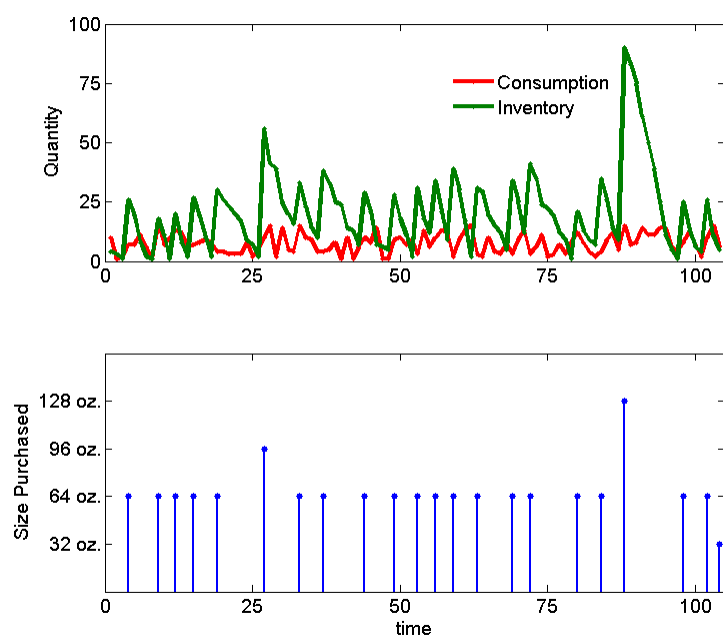


Figure 4.1: Simulated Histories Of Purchase and Consumption

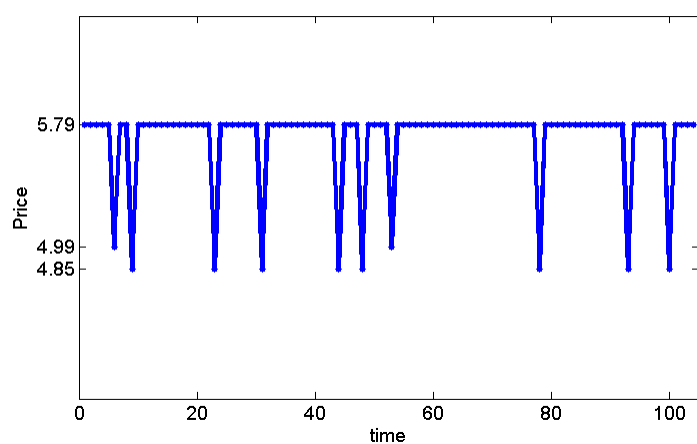


Figure 4.2: Simulated Price of a Randomly Selected Product

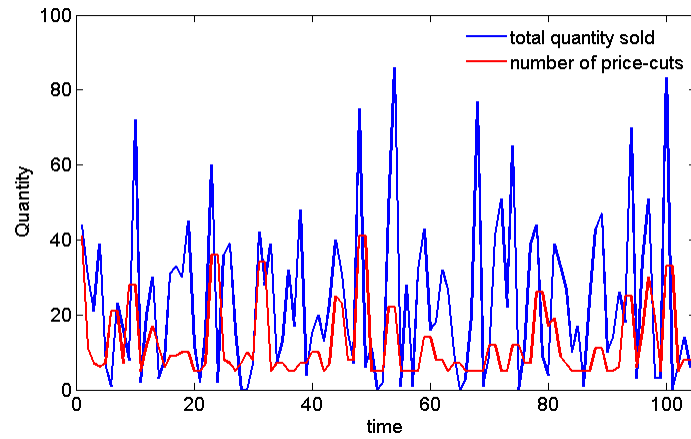


Figure 4.3: Correlation Between Number Of Price-cuts and Quantity Sold

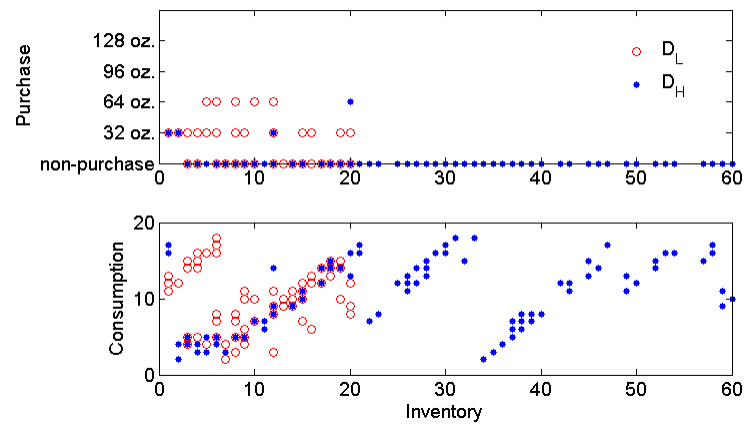


Figure 4.4: Consumption and Purchase Policies

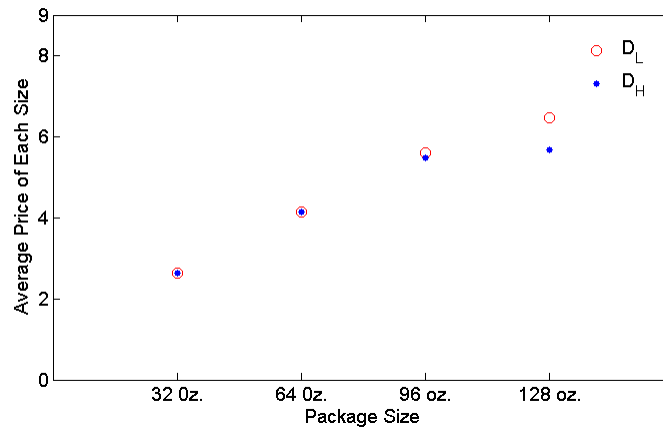


Figure 4.5: Store Pricing Policy

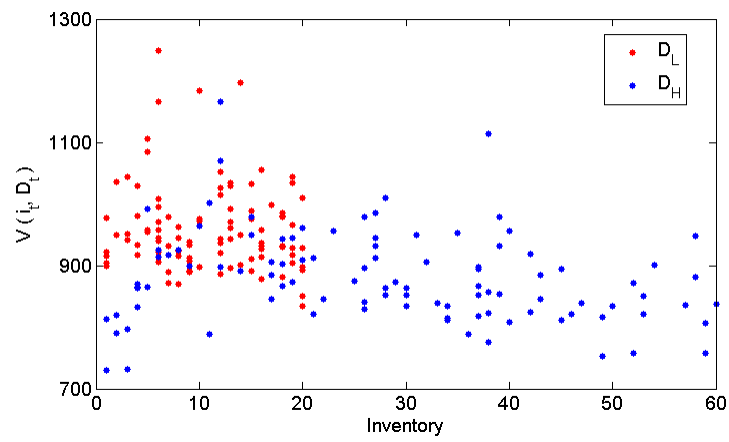


Figure 4.6: Consumer Value Function (The Number of Inventory Distributions = 2)

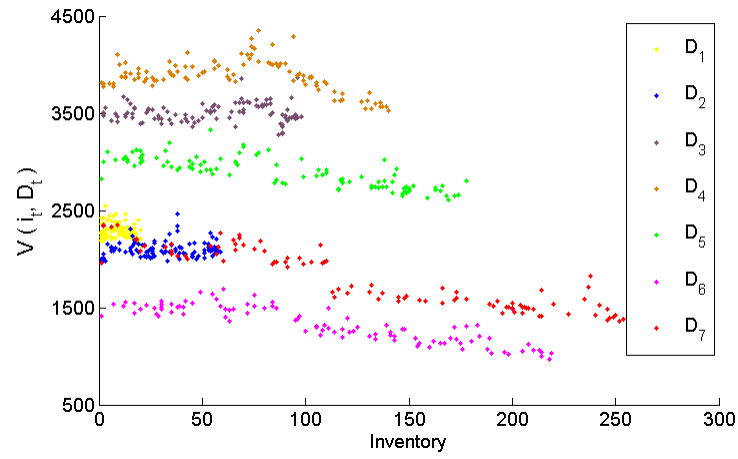


Figure 4.7: Consumer Value Function (The Number of Inventory Distributions = 7)

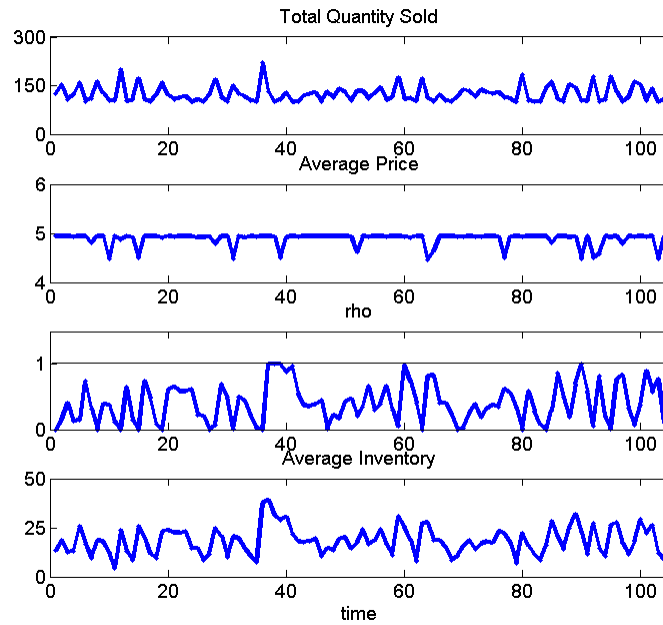


Figure 4.8: Simulated Market Evolution

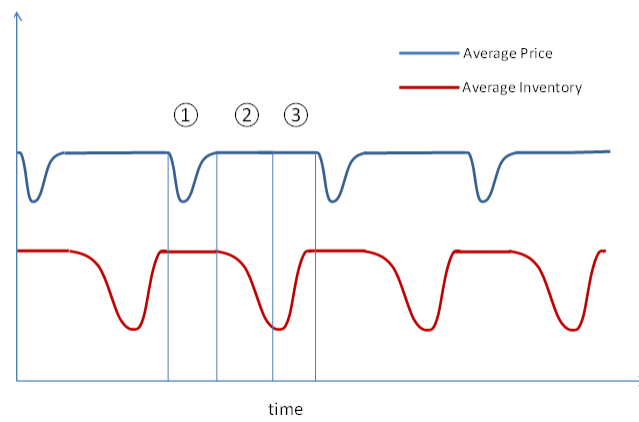


Figure 4.9: Patterns of Prices and Consumer Inventory

## Appendix A

*Derivation of Eq. (3.20)*

Notice that  $V(P_{t+1}, k-1, k-1+L)$ , the deviator's continuation value, can be written as the sum of cumulative inventory costs and consumption utilities in the next  $k-1$  periods, and the continuation value when she runs out of stock and forced to purchase at  $p^R$ :

$$\begin{aligned}
 & V(P_{t+1}, k-1, k-1+L) \\
 &= \sum_{i=1}^{k-1} -c(k-1-i)\delta^{i-1} + v \sum_{i=1}^{k-1} \delta^{i-1} + \delta^{k-1} \tilde{V}(p^R, 0, L, d_{k-1}=1) \\
 &= \sum_{i=1}^{k-1} -c(k-1-i)\delta^{i-1} + v \sum_{i=1}^{k-1} \delta^{i-1} \\
 &\quad + \delta^{k-1} \left( \sum_{i=0}^{L-1-K} -c(L-1-i)\delta^i + v \sum_{i=0}^{L-K} \delta^i + \delta^{L-K} V(P, K, K) \right) \tag{A.1} \\
 &= \sum_{i=1}^{k-1} -c(k-1-i)\delta^{i-1} + \sum_{i=0}^{L-1+K} (-c(L-1-i))\delta^{k-1+i} \\
 &\quad + v \left( \sum_{i=1}^{k-1+L-K} \delta^{i-1} \right) + \delta^{k-1+L-K} V(P, K, K)
 \end{aligned}$$

## Appendix B

### *A Brief Description of the Matlab Program*

The computer program has three loops:

The outer loop is a parameter search procedure, in which the estimated parameters are found if the weighted moments is globally minimized. This procedure is done by calling builtin '`fminsearch`'.

The middle loop is to simulate the time series of consumption and purchase of consumers, and the price sequence of all products, at each  $\Theta$ , and returns the simulated moments.

The inner loop solves consumer and store's policy function at each period, taking the outcome induced by the policies of the last period as the state variable of the current period.

## Bibliography

- Aguirregabiria, V. and P. Mira, 2007: Sequential estimation of dynamic discrete games. *Econometrica*, **75** (1), 1–53.
- Bagwell, K., 2007: The economic analysis of advertising. *Handbook of industrial organization*, **3**, 1701–1844.
- Bajari, P., C. L. Benkard, and J. Levin, 2007: Estimating dynamic models of imperfect competition. *Econometrica*, **75** (5), 1331–1370.
- Bell, D., T. Ho, and C. Tang, 1998: Determining where to shop: Fixed and variable costs of shopping. *Journal of Marketing Research*, 352–369.
- Bell, D. and J. Lattin, 1998: Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer edlp. *Marketing Science*, **17** (1), 66–88.
- Benkard, C. L., 2004: A dynamic analysis of the market for wide-bodied commercial aircraft. *The Review of Economic Studies*, **71** (3), 581–611.
- Bester, H. and E. Petrakis, 1995: Price competition and advertising in oligopoly. *European Economic Review*, **39** (6), 1075–1088.
- Boizot, C., J.-M. Robin, and M. Visser, 2001: The demand for food products: an analysis of interpurchase times and purchased quantities. *The Economic Journal*, **111** (470), 391–419.
- Bolton, R. and V. Shankar, 2003: An empirically derived taxonomy of retailer pricing and promotion strategies. *Journal of Retailing*, **79** (4), 213–224.
- Bolton, R., V. Shankar, and D. Montoya, 2010: Recent trends and emerging practices in retailer pricing. *Retailing in the 21st Century*, 301–318.

- Bonnet, C., P. Dubois, S. B. Villas Boas, and D. Klapper, 2010: Empirical evidence on the role of non linear wholesale pricing and vertical restraints on cost pass-through. *The Review of Economics and Statistics*, **(0)**.
- Bresnahan, T. F., 1987: Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, 457–482.
- Butters, G., 1977: Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, 465–491.
- Carranza, J. E., 2010: Product innovation and adoption in market equilibrium: The case of digital cameras. *International Journal of Industrial Organization*, **28 (6)**, 604–618.
- Chintagunta, P. K., A. Bonfrer, and I. Song, 2002: Investigating the effects of store-brand introduction on retailer demand and pricing behavior. *Management Science*, **48 (10)**, 1242–1267.
- Conlisk, J., E. Gerstner, and J. Sobel, 1984: Cyclic pricing by a durable goods monopolist. *The Quarterly Journal of Economics*, **99 (3)**, 489–505.
- Dixit, A. and V. Norman, 1978: Advertising and welfare. *The Bell Journal of Economics*, 1–17.
- Ericson, R. and A. Pakes, 1995: Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, **62 (1)**, 53–82.
- Fox, E. and S. Hoch, 2005: Cherry-picking. *Journal of Marketing*, 46–62.
- Gauri, D., K. Sudhir, and D. Talukdar, 2008a: The temporal and spatial dimensions of price search: Insights from matching household survey and purchase data. *Journal of Marketing Research*, **45 (2)**, 226–240.
- Gauri, D., D. Talukdar, and B. Ratchford, 2008b: Empirical investigation of the impact of loss leader promotion on store and category performance in grocery industry. Tech. rep., Working Paper, Syracuse University.

- Goettler, R. and B. Gordon, 2008: Durable goods oligopoly with innovation: theory and empirics. *manuscript, Carnegie Mellon University*.
- Grossman, G. M. and C. Shapiro, 1984: Informative advertising with differentiated products. *The Review of Economic Studies*, **51** (1), 63–81.
- Hall, G. and J. Rust, 2003: Simulated minimum distance estimation of a model of optimal commodity price speculation with endogenously sampled prices. *manuscript, Yale University*.
- Hendel, I. and A. Nevo, 2006: Measuring the implications of sales and consumer inventory behavior. *Econometrica*, **74** (6), 1637–1673.
- Ho, K., 2009: Insurer-provider networks in the medical care market. *American Economic Review*, *99*(1): 393–430.
- Ho, T., C. Tang, and D. Bell, 1998: Rational shopping behavior and the option value of variable pricing. *Management Science*, **44** (12-Part-2), S145–S160.
- Hong, P., R. P. McAfee, and A. Nayyar, 2002: Equilibrium price dispersion with consumer inventories. *Journal of Economic Theory*, **105** (2), 503–517.
- Hosken, D. and D. Reiffen, 2004: Patterns of retail price variation. *RAND Journal of Economics*, 128–146.
- Hosken, D. and D. Reiffen, 2007: Pricing behavior of multiproduct retailers. *The BE Journal of Theoretical Economics*, **7** (1).
- Ishii, J., 2011: Compatibility, competition, and investment in network industries: Atm networks in the banking industry. *manuscript, Stanford University*.
- Katz, M., 2007: Supermarkets and zoning laws. *Unpublished Ph.D. Dissertation, Harvard University*.
- Kohls, R. L. and J. N. Uhl, 2001: *Marketing of Agricultural Products*. Prentice Hall.
- Lal, R. and C. Matutes, 1994: Retail pricing and advertising strategies. *Journal of Business*, 345–370.

- Levy, D., M. Bergen, S. Dutta, and R. Venable, 1997: The magnitude of menu costs: direct evidence from large us supermarket chains. *The Quarterly Journal of Economics*, **112** (3), 791–824.
- Marshall, A., 1919: *Industry and Trade: A Study of Industrial Technique and Business Organization and of Their Influences on the Conditions of Various Classes and Nations*. MacMillan and Co., London.
- Maskin, E. and J. Tirole, 1988: A theory of dynamic oligopoly, i: Overview and quantity competition with large fixed costs. *Econometrica: Journal of the Econometric Society*, 549–569.
- McAfee, R. P., 1995: Multiproduct equilibrium price dispersion. *Journal of Economic Theory*, **67** (1), 83–105.
- McFadden, D., 1974: Conditional logit analysis of qualitative choice behavior.
- Nair, H., 2007: Intertemporal price discrimination with forward-looking consumers: Application to the us market for console video-games. *Quantitative Marketing and Economics*, **5** (3), 239–292.
- Narasimhan, C. and R. T. Wilcox, 1998: Private labels and the channel relationship: A cross-category analysis\*. *The journal of business*, **71** (4), 573–600.
- Nevo, A., 2001: Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 307–342.
- Nevo, A. and C. Wolfram, 2002: Why do manufacturers issue coupons? an empirical analysis of breakfast cereals. *RAND Journal of Economics*, 319–339.
- Nichols, L. M., 1985: Advertising and economic welfare. *The American Economic Review*, **75** (1), 213–218.
- Pakes, A., 2010: Alternative models for moment inequalities. *Econometrica*, **78** (6), 1783–1822.

- Pakes, A. and P. McGuire, 1992: Computing markov perfect nash equilibria: Numerical implications of a dynamic differentiated product model. National Bureau of Economic Research Cambridge, Mass., USA.
- Pakes, A., M. Ostrovsky, and S. Berry, 2007: Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *The RAND Journal of Economics*, **38** (2), 373–399.
- Pakes, A., J. Porter, K. Ho, and J. Ishii, 2011: Moment inequalities and their application. *Unpublished Manuscript, Harvard University*.
- Pesendorfer, M., 2002a: Retail sales: A study of pricing behavior in supermarkets\*. *The Journal of Business*, **75** (1), 33–66.
- Pesendorfer, M., 2002b: Retail sales: A study of pricing behavior in supermarkets\*. *The Journal of Business*, **75** (1), 33–66.
- Porter, R. H., 1983: A study of cartel stability: The joint executive committee, 1880-1886. *The Bell Journal of Economics*, 301–314.
- Richards, T. J., 2007: A nested logit model of strategic promotion. *Quantitative Marketing and Economics*, **5** (1), 63–91.
- Roy, S., 2000: Strategic segmentation of a market. *International Journal of Industrial Organization*, **18** (8), 1279–1290.
- Rust, J., 1987: Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, 999–1033.
- Smith, H. and O. Thomassen, 2012: A Model of Demand for Grocery Categories Mimeo, oxford University.
- Sobel, J., 1984: The timing of sales. *The Review of Economic Studies*, **51** (3), 353–368.
- Stahl, I. and O. Dale, 1994: Oligopolistic pricing and advertising. *Journal of economic theory*, **64** (1), 162–177.

- Stegeman, M., 1991: Advertising in competitive markets. *The American Economic Review*, 210–223.
- Stigler, G. J. and G. S. Becker, 1977: De gustibus non est disputandum. *The American Economic Review*, **67** (2), 76–90.
- Talukdar, D., D. Gauri, and D. Grewal, 2008: Extreme cherry pickers: Prevalence, profile and profit impact in the frequently purchased goods market. Tech. rep., Working Paper, SUNY Buffalo.
- Tirole, J., 1988: *The Theory of Industrial Organization: Jean Tirole*. MIT press.
- Varian, H. R., 1980: A model of sales. *The American Economic Review*, 651–659.
- Villas-Boas, S. B., 2007: Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, **74** (2), 625–652.
- Yang, Z., Z. Bi, and N. Zhou, 2005: The double jeopardy phenomenon and the mediating effect of brand penetration between advertising and brand loyalty. *Journal of Advertising Research*, **45** (2), 211.