Essays on Pricing and Consumer Demand in the Retail Sector

Author: Lucrezio Figurelli

Boston College Electronic Thesis or Dissertation, 2013

Boston College

The Graduate School of Arts and Sciences

Department of Economics

ESSAYS ON PRICING AND CONSUMER DEMAND IN THE RETAIL SECTOR

a dissertation

by

LUCREZIO FIGURELLI

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

August 2013

ESSAYS ON PRICING AND CONSUMER DEMAND IN THE RETAIL SECTOR

Dissertation Abstract

by

Lucrezio Figurelli

Dissertation Committee:

JULIE HOLLAND MORTIMER (chair)

HIDEO KONISHI

ARTHUR LEWBEL

Abstract

This dissertation consists of two independent chapters on pricing and consumer demand in the retail sector. In chapter 1 develop an empirical model of Consumer Supermarket Choice that enables identification of heterogeneous consumer travel costs and is suitable for a wide range of policy experiments and the study of local competition. Chapter 2 is a theoretical investigation on pricing patterns in multi-product retail markets, when boundedly rational consumers' choice of a store is based on the price and valuation of a subset of goods.

Estimation of demand systems in spatially differentiated retail markets is fundamental for understanding local competition and the impact of policy changes. It is also challenging, because shopping decisions consist of multiple dimensions: when to shop, where to shop and what to buy. In chapter 1 I develop an empirically tractable model of store choice in the supermarket industry and provide a way to identify consumers' heterogeneous travel costs without imposing restrictions on bundle choice. Using micro level data on a small market in New England, I estimate demand for stores using both a moment inequality approach and standard discrete choice techniques. I specify utility as a function of both store and bundle characteristics, and control for the endogeneity of expenditure on the bundle. I use the estimates of the discrete choice model to evaluate the welfare impact of 1) the closing of each individual store in the market and 2) the relocation of one of the stores. I find that travel costs are heterogeneous and marginally decreasing; that people like to shop at stores that are close, but also like to shop at multiple stores. Furthermore, people value stores differently (across consumers and shopping occasion) and trade off additional travel time for better store characteristics; utility differentials in preference for stores correspond to a distance ranging between zero and up to 3.3 miles. Variation in demand and substitution patterns across stores are explained by differences in store characteristics and by the shopping habits

and geographic distribution of heterogenous consumers. Changes in market structure, like store entry and exit can have significant impact on consumer welfare. For example, removal on one of the stores results in a loss in CS that ranges between 8% and 44%.

The assumption of rationality in retail shopping decisions appears very problematic when stores sell thousands of products and frequently vary their assortments and prices. Consumers are typically uncertain about prices at different stores and for a consumer to consider the entire distribution of bundles and prices might be a far too complex decision process. Furthermore, models with rational consumers are incapable of fully explaining important features of retail markets such as price dispersion, advertising and leader pricing. In chapter 2 I attempt to characterize optimal pricing by multi-product retailers when imperfectly informed consumers buy more than one product. The distinctive feature of the model is that there are two relevant moments to all purchase decisions. First, the choice of a store to visit, and second, the choice of the items to purchase. While consumers might rationally choose a store to best meet their specific needs and desires, the choice of the items to purchase is made only once in a store. Whether guided by impulse, contingent and unforeseen needs or in-store learning about a product, consumers often end up buying additional products which can generate higher profits for the stores. To examine the implications on retail pricing of this kind of behavior, I depart from a standard rational setup and introduce the concept of attractor goods. Using an an approach similar to that found in Osborne and Rubinstein (1998) and Spiegler (2006) I consider boundedly rational consumers whose choice between stores is based solely and entirely on the price and valuation of a subset of goods, the attractors. I show that retailer's pricing decisions have to take into account not only the direct effect of prices on a product's demand but also the effect on the demand for the other products sold in the store. The optimal pricing schedule will be a decreasing function of the goods' attractiveness, and pricing below marginal cost might be optimal for some goods. The model provides a rationale for the strategy of loss leader pricing and offers an intuitive explanation to countercyclical markups.

To my wife Anna, whose love and support have guided me throughout this work,
and to our daughters, Adelaide and Carolina, for the joy they brought into our lives.

Contents

List of Figures

List of Tables

Acknowledgements

I am very thankful to my advisor Julie Holland Mortimer, for her constant encouragement, advice and mentorship. I thank my advisors Hideo Konishi and Arthur Lewbel, for their tremendous guidance and advice. I thank Mark Rysman, Ariel Pakes, Gregory Lewis, Samson Alva and Donald Cox for helpful comments and discussions. I thank Inacio Bo for his technical assistance, and Emek Basker for her industry insights. Finally, I would like to thank seminar participants at the Harvard Industrial Organization Workshop, the Green Line IO meeting, the Boston College Dissertation Workshop, Applied Micro Meeting and Micro Theory Workshop, for helpful comments and suggestions. All estimates and analyses in the first chapter of this dissertation, based on SymphonyIRI Group, Inc. data, are by the author and not by SymphonyIRI Group, Inc.

Chapter 1

Store Choice in Spatially Differentiated Markets

1.1 Introduction

Retail markets are a major part of the US economy, with total sales of $3.8 trillion and over 14 million people employed in 2010. At the final stage of the distribution channel, these markets provide consumers access to a large number of final products. The spatial dimension of the industry makes stores inherently differentiated, and consumers choose stores based both on the price of available products and travel costs. Estimation of demand systems in these markets needs to account for the multi-dimensional nature of consumers' choice in order to accurately predict substitution patterns and price elasticities. These, in turn, are fundamental for understanding store location decisions, the effect of mergers, and the welfare impact of zoning regulations and counterfactual industry structures.

In this paper I develop and estimate an empirically tractable model of store choice in the supermarket industry that accounts for the spatial dimension of the market by explicitly modeling consumers' heterogeneous travel costs of driving to the stores. The choice of the supermarket industry is particularly relevant, as consumers visit supermarkets frequently, on average more than twice a week, and spend a large portion of their income on groceries and other household supplies. Frequency of trips and persistency of behavior make the supermarket industry an ideal context for the study of consumer store choice, as individuals, at least on average, will tend to make more informed decisions once they have to repeat them over time.

The choice of a store (location) is only one dimension of a more complex shopping decision (timing-location-bundle), and most of the difficulties in the study of store choice arise from having to deal with the other dimensions as well. The agents' choice set can be very large and complex, as supermarkets carry thousands of products, and retail outlets differentiate by location, so that the choice set will be different for different consumers. Shopping decisions are inter-temporally dependent, as bundle choices depend on previous purchases as well as current and future stocks, and consumers have imperfect information about prices at different stores.

The empirical approach I take is to study the consumers' choice of a store in a way that allows to abstract from the choice of a bundle. This requires some restrictions on the nature of consumers utility, but is very general in terms of how consumers choose what to buy and where to buy it. The motivating assumption of such an approach is that, although consumers have imperfect information about the exact bundle they will end up buying, specific needs guide their decisions, and they ultimately make their store choice based on a bundle they are planning to buy. Under this assumption store choice is essentially a discrete choice problem, and I can identify how location, store characteristics and bundle characteristics affect consumers' decisions.

Starting from a very general model of consumer supermarket choice, I introduce two sets of assumptions that allow me to estimate demand for supermarkets without imposing restrictions on the choice of the bundle. I start by considering a multinomial logit specification that allows for a flexible model of consumers' heterogeneity in travel costs. I construct trip specific price indices and control for endogeneity of price and expenditure to address the potential biases that could arise in presence of expectational and measurement errors. To further support my results, I then introduce a less restrictive moment inequality approach, first introduced in the context of supermarket choice by Katz (2007).[1] The inequality estimator does not require a parametric distribution of the disturbances, and is robust against certain types of measurement errors and consumers' expectational errors. However, my focus on travel costs ultimately stems from the desire to understand the welfare effects of policy decisions such as land use regulations. Addressing these questions requires choice probabilities under counterfactual market conditions, which require a distribution of the disturbance.[2] Therefore, I primarily focus on a discrete choice setup where a few reasonable assumptions allow me to exploit point identification and the ability to compute choice probabilities.

Using detailed micro-level data on trips and purchases made by a panel of households in a small market in New England, I estimate demand for supermarkets under the two methodologies. While both inequalities and discrete choice agree on how consumers trade off their travel costs for better store characteristics, I find that in my application, where only a limited number of alternatives (i.e. stores) can be used to construct the moments, the inequality approach fails to characterize the heterogeneity in consumers' travel costs, and provides too little information about how

[1]Results from this unpublished job market paper are reported in Pakes (2010); Katz's work is there presented as a motivating example to the methodology.

[2]A number of studies using inequality conditions have run counterfactual analysis. These studies however, use the methodology to estimate cost and supply side parameters which are then combined to estimates of a demand system. See, for example, Ishii (2005) Crawford and Yurukoglu (2011) and Ho, Ho, and Mortimer (2012).

people value the individual stores. Conversely, the discrete choice approach allows to identify consumers' heterogeneity in travel costs, and does not require any restrictive assumptions on the choice set of stores when data are available on all stores visited by a household.

I specify utility as additively separable in travel costs, store characteristics and utility from the bundle. I note that both differences in the shopping experience and availability of substitute products might result in utility from the bundle being different at different stores, so that people might "select" into particular stores depending on the bundle. To address this issue I propose a control function approach, where I allow the value of a purchase to vary by store and bundle characteristics. The approach exploits significant variation in bundle characteristics, and in the market I study it allows me to control for endogeneity of the expenditure coefficient.

In contrast to differentiated product markets, where the price of a product enters a consumer's utility as a product characteristic, the price level at a supermarket affects a consumer's utility in two ways: as a store characteristic, affecting a consumer's mean utility at a store, and as a bundle characteristic, affecting the expenditure on the bundle. While the issue of endogenous prices has been extensively addressed in the literature, to the best of my knowledge, this is the first paper to directly address endogeneity of expenditure. If unobservable store characteristics are correlated with price, then they are also correlated with expenditure; if the difference in utility from a purchased bundle varies across store and bundle characteristics, then expenditure will be endogenous as well. In the market I study variation over time in (relative) prices is very limited, and the inclusion of store fixed effects alleviates much of the issue of price endogeneity. Conversely, I note that price and store size are positively correlated. Thus a major concern arises if households prefer larger stores for larger bundles. By including an interaction between store size and "real quantity" purchased, I estimate the coefficient on expenditure controlling for bundle size.

The model predictions match very closely both the number of trips and the overall revenue at the stores. Implied substitution patterns are fairly rich, as they depend on the relative distribution of household locations and travel costs across different demographic segments, and on the relative shopping habits within these segments.[3] Distance plays a major role in a household's decision about which store to shop at, as people tend to concentrate their trips to stores that are close to their residence. Furthermore, people value stores differently and trade off additional travel time for better store characteristics. Individual store fixed effects reflect how consumers value both the location and the attributes of a store. Differentials in preference for stores correspond to a distance of up to 3.3 miles, or equivalently a driving time of 19.5 minutes, between the most and least preferred store in the market. Limited variation in store characteristics and relative prices hampers my ability to estimate the effect of price on the mean utility from a store. A small negative coefficient on bundle expenditure suggests that consumers are sensitive to price in proportion to their anticipated bundle. At the same time, selection is important, as consumers show a strong preference for variety and assortment, and tend to shop at larger stores for larger bundles.

The model's specification enables the study of the welfare consequences of various policy changes across different groups of consumers. I use the estimates of the multinomial logit specification to study the welfare impact, as well as the effect on trip and expenditure shares, of 1) the closing of each one of the stores and 2) the relocation of one of the stores in the market. These two counterfactuals exemplify a large number of questions that can be answered within this framework. As the welfare measures I

[3]Note that the IIA property holds only at the individual trip level, as both the bundle and the household's location and travel costs affect a consumer's decision.

construct are based on the assumption that consumers do not adjust to the change in the set of alternatives by changing their shopping habits, they should be considered an upper bound to the welfare loss, or alternatively a lower bound to the welfare gain. These measures should also be considered as short run, as I assume no response by stores.[4]

I find that the removal of a store produces significant welfare losses across all segments of consumers, ranging from 8.2% to 44.4% reduction in consumer surplus depending on the store considered. The impact however is extremely different across consumer segments, depending on differences in travel costs, geographic distribution, and distribution of bundle characteristics across consumer segments. The replacement of one of the stores has more subtle effects then its simple removal: not only will the original location affect how people substitute away from the store, but also the new location will affect people substituting to the store from different locations. By computing trip-specific choice probabilities and expected expenditures, I construct predicted expenditures and trips at different stores under the alternative counterfactuals. As the IIA property holds only at a trip level, substitution patterns across stores reflect the geographic distribution of consumers as well as the distribution of their bundle choices. Competition among stores as implied by the model fully captures the spatial dimension of the market and the observed heterogeneity in consumer habits.

This paper contributes to the large literature on retail markets and spatial competition by proposing a model of consumer store choice in the supermarket industry that allows to identify consumers travel costs and is suitable for policy analysis. Previous

[4]This might be a justifiable restriction in the industry: physical limitations prevent stores to change most of their characteristics in the short run, and pricing strategies in the supermarket industry are typically set at a higher level than the individual store, so that a supermarket's reaction to a change in local competition is very unlikely to radically alter its pricing strategies.

studies of consumer store choice have generally used various reduced-form approaches or have reduced the dimensionality of the problem by simplifying the choice of bundles. In the marketing literature the main focus is on retailers, usually considering segments of consumers, aggregate effects of choice and broad strategy questions.[5] Furthermore, the use of probabilistic models does not allow for policy analysis under counterfactual market conditions. [6]

Industrial organization studies of the supermarket industry have often limited their attention to how the spatial dimension affects market structure. Smith (2004), for example, measures the extent to which cross-elasticities between stores enhance market power by chain stores. Beresteanu and Ellickson (2006) analyze competition in the industry using a dynamic model of chain level competition. Ellickson (2007) studies the role of endogenous fixed costs. Griffith and Harmgart (2008) evaluate the effect on market structure and consumers' surplus of restrictive land use regulations without accounting for consumer travel costs. Supermarket choice has also been considered in a number of papers aimed at measuring the biases of standard consumer price indices. Hausman and Leibtag (2009) for example, study the choice across different formats to measure the outlet substitution bias of the CPI. Griffith, Leibtag, Leicester, and Nevo (2008) instead, look at the "timing" and "quantity" biases that arise from ignoring consumers' choice of how much and when to buy.

A few papers in the empirical literature on spatial competition have considered

[5] For example, Carpenter and Moore (2006) study the effects of consumer demographics on retail format choice; always in the context of store format, Reutterer and Teller (2009) study the role of the shopping occasion, operationalized by different trip types; Rhee and Bell (2002) consider cherry-picking and main store allegiance using transition probabilities. Ellickson and Misra (2008) estimate a discrete game of firms pricing strategy to explain how local market conditions and competitors' choices affect supermarket chains' pricing strategies.

[6] Applications of the Dirichlet and Negative Binomial models include Wrigley and Dunn (1984), Wrigley and Dunn (1985), Keng and Ehrenberg (1984); the Dynamic Markov model was introduced in Burnett (1973). Leszczyc, Sinha, and Timmermans (2000) propose a dynamic hazard model that accounts for timing and location using a factor analytic structure of store attributes and location.

consumer store choice, generally in contexts where the bundle of products to choose from is small. Davis (2006), for example, studies spatial competition and market power in the movie theater market. Using aggregate data on shares and the observed geographic distribution of consumers he extends the random coefficients model by Berry, Levinsohn, and Pakes (1995) to account for consumers' preferences for geographic proximity. Most closely related to the current work, Katz (2007) was the first to integrate micro-level data with driving time, and estimates consumer travel costs using a moment inequality approach. The model, however, does not allow to study consumer behavior under counterfactual conditions and is not suitable for policy analysis.

The remainder of the paper is organized as follows. I first provide a brief description of the industry in section 2. In section 3 I describe the data, provide descriptive statistics for the sample and discuss some of the patterns observed in the data. In sections 4 and 5, respectively, I introduce the general model and discuss the empirical implementation. I discuss the results in section 6. Section 7 presents the counterfactuals and section 8 concludes. Appendices A.1 and A.2 present the GMS procedure by Andrews and Soares (2010) as adapted here for inference using moment inequalities, and discuss construction of the price indices used in the paper.

1.2 The Industry

The supermarket industry is a multi-billion dollar business with sales of over $584 billion and 3.4 million employees in 2011. According to the Annual Retail Trade Report from the U.S. Census Bureau, sales at grocery stores accounted for 13.6% of total retail sales in 2010 (12.9% supermarkets only). This value was third behind the total retail sales of motor vehicle and parts dealers and general merchandisers, respectively accounting for 19.4%, and 15.8% of the market.

The Food Marketing Institute, the industry trade organization for food distribution and retail, defines a traditional supermarket as a store offering a full line of groceries, meat, and produce with at least $2 million in annual sales and up to 15% of their sales in general merchandise. In 2010, there were 36,569 supermarkets in the US carrying an average of 38,718 items. The median size of a store was 46,000 sq. ft., in a slightly decreasing trend since the peak of 48,750 sq. ft. in 2006. According to USDA estimates, Americans spent 5.5% of their disposable income on food-at-home in 2011. The average transaction expenditure in a supermarket was 26.78$, and consumers made an average of 2.2 trips per week to a supermarket.[7] Outside of the supermarket category fall convenience stores, the other main grocery store category, offering a limited line of high convenience items and accounting for sales lower than $2 million.

In the last decades innovations in the industry have brought to the emergence of a wide variety of (supermarket) formats including superstores, "fresh stores" and limited-assortment stores. Superstores are large supermarkets, with at least 30,000 sq. ft., generating annually $12 million or more in revenues, and offering an expanded selection of non-food items as well as specialty departments and extensive services. Fresh stores emphasize perishables and offer center-store assortments that differ from those of traditional retailers, especially in the areas of ethnic, natural and organic (Whole Foods, Publix GreenWise, The Fresh Market). Limited-assortment stores are low-priced grocery stores with limited assortment of center-store and perishable items (Aldi, Trader Joes, and Save-A-Lot).

Over the last few decades other types of retailers have started to offer food and grocery items. These non-traditional grocery retailers include mass merchandisers

[7]Source: http://www.fmi.org/research-resources/supermarket-facts.

and wholesale clubs. Mass merchandisers like Walmart, Kmart, and Target, are stores selling primarily hardlines, clothing, electronics, and sporting goods but also grocery and non-edible grocery items. Wholesale clubs are membership based retail/wholesale stores with a varied selection and limited variety of products presented in a warehouse-type environment usually carrying a grocery line dedicated to large sizes and bulk sales (Sams Club, Costco, BJs).

The supermarket industry today works on high-volume low-margins, with net profits of 1.09% in 2010, and is characterized by intense competition in price and services offered across and within a wide range of store formats. Modeling supermarket choice can be very hard, and requires a good understanding of the structure of local geographic markets and substitution across formats. The emergence of non-traditional grocery stores has made the task even more complicated, as multipurpose shopping makes it very difficult to identify what caused a consumer to choose a store.

1.3 The Data

I use the IRI Marketing Dataset (see Bronnenberg, Kruger, and Mela (2008)) which contains data on store sales and individual consumer purchases for 31 packaged goods categories for the years 2001 through 2007. The store sales data contains product sales, pricing, and promotion data for items sold in 50 U.S. markets. In two U.S. markets (Eau Claire, Wisconsin and Pittsfield, Massachusetts), the store level data are supplemented with panel-level purchase data and cover the entire population of stores. Further information is available regarding store characteristics such as location, type and estimated activity, as well as detailed household level information and market demographics.

As data on trips are available only for the micro level data, I restrict my attention to a panel of households in Pittsfield, Massachusetts. Using the longitude and latitude data for households residence and store location, I integrate the data by calculating distance and driving time to the stores using Google API (Application Programmable Interface). The dataset records all trips made by the households to the entire population of stores, including traditional supermarkets, convenience and drug stores, department stores and mass merchandisers. Due to confidentiality agreements, stores cannot be identified by chain name; however, each store is assigned with unique store and chain identifiers that allow to detect chain membership of a store.

Data on trips made by the panelists are separate form the category data, and include a store identifier, the time and day of the purchase and the total dollar amount spent. In some cases, trip files have multiple records very close in time for the same panelist at a given store; this could be for a variety of reasons, not necessarily indicating separate trips. As a rule of thumb I consider consecutive records within two hours as being one trip only. Panel data from the category files contain a product identifier, the quantity, price and store of purchase for all products sold within the categories at the weekly level. I recover the bundle of products purchased in a given trip to a store by aggregating all items bought by a household at the store in a given week; if a household visits a specific store more than once in a given week, I split the weekly expenditure across categories proportionally to the trip totals; that is, if a household bought a box of Cheerios for three dollars in one of two trips to a given store, the three dollars will accrue to the "cold cereals" category expenditure in the two trips proportionally to their relative totals. Fortunately a vast majority of household/week observations involve only one trip per week to a specific store.

In the national sample of store sales data 17604 items were sold in 31 categories and 50 markets. Of these, 4851 were sold in Pittsfield supermarkets, of which 4054 were sold in more than one store, and 3162 in more than one chain. Items sold in more than one store account for almost 96% of transactions and more than 99% of expenditures; items sold in more than one chain account for over 86% of transactions and almost 96% of expenditures. I use the aggregate sales data to construct individual trip specific price indices as a weighted average of category specific indices using as weights the trip share of expenditures in the categories. A separate appendix discusses the construction of price indices in more detail.

1.3.1 The Sample

By considering only panelists who are consistently active every month of a given year, and for which I possess detailed geographic and demographic information, I get a sample of 2362 households, making 285,309 trips in 2003. I focus on 2003 because it is the year in which the largest number of households are consistently active throughout the year. In the context of moment inequalities I further restrict the sample to a subset of 1213 households active for the entire seven years of the dataset to ensure a larger number of observations for each household. Although the number of observations and the variability in consumer demographics is different, the two samples are very similar in terms of the descriptive statistics and data patterns I now present for the larger sample.

Out of the 285,309 trips made in 2003, 259,460 were made to traditional supermarkets. Substitution to other retail formats is extremely low, especially to mass merchandiser stores, accounting for only 666 trips.[8] Department stores typically offer a combination of grocery products and general merchandise making analysis of store choice a lot more complicated; however, in the market analyzed most of grocery shopping is concentrated at traditional supermarkets, making it safe to exclude

[8]Note that the actual number of trips to department stores might well be larger than the number observed; trips to department stores, and other stores not in the national IRI sample of stores, are included in the sample only if they involve purchases within the 31 product categories.

department stores from the analysis. I also exclude the 25,091 trips were made to drug and convenience stores, accounting for below 3% of expenditure. Drug stores are not close substitutes to supermarkets; they are smaller and less assorted, and households planning purchase of a specific bundle rarely view a drug store as a viable option. Additionally, people visit drug stores for alternative reasons to that of buying groceries, and a large number of the trips observed might be driven by the need of prescription drugs or other contingencies.

A possible concern that might be raised is that of sample selection. In general, one might expect over-participation of price sensitive consumers seeking the benefits of participating into the program,[9] and under-participation of time sensitive consumers. If selection is based on unobservables, the magnitude and importance of the selection problem, and its impact on a model's estimates, are very hard to assess. Collection of the panel data by IRI is based on either showing a card at a participating retailer, or self reporting by means of a scanner. While participation in the sample can be time consuming for scanner panelists, making that of sample selection a possible concern, this is not the case for consumers using a card. Given the lower compliance rate of scanner panelists, IRI increased the use of cards over the years, and all the households in my sample were card panelists. Very similar to the use of a loyalty card, the use of the card is very simple, and the issue of sample selection, if present at all, is going to be marginal.

In the final dataset I observe 259,415 trips made by 2362 households to the seven supermarkets in Pittsfield. Throughout the paper I label the supermarkets with the letters A to G for confidentiality purposes; the alphabetical ordering chosen reflects the stores' estimated revenues, store A having the highest revenues. The seven stores

[9] Although I do not possess details about the IRI program, consumers participating in such programs are typically rewarded by the data collecting agency with reward points to be redeemed in merchandizing.

belong to four chains and are all participating stores in the IRI's national sample.

Table 1.1 provides descriptive demographic information as well as information on expenditure, number of weekly visits, distance and driving time to the store visited. The sample is fairly representative, with an average income of $48,000 and a median of $40,000; the median family size is 2, and the average size is 2.53; 21% of the household are composed of one person; 63% of the heads of a household are married, 51% are over 55 years of age, 5% is under 35; 26% of the households have young children. On average people drive 8.5 minutes and 3.1 miles to visit a supermarket. Average expenditure on a trip is $44.6, the median is $27.1. Households visit a supermarket on average 2.11 times a week, a their median and mean weekly expenditure are respectively $79.1 and $94.2. Figure 1 shows the geographic distribution of households and stores. Households in the sample are well representative of the relative population in Pittsfield census blocks, and the variability in households' driving time to the stores is extremely high.

1.3.2 Data Patterns

I now present some of the patterns observed in the data driving the results of my empirical implementation. Distance is a major factor in households' decision about which store to shop at, as people tend to concentrate their trips to stores that are close to their residence. This can be seen in panel (a) of Figure 2, which shows the distribution of driving time to a supermarket; a large portion of trips involve driving less than 10 minutes, and very few trips involve driving more than 20. This information however, does not clarify whether people actually drive to the closest stores. Panel (b) shows the distribution of the number of supermarkets visited by a household within a year; a vast majority of people shop at multiple stores, with only 5% of the households shopping in less the three stores, and over 80% at more than three stores; only 83 people in the sample, representing 3.5 percent of the sample, shop at only one store.

As the number of stores visited does not provide us with information about how people actually distribute their trips across different stores, panels (c) and (d) of Figure 2 look at the concentration of household trips across stores. Households seem to substitute between stores across different occasions and most households do not exhibit loyalty to a single store. Panel (c) shows the distribution of ranking of the trips in terms of driving time from a household's residence, rank 1 denoting a trip to the closest store. The mode of the distribution is 1, and households concentrate over 45% of their trips to the two closest stores. More informative about differences across individuals is panel (d), which shows the distribution of the household level Herfindahl Index of stores visited.[10] Not only do people shop at different stores, they also concentrate their visits at multiple stores. Over 60% of the households have an index value lower than 0.5, representing a consumer splitting trips evenly between two stores, and 33% have an index lower than 0.36, representing a 40-40-20 split across three stores.

Figure 3 looks at trips and expenditures in some more detail. Panels (a) and (b) show respectively the distribution of trip and weekly expenditure. Expenditures vary significantly across households and shopping occasions. This variation can help explain why people substitute between stores at different occasions: larger bundles might be associated to a stronger preference for variety and/or lower prices. People's decisions will reflect the trade-offs they face between prices, driving time and assortment, and people might spit their shopping needs across multiple stores. Panel (c) shows the distribution of the number of weekly visits to a supermarket. While

[10] I calculate the household level Herfindahl-Hirschman index as $HHI = \sum_i s_j^2$, where s_j denotes the share of visits made to store j out of all visits made by the household.

the mode of the distribution is one, a majority of household-weeks involves multiple trips. Panel (d) of Figure 3 presents how average expenditures differ across stores and income groups. Average expenditure appears to be increasing in income at all stores and varies significantly across stores. For all income classes average expenditure is highest at store A and lowest at stores F and G; for stores B through E the average expenditure is comparable.

Consumers are extremely heterogeneous in their shopping habits, in how often they visit a store and in how much they spend. These differences depend in part on consumers demographics, such as income and family size, in part on the choice set of stores, in terms of their relative location with respect to the households residence, and in part in heterogeneity in preference for shopping. Assessing whether the variation in household's habits depends on demographic factors rather than location or difference in preferences can be very important when analyzing the impact of a policy or other decision that affects the consumers' choice set. I here investigate the issue by means of reduced form regressions of average expenditure and average number of weekly trips on a number of observable demographics and a measure that captures how close (in terms of driving time) to the stores is a household location.

As a measure of a household's proximity to supermarkets I consider an index that: 1. decreases in a store's distance and 2. increases in the number of stores. Let d_{hs} denote driving time for household h to store $s = 1, ..., S$; I use

$$L_h = \sum_s \frac{1}{r_s} d_s^{-1/2},$$

where r_{hs} denotes the rank of store s in terms of distance, and a higher value indicates a better location.[11]

[11] I also considered the more general specification $L_h = \sum_s \frac{1}{r_s} d_s^{-\alpha}$. However, convex specification of travel costs ($\alpha \geq 1$) could not explain the observed variation in behavior. Conversely, the concave

Table 1.2 shows the results of the two regressions. A larger average expenditure is associated with higher income, larger families, home owners, married couples, well educated people and people with longer working hours. Larger bundles are generally associated with a lower number of trips, although large families and married couples tend to both spend more and more often. Higher values of the index on location (that is, a better location) are associated to more frequent trips and lower expenditure. Although the coefficients on income, family size, working hours and location are significant, the overall fit of the regressions is rather poor, suggesting that a major determinant of households shopping habits is heterogeneity in preferences.

1.4 The Choice Model

I now introduce a general model of consumers' shopping decision that makes some simplifying assumptions on the nature of consumers' utility, but is otherwise general in terms of how consumers choose what to buy and where to buy it. The model allows to analyze the determinants of store choice without making restrictive assumptions on the choice of the bundle. Throughout the rest of the paper I interchangeably use the words individual, consumer and household and denote individuals (households) by h to avoid confusion between observations and individuals.

Assume that only stores within a reasonable distance are in a consumers choice set and utility is additively separable in utility from bundle and store characteristics, expenditure on the bundle and driving time. What a reasonable distance is will be different for different households, and will typically depend on the transportation means available to them, as well as on other market and household characteristics. Characteristics that affect consumers' utility might be the size and assortment of a

specification chosen($\alpha = 1/2$) produces results that are very close in terms of fitting the data, to the use of a set of dummies with actual distances.

store, availability of parking space, number of employees, number of checkouts, an overall price level at the store.

Households periodically decide to visit a store to buy food, groceries and other household supplies. Whether they visit a store at all, and what they eventually buy depends on their preferences and on what they currently stock and plan to stock at their homes. In each period the consumer makes a choice whether to visit a store, which groceries to buy and where to buy them. Let the utility an individual h gets from buying bundle b at store s in time t be equal to:

$$u_{hsbt} = v_{hsbt} + \alpha_h e(b, p_{hst}) + \gamma_h(d_{hs}) + X_{hs}\beta + \xi_{hs} + \varepsilon_{hst}. \tag{1.1}$$

Total utility depends on utility from the bundle v_{hsbt}, which is different for different households and can vary across stores and shopping occasion. X_{hs} and ξ_{hs} represent respectively observed store characteristics and an unobserved component that affect the utility of a consumer. Utility depends also on expenditure on the bundle $e(b, p_{hst})$ and the time it takes to drive to the store d_{hs}. In a given time period, the utility of not visiting a store is a function of past purchases:

$$u_{h0t} = \sum_j f_j(b_{t-j}) + \varepsilon_{h0t} \tag{1.2}$$

The utility specification in (1.1) puts no restriction on utility from the bundle; bundle choice can be dependent across shopping occasions, as bundles bought in the past and available stocks might affect current choice, and consumers can split their shopping needs across multiple stores. In general, the choice set will not be the same for all consumers, as it depends on the consumers location and market characteristics. All else equal, consumers will prefer stores that are closer, cheaper and have better characteristics. Consumers might go to a more expensive closer store if the saving in travel time outweighs the increase in expenditure. Consumers might go to a more distant store for a larger bundle if the store is cheaper or better assorted. Consumers are heterogeneous in their preferences over stores and over bundles, and their preferences can vary over time.

Consumers are typically uncertain about prices at different stores and about the bundle they will end up buying. While in principle one could define expectation over all possible bundles and assume that the consumer bases his choice on that expected utility, this alternative is infeasible in practice, and ultimately unrealistic. Considering the entire distribution of bundles they might buy when visiting a store might be a far too complex decision process. Furthermore, estimating the choice of a store simultaneously with the choice of a bundle would require restrictive and unrealistic assumptions, so that misspecification of bundle choice is bound to happen and might result in severe biases of the estimates of travel costs and preferences for stores.

Assume instead that consumers make their choice based on a planned bundle. Although product prices and availability, stock-outs and promotions can lead consumers to buy substitute goods which were not used to formulate their store choice, specific needs will most likely guide their decision. If consumers make their choice of a store based on a bundle they are planning to buy, store choice is essentially a discrete choice problem, and the researcher can study this "conditional" choice using standard discrete choice techniques without imposing unrealistic restrictions on bundle choice.

The assumption that consumers base their store choice on a planned bundle might seem somewhat restrictive; however, we should think about the planned bundle as a "shopping list", indicating the products in a household's needs, but not specifying the individual brands. To the extent that stores in the market considered are fairly homogeneous, the approach naturally addresses issues of product availability, since even if a bundle observed was not available for purchase at another store, supermarkets will generally carry substitutes to the specific brand, so that a similar bundle will be available at the alternative store. As in any other discrete choice model, only differences in utility affect consumers' decisions, and to consistently estimate the determinants of store choice all that matters is to correctly pick up the differences in utility across stores: to the extent that observable bundle and stores characteristics allow us to capture this difference in utility, the estimates of mean utility from store characteristics and disutility from driving time will not be affected by the assumption of a planned bundle.

A commonly used assumption here would be to assume that utility from a given bundle is the same whether you buy it at a store or another. In this case the determinants of store choice would be independent of the bundle considered, and no further assumption would be necessary. I note, however, that utility from a given bundle might be different at different stores, either because of difference in the shopping experience (e.g. people prefer a larger store for a larger bundle), or because of the availability of substitute products. Therefore, people "select" into particular stores depending on the bundle they are planning to buy, and estimating a model without taking into account how different bundles affect consumers' decisions might result in biased estimates. To address this issue I propose a control function approach, where I allow the utility of a purchase to vary by store and bundle characteristics. The approach exploits significant variation in bundle characteristics, and in the market I study allows me to control for the endogeneity of the expenditure coefficient.[12]

Optimality of choice requires that a consumer planning to buy bundle b at time t

[12]The approach can be easily generalized to address additional selection biases arising in markets with less homogeneous competitors.

will choose to visit store s if and only if

$$u_{hsbt} \geq u_{hs'bt} \text{ for all } s', \text{ and } u_{hsbt} \geq u_{h0t} \tag{1.3}$$

Alternatively, a consumer will choose not to visit a store whenever

$$u_{h0t} \geq u_{hsbt} \text{ for all } s. \tag{1.4}$$

In the following section I introduce two alternative methodologies that, by adding assumptions to the general model, allow us to study how location, store characteristics and (eventually) bundle characteristics affect consumers choice of a store.

1.5 Empirical Implementation

The choice model presented above helps us understand the nature of consumers' decision process when they choose where to shop. However, the bundles and prices the consumers used to make their store choice are not the bundles and prices they end up buying and paying, which are what the econometrician observes, and further assumptions will be necessary. To make the model suitable for estimation, we need to specify how utility from a bundle differs across stores, and a relation between the bundle the consumer used to formulate his store choice and the bundle he actually purchased.

Specifying how utility from a bundle differs across stores is important to avoid the selection biases that might arise if either the shopping experience or the availability of substitute brands make differences in utility across stores vary in bundle characteristics. The appropriate assumption to use will typically depend on the specific application. In the market I study supermarkets are extremely homogeneous in terms of the product they sell, and I only need to worry about the shopping experience. In particular, as I discuss below in more detail, I observe that consumers show a strong preference for variety and assortment, and tend to shop at larger stores for larger bundles. Conversely, specifying a relation between the planned bundle and the bundle actually purchased allows us to include observable bundle characteristics in the analysis and to construct expenditure at the alternative stores. Note, however, that even assuming that the planned bundle and the bundle actually purchased coincide (i.e. no expectational error) only imposes conditions on those characteristics that affect a consumer's decision (e.g. total quantity, within categories purchases, etc.).

I now consider the empirical implementation under both standard discrete choice techniques and a less restrictive moment inequality approach. Both methodologies will use optimality conditions (1.3) and (1.4) to estimate demand for stores in a way that the bundle choice will not be part of the econometric implementation. I first consider a discrete choice specification that allows for a flexible model of consumers heterogeneity in travel costs and is suitable for policy analysis. As the model does not directly allow for measurement and expectational errors, I construct trip specific price indices and control for endogeneity of price and expenditure. To further support my results, I then introduce a less restrictive moment inequality approach that does not require a parametric distribution of the disturbances, and is robust against certain types of measurement errors and consumers expectational errors.

1.5.1 Discrete Choice

The discrete choice specification uses the necessary conditions of optimal store choice as a basis for estimation. The model requires the we specify a parametric distribution for the random part of the utility, and does not directly include measurement and expectational errors. I partially deal with these potential biases by constructing trip specific price indices and controlling for endogeneity of price and expenditure. The model additionally requires that we specify the utility of the outside option and the choice set of stores. Estimation involves finding the value of parameters for which the likelihood function is maximized. The ability to compute choice probabilities in turn, allows to use the estimates of the model to compute the choice probabilities under counterfactual market conditions, and evaluate the welfare impact of changes in the market structure.

By optimality conditions (1.3)-(1.4), a consumer planning to buy bundle b at time t will choose to visit store s if and only if

$$u_{hsbt} \geq u_{hs'bt} \text{ for all } s', \text{ and } u_{hsbt} \geq u_{h0t}$$

They will alternatively choose not to visit a store whenever

$$u_{h0t} \geq u_{hsbt} \text{ for all } s.$$

To make the model suitable for estimation I require the following additional assumptions:

DC1 Difference in utility from the bundle across stores depends only on observable store and bundle characteristics. That is

$$v_{hsbt} = v_{hbt} + f(b, X_s), \tag{1.5}$$

DC2 The utility of not visiting a store is a function of past purchases of a "real" grocery unit:

$$u_{h0t} = \sum_j b_j q_{t-j} + \varepsilon_{h0t} \tag{1.6}$$

DC4 The distribution of ε_{hst} belongs to the GEV class.

Under assumption DC1 I restrict difference in utility across stores to depend on on observable store and bundle characteristics. Assumptions DC2 and DC3 specify, respectively, the utility of the outside option and the distribution of the disturbances. As the model does not directly include measurement and expectational errors, consumers have no uncertainty, and the prices and bundles the econometrician observes are the prices and bundles consumers used to guide their choices. Note, however, that under assumption DC1, this is only imposing that bundle characteristics affecting consumers' decision are "observable" via the bundle purchased. On the other hand, the assumption that consumers make no expectational errors can be relaxed by imposing an orthogonality condition.

Let ε_{ht} denote the vector of disturbances ε_{hst} for $s \in S_h$, with joint distribution $f(\varepsilon_{ht})$. Under assumptions DC1-DC5, the probability consumer h planning to buy bundle b will visit store s is equal to:

$$P_{hsbt} = \int_{\varepsilon_{ht}} I(u_{hsbt} > u_{hs'bt} \forall s' \neq s) f(\varepsilon_{ht}) d\varepsilon_{ht} \tag{1.7}$$

Estimation By assuming that the structural error ε_{hst} is distributed extreme value, type I, the discrete choice model reduces to a multinomial logit which I estimate via Maximum Likelihood.[13] As a time unit I use a week, and construct an observation for weeks in which a household has not visited a store. As I do not observe a planned bundle for these weeks, I use a household-specific "average bundle" to construct utility in the stores in weeks with no visits. I also construct a variable specifying whether a trip is the first or a subsequent trip during the week and set the choice probability of

[13] I also considered nesting choices of a store separately from the outside option. The data however rejected this nesting structure with estimated λs of 1.

choosing the outside option to zero for subsequent trips within a week.[14] To specify the utility and choice probability of the outside option, I use store level price indices to construct quantities purchased of a real grocery unit in previous weeks.

The number of stores in the market is too low for consistent estimation of utility from store characteristics. Store dummies thus capture mean utility from both observable and unobservable characteristics. This limitation in the data inhibits my ability to infer a price elasticity and model the supply side. The model specification however is easily applied to larger markets where larger variability in store characteristics would allow to identify their effect on consumer utility.

I model consumers' heterogeneity as a function of observables. Heterogeneity enters the model in several ways. First, consumers are heterogeneous in their propensity to visit a supermarket. I specify utility from the outside option as linear in the real quantity purchased in previous weeks relative to the average quantity purchased by the household. Similarly, I use a household-specific average bundle to construct utility in the stores in weeks with no visits. Finally, and most importantly, consumers are heterogeneous in their disutility from driving to a store, not only because of differences in preferences, but also because of differences in location: in my actual specification I model disutility from driving time as quadratic,[15] with an household specific slope that varies with observable consumer demographics, that is:

$$\gamma_h(d_{hs}, z_h) = \gamma_0 d_{hs} + \gamma_1 d_{hs}^2 + \gamma_z' z_h d_{hs}.$$

The model in principle allows for unobserved consumer heterogeneity and the

[14]The procedure used here is equivalent to that used in case of product availability, where choice probability for a product which is not available is set to zero by specifying a price equal to infinity.

[15]The choice of a quadratic specification was driven by the data. I tried other functional forms like the square root and the log, but these specifications did not fit the data better than a simple linear specification. Conversely, the fit including a quadratic term improves significantly, and the estimate of the coefficient on the quadratic term is robust to the model's specification.

use of random coefficients. Note, however, that unobserved heterogeneity in travel costs and preference for stores cannot be separately identified, and a model including heterogeneity in consumers' preferences for stores might well be over-specified. Furthermore, availability of detailed micro data allows me to directly model heterogeneity in travel costs as a function of observable consumer characteristics and as a robustness check that unobserved heterogeneity in preference for stores does not affect my estimates, I estimate a model in which I introduce a dummy to account for an individual's experience at the store; as I will discuss later when commenting the results, the inclusion of such a dummy does not affect my estimates.

For each bundle b bought at store s in week t, I observe products purchased and expenditure $e_j(b, p_{st})$ within 31 packaged goods categories, plus a residual expenditure equal to the difference $e(b, p_{st}) - \sum_j e_j(b, p_{st})$. Goods purchased in individual trips are extremely different both in the category expenditure mix and in the within category product mix: different households not only allocate their spending differently across categories, but also choose different products within a given category. To construct expenditure at alternative stores in a way to better reflect a trip/bundle specific difference in expenditure, the construction of trip and household specific price indices seems most appropriate. In my main model specification I construct expenditure at the alternative stores for bundle b using category price indices $\tilde{p}_{jst}$, so that, for all $s' \neq s$

$$\tilde{e}_{bs't} = \sum_j e_j(b, p_{st}) \big(\tilde{p}_{js't} / \tilde{p}_{jst}\big).$$

To the extent that categories are sufficiently small and relative within category prices do not significantly vary across stores, the use of category level price indices is almost equivalent to the use of individual specific price indices. Even if this were not the case, category specific price indices should be preferred for the likely better quality of the data used in their construction. Individual households buy only a limited number of products, and are subject to continuous promotions making it very hard for the researcher to construct an index based on that limited information. Finally, the values for expenditure at alternative stores implied by different indices are very close, and the estimates are robust to the way I construct the indices. Details on the construction of price indices are presented in Appendix A.2.

A typical concern estimating demand in differentiated markets is that of price endogeneity when unobservable characteristics are correlated with price. A way to address this issue is the inclusion of store fixed effects capturing mean utility from store observed and unobserved characteristics. In the presence of time variation in unobservables and prices, however, fixed effects do not solve the endogeneity problem. Nevertheless, the time variation both in store level and store-category level price indices is very limited in the sample, and mainly due to trend and seasonal components. As a robustness check I also consider a model specification with time varying store effects, but the results are too similar to the more restrictive model to justify the extra computational burden.

The price level of a store affects the way consumers decide which store to visit in two ways: as a store characteristic, affecting a consumers mean utility at a store, and as a bundle characteristic, affecting the expenditure on the bundle. While in the absence of significant variation in price over time, the inclusion of store fixed effects addresses the endogeneity of price, expenditure might be endogenous as well. A major concern arises if households prefer more expensive stores for larger bundles; as counterintuitive as this might seem, if households prefer larger stores (more assortment) for their major shopping trips (higher expenditure) and larger stores are pricier than smaller stores, then the estimated coefficient on expenditure will result to be positive. As in the market I consider price and store size are positively correlated, I take a control function approach by including an interaction between store size and "real quantity" purchased to estimate the coefficient on expenditure controlling for bundle size.

Given all of the above, a consumers' utility of visiting store s at time t is given by:

$$u_{hsbt} = \alpha e_{bst} + \xi_s + \beta_q q_{bt} sz_s + \gamma_0 d_{hs} + \gamma_1 d_{hs}^2 + \gamma_z' z_h d_{hs} + \epsilon_{hst}.$$

where sz_s denotes the size of store s. The utility of the outside option is set equal to

$$u_{h0bt} = \begin{cases} \sum_j b_j \tilde{q}_{t-j} + \varepsilon_{h0t} & \text{first trip of the week} \\ -\infty & \text{otherwise} \end{cases}$$

I normalize the scale of the utility by setting equal to zero the constant term in the utility from the outside option. Under the assumption of an extreme value, type I, distribution for the error terms ε_{ist}, the choice probabilities have the closed form expression:

$$P_{hsbt} = \frac{\exp[\hat{u}_{hsbt}]}{\exp[\hat{u}_{h0bt}] + \sum_{s'=1}^{S} \exp[\hat{u}_{hs'bt}]}$$

As utility is linear in the parameters the log-likelihood function is globally concave and attains a unique global maximum.

1.5.2 Moment Inequalities

The use of discrete choice techniques have been recently criticized because expectational and measurement errors typically result in downward-biased estimates of price elasticities and over-prediction the impact of driving time.[16] To further sup-

[16]This might be the case, for example, if individuals shop at stores that are cheaper for the goods they are interested in- the presence and severity of the bias, however, depends on how one constructs

port the validity of the discrete choice specification, I introduce here a less restrictive moment inequality approach, first introduced in the context of supermarket choice by Katz (2007). The inequality estimator does not require a parametric distribution of the disturbances and is robust against certain types of measurement errors and consumers' expectational errors. Although the inequality approach does not allow to analyze the welfare impact of changes in the market structure, the methodology is a natural starting point to study the impact of driving time on store choice, and comparing its results to those of the more restrictive discrete choice specification will help us understand the implications of the assumptions used.

The idea behind the use of Moment Inequalities is that choice models generate inequalities which can be used as a basis for estimation. The typical advantages of such a methodology are that it does not require to restrict the agents' choice sets, as the researcher needs only to focus on a subset of reasonable alternatives, and it does not require a parametric distribution of the disturbances. Furthermore, by looking only at averages in differences in utility, the model can allow for expectational and measurement errors. The generality of this approach, however, comes at the cost of partial identification: while the actual estimates of the identified set can often be a singleton, inference consists in finding the boundaries of the identified region. I here present a model that is similar to Katz (2007) in terms of the optimality conditions used for estimation, but differs from it in the assumptions made on consumer choice.[17]

The inequalities model moves from the necessary conditions of optimal store choice to generate inequalities that are true for any store chosen and any alternative store no matter what the bundle chosen is. Focusing on the determinants of the store

the price index. See also Pakes (2010).

[17] While Katz's focus was the identification of consumers' travel costs, the focus of this paper is how these travel cost relate to consumers' preference for stores, and how this relation is affected by the nature of the bundle considered. As a result, my model allows the utility from a purchase to vary in bundle and store characteristics.

choice made by a consumer planning a particular bundle, the bundle choice will not be part of the econometric implementation. The researcher will need to consider only a small number of plausible alternatives rather then the entire set of stores. In fact, the choice set can be very large and will typically be different for different consumers, and having to specify for each consumer the choice set of stores, as well as the outside option can be problematic when data are unavailable for all the relevant stores.

As the model allows for expectational and measurement errors, I need to introduce some additional notation. Denote respectively $E_h[\cdot]$ and $\mathcal{I}_h$ the consumer's expectation operator and information set. We can rewrite optimality conditions (1.3) and (1.4) in terms of consumers' expectations. A consumer planning to buy bundle b at time t, will decide to visit store s if and only if

$$E_h[u_{hsbt}|\mathcal{I}_h] \geq E_h[u_{hs'bt}|\mathcal{I}_h] \text{ for all } s', \text{ and } E_h[u_{hsbt}|\mathcal{I}_h] \geq E_h[u_{h0t}|\mathcal{I}_h]$$

Consider the difference operator Δ and define $\Delta f_{dd'z} \equiv f(d,z) - f(d',z)$; then the inequality can be written as:

$$E_h[\Delta u_{hss'bt}|\mathcal{I}_h] \geq 0 \tag{1.8}$$

Denote (h, s, b, t) a trip made by consumer h to store s planning to buy bundle b in time t. For each observation in the data one can construct alternatives (h, s', b, t) given a rule for selecting s', and for every such rule one can construct a moment whose expectation is positive.

As product availability, pricing and promotions, will typically lead consumers to buy substitute goods which were not used to formulate their store choice, the planned bundle will typically be different from the bundle observed. However, consumers build into their expectations pricing and availability of products, so that difference between the planned bundle and the bundle observed is the result of an expectational error. Additionally, the econometrician only observes the prices actually paid by the consumer, possibly with some measurement error, and not the prices the consumer used to form his expectations.

Suppose now that the consumer was planning b^* but once in the store bought b; then consumer's optimal choice generates the following inequality conditions:

$$E_h[\Delta u_{hss'b^*t}|\mathcal{I}_h] \geq 0.$$

Denote respectively ε^e_{hst}, and ε^m_{hst} the consumer's expectational error and the econometrician's measurement errors; we have that

$$\Delta u_{hss'bt} = E_h[\Delta u_{hss'b^*t}|\mathcal{I}_h] + \varepsilon^e_{hst} \tag{1.9}$$

and

$$\Delta \tilde{e}_{ss'bt} = \Delta e_{ss'bt} + \varepsilon^m_{hst} \tag{1.10}$$

where $\Delta \tilde{e}_{ss'bt}$ denotes the observable difference in expenditure. To construct the moments using an observable version of inequality conditions (8) I require the following assumptions:

MI1 Difference in utility from the bundle across stores depends only on observable store and bundle characteristics. That is

$$v_{hsbt} = v_{hbt} + f(b, X_s)$$

MI2 Consumer have idiosyncratic preferences for unobserved store characteristics around a time and consumer invariant mean. That is

$$\xi_{hst} = \xi_s + \varepsilon_{hst}$$

MI3 The expectational error has mean zero conditional on the consumer's information set $\mathcal{I}_h$. That is

$$E_h[\varepsilon^e_{hst}|\mathcal{I}_h] = 0$$

MI4 The measurement error on prices has mean zero and is uncorrelated with any positive function $g(\cdot)$ of (instrumental) variables w_h in the consumers information set. That is

$$E[\varepsilon^m_{hst}|g(w_h)] = 0$$

Under assumption MI1 I restrict the difference in utility from the bundle across stores to depend on store and bundle observable characteristics. Assumption MI2 simply gives an interpretation to the idiosyncratic component in (1.1), without though specifying a distribution for it. Assumptions MI3 and MI4 impose conditions on the conditional mean respectively of the expectational error ε^e_{hst} and of the measurement error ε^m_{hst}. Under assumptions MI1-4 we can rewrite the utility in (1.1) as follows:

$$u_{hsbt} = v_{hbt} + f(b, X_s) + \alpha_h e(b, p_{hst}) + X_{hs}\beta + \gamma_h(d_{hs}) + \xi_s + \epsilon_{hst}.$$

Inequality condition (1.8) is then

$$E_h[\Delta u_{hss'bt}|\mathcal{I}_h] = E_h[\alpha_h \Delta e_{ss'bt}|\mathcal{I}_h] + \Delta f(b, X_{ss'}) + \Delta X_{ss'}\beta + \Delta\gamma_h(d_{hss'}) + \Delta\xi_{ss'} + \Delta\epsilon_{hss'} \geq 0$$

Denote $\Delta\widetilde{u}_{hss'bt}$ the difference in utility observable by the econometrician; then

$$\begin{aligned} E[\Delta\widetilde{u}_{hss'bt}|g(w_h)] &= E[\Delta u_{hss'bt}|g(w_h)] + \alpha_h E[\varepsilon^m_{hst}|g(w_h)] \\ &= E_h[\Delta u_{hss'b^*t}|\mathcal{I}_h] + E_h[\varepsilon^e_{hst}|\mathcal{I}_h] \\ &= E_h[\Delta u_{hss'b^*t}|\mathcal{I}_h] \geq 0. \end{aligned}$$

As $\Delta\widetilde{u}_{hss'bt}$ and w_h are observable, our moment for estimation is

$$E[\Delta\widetilde{u}_{hss'bt}|g(w_h)] \geq 0. \tag{1.11}$$

Estimation The model is partially identified and estimation focuses on finding the boundaries of the identifiable set. If the model is also linear, the identifiable set will be convex. Suppose $f(b, X_s)$ is linear in parameters denoted by λ. Estimation involves finding values of the parameters $(\alpha_h, \lambda, \beta, \gamma_h)$ along with mean utility from unobservables (ξ_s) such that the moment conditions are satisfied. If no such value of the parameters exists, similarly to a GMM setup, one picks the value of $\theta = (\alpha_h, \lambda, \beta, \gamma_h, \xi_s)$ that is "closest" to satisfying all the moments. For estimation of the identified set I follow Pakes, Porter, Ho, and Ishii (2011), whereas for specification testing and construction of the confidence sets I use one of the moment selection criteria considered in Andrews and Soares (2010). The idea behind moment selection is that we require that only the binding moments are considered for inference (note that at the boundary points of the identified set a subset of moments will be binding).

An observation in the dataset (h, s, b, t) is a trip made by household h $(h = 1, .., nh)$, at store s $(s = 1, .., S)$, buying a basket of goods b at time t. For each observation (h, s, b, t), one can construct alternatives (h, s', b, t) given a rule for selecting s'. Every rule thus generates a moment to be used for estimation. In general, depending on the rules used to construct the moments, the number of observations will vary across different moment conditions, and the researcher has to take this into account for inferential procedures.

I considered several sets of rules for constructing the moments, and I considered moments both at the population level and at the household level, similarly to a panel setup in a GMM context. To increase the number of trips observed at the household level I consider only households who are consistently active during the entire seven years of the dataset, so that in the final sample used for inequalities I observe 1213 households making over one million trips over seven years.

Let i denote an observation (h, s, b, t) in the data, and let y_i denote the measurable difference in utility $\Delta\widetilde{u}_{hss'bt}$. Consider rules $j = 1, ..., p$ for selecting alternatives (s'): the choice model implies that, for any positive functions $g(\cdot)$ of (instrumental) variables w_i

$$Em_j(y_i, \theta_0) \otimes g(w_i) \geq 0 \quad \text{for } j = 1, ..., p \text{ and } i = 1, ..., N \tag{1.12}$$

Let n_j denote the number of observations for moment j, and consider the empirical moment functions:

$$m_j(y, \theta) = \frac{1}{n_j} \sum_{i}^{n_j} m_j(y_i, \theta) \tag{1.13}$$

where

$$m_j(y_i, \theta) = \alpha_h \Delta\tilde{e}_{ss'bt} + \Delta f(b, X_{ss'}|\lambda) + \Delta X_{ss'}\beta + \Delta d_{hss'}\gamma_h + \Delta\xi_{ss'}.$$

If there exist a set of values of $\theta = (\alpha_h, \lambda_s, \beta, \gamma_h, \xi_s)$ such that the moment conditions are satisfied, the estimate of the identified set will consist of those values; formally, one picks

$$\hat{\Theta} = \{\theta : m(y, \theta) \geq 0\} \tag{1.14}$$

If no such value of the parameters exists, similarly to a GMM setup, one picks the value of θ that is "closest" to satisfying all the moments. Let $D_F(\theta)$ denote the matrix of diagonal elements of the covariance matrix of the moments. Also, let $(\cdot)_- = \min\{\cdot, 0\}$ and consider a consistent estimate $\hat{D}$ of $D_F(\theta)$. Estimation consists in finding (one dimension at the time) either

$$\hat{\Theta} = \arg\min_{\theta \in \Theta} ||(\hat{D}^{-1/2} m(y, \theta)_- || \tag{1.15}$$

or

$$\hat{\Theta} = \arg\min_{\theta \in \Theta} ||(m(y, \theta)_- || \tag{1.16}$$

Household level moments are constructed similarly. Let n_h denote the number of trips made by household h. For each trip i made by household h we have that

$$Em_{h,j}(y_i^h, \theta_0) \geq 0 \quad \text{for } j = 1, ..., p, \; h = 1, ..., nh. \tag{1.17}$$

where j denotes one of the alternatives above. The empirical moment functions in this case are:

$$m_{h,j}(y^h, \theta) = \frac{1}{n_h} \sum_{i}^{n_h} m_{h,j}(y_i^h, \theta) \tag{1.18}$$

While Pakes, Porter, Ho, and Ishii (2011) show that both estimation procedures in (1.15) and (1.16) lead to consistent estimates, the construction of confidence sets in inequalities setups can be computationally demanding and is still under debate in the literature.[18] Andrews and Soares (2010) introduced a generalized moment

[18]Numerous papers have considered inference and confidence sets for inequality models. A non-comprehensive list of theoretical papers in this area includes Andrews and Soares (2010), Andrews,

selection (GMS) procedure which provides confidence sets that are not asymptotically conservative. I adapt one of the selection criteria suggested by the authors to the case in which moments have a different number of observations, and introduce a way to characterize a $1-\alpha$ confidence interval by means of a grid search around the vertices of the estimated set. Appendix A.1 describes the procedure in more detail.

1.6 Results

I now present the results from the two alternative methodologies. I report the estimates of the multinomial logit model first and then compare them to those of the inequality approach.

1.6.1 Multinomial Logit

Table 1.3 presents the estimates from a variety of multinomial logit specifications. The table reports standard errors in parentheses to the right of the coefficient estimates. Under all model specifications I tried, the utility from the outside option is decreasing in the "real" purchases from previous weeks, as expressed by the negative estimates of $b_1 - b_4$. While ex-ante agnostic about the sign of these coefficients, the negative sign suggests that people are persistent in their habits. The probability that a household visits a store in a given week is increasing in previous weeks purchases. This is true even after controlling for individual characteristics.[19]

Mean utility from store characteristics is estimated very precisely, and moving from one specification to another only changes the scale, keeping the relative difference in mean utility between two stores, and thus the respective choice probabilities,

Berry, and Jia (2004), Andrews and Shi (2012), Beresteanu and Molinari (2008), Chernozhukov, Hong, and Tamer (2007).

[19] Models (1)-(5) use the relative quantity purchased by the household. Results using the absolute quantity (not reported here) were very similar.

unchanged. The estimates highly conform to my expectations, reflecting both the location of a store (in terms of the amenities close to it) and the value of its attributes. The model predictions match very closely both the number of trips and the model overall revenue at the stores (see tables 1.6 and 1.7).

Disutility from driving time is estimated to be negative in all specifications, and the quadratic term in models (2) through (5) is significantly positive, suggesting decreasing marginal travel costs consistently with Davis (2006). Consumers are significantly heterogeneous in their value of driving time. Travel costs decrease in income. as expressed by a negative estimate of γ_Y, entering the model interacting distance d with the inverse of income. Disutility from driving time is higher for larger families, families with no children, senior households, families with a high working load and families with a higher education level. Families with one or more children, young households, families with a low or medium work load and married couples have lower disutility. Identification comes from observed household behavior rather than from the differential effect on demand of the nearby population and the estimates are all significant and robust to the model specification.[20]

Driving time is expressed in minutes to go and come back from a store, and estimates of travel costs and utility at a store appear very reasonable. The average household would be willing to drive between 23.6 minutes for model (1), and 19.5 minutes for models (3) and (4), to visit store A rather than store G. As stores G and A are respectively the most and least preferred stores in the market, these numbers imply that differentials in preference for stores in the market correspond to a distance ranging between zero and up to 3.3 miles.[21] The numbers reported in the last row

[20]The only coefficient not estimated precisely is γ_{EDU2}, involving a residual group of households, not reporting their educational attainment, which I didn't want to include in the reference group (low education).

[21]The distance is based on the average trip made by the average household driving at a speed of 20 miles per hour.

of table 1.3 are constructed using store G as a starting location, so that the actual values will be smaller moving away from that location because of decreasing marginal travel costs.

Most problematic is the estimation of the coefficient on expenditure on the bundle. As consumers should consider the relative expenditure across stores for a given bundle, a positive α cannot be explained in economic terms. Conversely, consumers persistently shop at different supermarkets, so that a valid model of store choice should yield a small but negative estimate of α. The issue here might be one of endogeneity of expenditure; as in the market I consider price and store size are positively correlated, the positive estimate of α in models (1) and (2) likely arises from omission of a "taste for assortment" when buying larger bundles. In model (3) I include an interaction between store size and "real quantity" purchased to estimate the coefficient on expenditure controlling for the size of a bundle.[22] The negative estimate of α when controlling for bundle size is robust to the model specification, and all coefficient estimates are robust to the choice of the price index.[23]

The effect of the price level of a store is fully captured by the store fixed effects, ξs. An endogeneity issue arises if variation in the price level over time is correlated with a store's unobservable characteristics; time variation in price however, is very limited in the sample, and is uncorrelated with households' characteristics and driving time to the stores (which are time invariant). As a robustness check, model (4) reports the estimates of a model with time varying store effects. The model is identical to model (3), but I now spit the sample in ten 5-weeks periods and allow store effects to vary over these time periods. The resulting estimates are almost identical to those

[22] I could have used here alternative functional forms. However, in the market analyzed variability in store size is very limited, and the use of an interaction term produces results equivalent to the use of dummies for size categories, provided monotonicity in their coefficients.

[23] This is not surprising, since households persistently shop at different stores and different indices results in very small changes in relative expenditures.

of model (3). For model (4), the table reports the mean values of store effects over time; variation in the individual values (not reported here) is very limited, and does not suggest any systematic change (over time) in the relative utility from the stores.

A final issue regards households' heterogeneity in preference for the stores. The model allows for heterogeneity both over time and across household via the i.i.d. error term ε_{hst} (note that $\xi_{sh} = \xi_s + \varepsilon_{hst}$), but violation of the i.i.d. assumption might bias the estimates of the driving time coefficients. A natural way to address this issue would be the use of random coefficients; while feasible in principle, the inclusion of random coefficients would be computationally burdensome, and the model might well be under-identified. Instead, as a robustness check, model (5) presents the estimates of a model in which I include a very specific form of heterogeneity; I introduce a dummy to account for an individuals experience at the store, taking the value of one if an individual has visited the store at least ten times. Overall estimates of such a model should be disregarded because of dependence of the right hand side variable (the dummy) on the dependent variable (the store choice). However, the signs and magnitudes of coefficients closely match the estimates from the other models, suggesting that, if there is indeed unobserved heterogeneity in consumers' preference for stores, it ultimately does not affect the other coefficients[24].

1.6.2 Inequalities

Table 1.4 presents the estimates from a parsimonious model specification in which households are homogenous in both their travel costs (that is $\gamma_h = \gamma$), and decide to

[24]Not surprisingly the value of the likelihood function increases significantly, and the coefficient on experience "outweighs" the store fixed effects. This specification however, cannot be used in counterfactual analysis because of the endogeneity of the experience coefficient; if a store for which an individual has experience is removed or moved to another location it is unreasonable to assume that the individual will not gain experience in another alternative place, and/or "reset" his experience at the store that is being moved.

shop at differentiated stores where mean utility from store characteristics, captured by store fixed effects ξ_s, for $s \in \{A, B, ..., G\}$, is the same across households.[25] Given the limited number of alternatives which could be used to construct the moments, it was only by comparing store pairs that I was able to identify store fixed effects. Under this specification I use all the visits made to each individual store s to construct 7-by-6, 42 moments, using as alternative stores s', for each store $s' \neq s$. Under this model specification the estimate of the identified set turned out to be a set.

Table 1.4 reports the highest and lowest values in the identified set for each of the parameters, and 95% confidence bounds are reported in parenthesis. The last row of the table further reports the implied additional time a person would drive to visit store A rather than store G. As a level normalization I set the utility from visiting store A equal to zero, and as a scale normalization I set the coefficient on expenditures α to -1.[26] A more detailed description of the identified set and more details on the characterization of confidence bounds for set estimates are discussed in Appendix A.3 and Table A.1.

The estimates of the model conform to expectations, with a significantly negative utility from driving time, and a magnitude that seems reasonable in relation to mean utility from store characteristics. These values are very similar to those found with the multinomial logit specification, and the two models agree on how consumers trade off their travel costs for better store characteristics. In particular, the estimates of the inequality model suggest that a household would be willing to drive between 8 to 19 additional minutes to visit store A rather than store G. The upper bound

[25] I also considered alternative rules to construct the moments, heterogeneous households, consumer types, and individual household coefficients. I tried to instrument moments using a constant, income, family size, children, working hours, marital status, age, education level and positive transformations of them. Flexible specifications however, resulted in unbounded estimates or rejection of the model.

[26] This form of scale normalization is necessary because, differently from discrete choice techniques, there is no variance of the error term scaling the utility.

of 19 minutes is extremely close to the value of 19.5 minutes from the multinomial logit. Furthermore, the quadratic specification of travel costs in the multinomial logit implies that the value from a household's residence would be lower, and the inequality model both scales the utility by normalizing the coefficient on expenditure and does not control for bundle size, which could result in overestimation of travel costs and, therefore, underestimation of the store G to A value.

The estimates of store fixed effects capture some aspects of the data. Recall that the labeling of stores A-G corresponds to the estimated overall revenue at the stores. Revenue at a store is determined both by its characteristics and its location; we expect values of mean utility from a store to reflect revenues, even if not monotonically. Consistently, we observe that stores A, C, D and E are preferred to stores F and G everywhere in the set. These estimates however, provide too little information on how people value these stores, and the assumption of homogeneous consumers seems too restrictive. I further considered weighting observations by a household's number of trips, or using instruments to further restrict the estimate of the identified set, but both directions resulted in no parameter value satisfying the moment conditions and rejection of the model.

1.7 Counterfactuals

The multinomial logit model can be used for a wide range of policy analysis regarding zoning regulations, and can also be applied to the study of local competition and store location decisions. I use the estimates of model (3) (reported in table 1.3) to evaluate the welfare impact, as well as the effect on trip and expenditure shares, of 1) the closing of each individual store in the market and 2) the relocation of store B to a likely new location 6 miles away north-east of the original location. These two counterfactuals exemplify a large number of questions that can be answered within this framework.

Denote by i a trip (h, s, b, t). Under a multinomial specification the expected consumer surplus takes the convenient closed form

$$E(CS_i) = \frac{1}{\alpha_{Y_h}} \log \big(\sum_{s \in S \cup 0} \exp[\hat{u}_{hsbt}] \big) + c,$$

where α_{Y_h} is the marginal utility of income of household h, $\sum_{s \in S \cup 0} \exp[\hat{u}_{hsbt}]$ is the denominator of the logit choice probability, and $\hat{u}_{hsbt}$ for $s = 0$ denotes the expected utility of the outside option.[27]

Under the logit specification it is straightforward to analyze the impact on consumer surplus of a change in the set of alternatives, or in one of the attributes of one of the alternatives. Let the superscripts 0 and 1 denote respectively before and after a change; then, for any trip i, the percentage change in consumer surplus is given by:

$$\%\Delta E(CS_i) = \log \big[\big(\sum_{s \in S \cup 0} \exp[\hat{u}^1_{hsbt}] \big) / \big(\sum_{s \in S \cup 0} \exp[\hat{u}^0_{hsbt}] \big) \big] - 1 \tag{1.19}$$

For each trip I construct after-the-change choice probabilities and expected expenditures at each of the alternatives $s = A, ..., G$ by using the standard logit formulas:

$$P^1_{i,s} = \exp[\hat{u}^1_{hsbt}] / \big(\sum_{s' \in S \cup 0} \exp[\hat{u}^1_{hs'bt}] \big) \tag{1.20}$$

and

$$E^1(e(b, p_{st})) = \tilde{e}_{bst} P^1_{i,s} \tag{1.21}$$

[27] Rosen and Small (1981) show the result when the error terms are iid extreme value and utility is linear in income (details can also be found in Train (2009)). All results that follow measure percentage changes in consumer surplus that are independent of income elasticity.

where $P^1_{i,s}$ denotes the probability that household h visits store s after the change.

Using (1.19)-(1.21), I construct aggregate measures of consumers welfare change as well as change in the number and share of trips and change in expenditure shares. I do so by summing expected consumer surplus, choice probabilities and expected expenditures over the trips I observe keeping the bundles fixed. I evaluate the welfare change overall, and by household types defined by income, family size, number of children,working hours code, age, education and marital status.

The welfare measures I construct are based on the assumption that consumers do not adjust to the change in the set of alternatives by changing their shopping habits, such as the frequency of their trips and their average expenditure. As long as individuals react to a change in the set of alternatives the measures I provide should be considered an upper bound to the welfare loss (or alternatively a lower bound to the welfare gain) due the removal of (each) one of the stores, or the replacement of store B.

These measures should also be considered as “short run”, as I assume no response on part of the stores. While physical limitations do not allow stores to change most of their characteristics at least in the short run (a store for example cannot significantly change its size), it is possible that a store would reconsider its pricing strategies in response to a major change in the number of visitors. Although the model I estimate could be used to model the supply side of the industry, for example by adding a pricing equilibrium condition, the limited variation in store characteristics in my sample hampered my ability to estimate price elasticities. Note however, that pricing strategies in the supermarket industry are typically set at a higher level than the single store and for wider areas than the small town of Pittsfield. Additionally, a supermarket’s response to a change in local competition is very unlikely to radically alter its pricing strategies.

Table 1.5 presents the welfare impact overall and on different segments of consumers under both counterfactuals. Columns labeling A through G refers to which store has been removed, column B-B' refers to the replacement of store B to a new location. The removal of a particular store affects households differently; households who live closer, shop more frequently and spend more in a given store will suffer more than those who live further away and shop there less frequently. As households are otherwise homogeneous in their preference for stores, difference in the welfare loss from store removal across groups is fully explained by one or more of the following reasons: 1) difference in travel costs; 2) difference in location; 3) difference in size of the planned bundles.

The last row of table 1.5 shows the overall welfare impact of store removal on the population considered. Removal of store B produces the largest welfare loss (44.4%), followed by removal of stores A (43.9%) and C (40.3%); removal of the other stores produces a much smaller loss. The overall loss is determined mainly by the number of trips at the stores, being stores B, A and C respectively the most visited stores. The welfare effects however, vary significantly across household groups.

The first three rows of table 1.5 show the effect on households who's income is below 25,000$ (low), between 25,000$ and 65,000$ (medium), and over 65,000$ (high). Wealthier families are more affected than low income households from the removal of store A (51.4% vs. 37.3%), while the opposite goes for store B (38.7 vs. 48.9). This difference is due to the combination of two factors: first, wealthier families buy larger bundles (and preference for store A is stronger for larger bundles); second, wealthier families are located relatively closer to store A.

Differences across households grouped by family size and age groups are mainly driven by travel costs. With a few exceptions from the removal of stores C, E and G, welfare losses are higher for larger families; similarly, young households have a lower welfare loss under all the counterfactuals relative to adult and senior households. While families with children have generally lower travel costs, relative welfare effects differ significantly depending on the store removed, because of location and shopping habits. Families with children, married couples, households of higher income and with higher education suffer relatively more from the removal of store A and relatively less from the removal of stores F and G; these differences are driven by the relative shopping habits of these groups, rather than their travel costs.

Replacement of store B to a new location produces significantly lower welfare losses than its removal. The overall welfare loss from its replacement to a new less convenient location is 26.9% versus the 44.4% loss from its removal. The distribution of these losses across households however, does not match that of the removal of the store. Changes in the ranks of losses are due to the distribution of household location and on how the additional (or reduced) distance differently affects the households.

Tables 1.6 and 1.7 report observed and predicted expenditures and trips, both before and after the removal of one of the stores. The first two rows of both tables report respectively the actual and predicted measures. Although the logit specification matches choice probabilities, and not expenditures, predicted expenditures in the baseline match pretty closely observed expenditure in the data; as expected, the match is extremely precise for trips.

Rows indexed A through G report counterfactual trips and expenditures predicted after removal of the corresponding store. Note that the IIA property here holds only at the household level and for a given bundle, as both planned bundle size and location affect a consumer's decision. Reading the numbers along the columns highlights the gains for each store from the removal of one of the competitors. Store A benefits more from the removal of store C than B, and has a very small benefit from the removal of the other stores. Store B benefits almost equally from the removal of stores A and C in terms of trips, but the increased expenditure from people substituting away from store A is much larger. For stores D through G the removal of store B is way more beneficial than that of store A.

These results appear even more clear in tables 1.8 and 1.9. Table 1.8 reports the change in the share of trips under the different counterfactuals. Reading along a row will express how the trips originally made to one store are substituted to other stores; reading along a column expresses how the removal of a competitor affects a particular store. Similarly, table 1.9 reports effects on the overall revenue shares.

Table 1.10 summarizes for the effect of the replacement of store B to a new location. The replacement of a store has more subtle effects then its simple removal: not only will the original location affect how people substitute away from the store, but also the new location will affect people substituting to the store from different locations. Again, both location and bundle size contribute to the results in table 1.10. Looking at the change in the share of trips, we see that stores C and D have the highest gains followed by store A, store B looses a big portion of its original share, and all other stores have a substantial gain. The effect on expenditures share however is (not) surprisingly different; store C, who had the highest increase in trips, gains less than stores A and D.

I finally compare the welfare measures and the counterfactual visits and expenditures reported above for the full model with heterogeneous consumer travel costs, to those implied by a simpler model with homogeneous consumers and no control for the size of the bundle. The comparison helps to better understand how the inclusion of heterogeneity, and controlling for bundle size, both yield a significantly richer welfare analysis and provide more realistic substitution patterns.

Tables 1.11 compares the welfare impacts under the two counterfactuals using the estimates of models (1) and (3) from table 1.3. Looking at the overall impact, we find that the welfare losses implied by model (1) are significantly lower than those implied by model (3). The difference is presumably due to linear specification of travel costs in model (1), ignoring the decreasing nature of marginal travel costs, and the omission of a control for the bundle size, that in the larger model allows consumers to "select" to a more suitable store depending on their planned bundle.

Table 1.11 also looks at welfare differentials between the two models across income classes and family size. Not surprisingly, differentials implied by model (2) are significantly smaller in most cases. For example, the welfare loss from the removal of store A ranges from -35.7% to -46.5% for model (1) and from -37.3% to -51.4% for model (3); as the estimates from latter model suggest travel costs which are decreasing in income (that alone would imply a smaller gap), the wider gap in the welfare losses is entirely due to the inclusion of the control for bundle size. Conversely, when we look at family size, it is the inclusion of heterogeneity that leads to the larger gap in model (3) predictions. For example, the welfare loss from the removal of store A increases in family size from -38.1% to -41.9% for model (1) and from -36.6% to -51.7% for model (3); this is because not only larger families do buy larger bundles, but also experience higher travel costs according to model (3) estimates.

Table 1.12 looks instead at the fit of the two models and their implied substitution patterns. Both models are very accurate in matching the number of trips observed in the data, but produce significantly different predictions for expenditure shares. In particular, model (1) does a very bad job at matching expenditures at store A, as the model cannot explain why people buying larger bundles select to that store from longer distances. The second part of table 1.12 reports the "normalized" change in the share to make possible comparisons between the two models. Substitution patterns appear significantly different between the two specifications; for example, both models suggest that after the removal of store A, 72% of expenditure at that store would be substituted to stores B and C, but the two models disagree in how this expenditure will be substituted to the two stores. As model (1) ignores the heterogeneity in consumer travel costs and shopping habits, the richer specification of model (3) is able to capture more features of the market and should thus be preferred as long as these features produce significantly different substitution patterns.

1.8 Conclusions

Policy oriented empirical analysis of retail markets requires understanding how travel costs affect consumers' decisions. Empirical challenges in the context of supermarket choice make the use of standard techniques infeasible or require restrictive assumptions on bundle choice. This paper examines two alternative methodologies that allow to abstract from bundle choice, and develops a multinomial logit specification that allows for a flexible model of consumers' heterogeneity in travel costs.

Adding to the existing literature on spatial competition, variation in demand and substitution patterns across stores are explained not only by differences in store characteristics and the geographic distribution of consumers, but also by the shopping habits and relative geographic distribution of household demographic types. Using micro level data on store and household locations, consumers' heterogeneity in travel costs is directly identified as a function of observable consumer characteristics. By specifying utility as a function of both store and bundle characteristics, I control for the endogeneity of expenditure on the bundle, and this flexible specification can be generalized to address additional selection biases arising in markets with less homogeneous competitors.

Consistently with Davis (2006) I find that travel costs are quadratic and marginally decreasing. Both the quadratic term and heterogeneity parameters are robust to a variety of model specifications. Identification however, comes from the observed differences in behavior across geographically dispersed household types, rather than from the differential effect on demand of the nearby population. Store fixed effects reflect how consumers value both the location and the attributes of a store. Households are sensitive to price in proportion to their anticipated bundle. However, they seem to substitute between stores across different shopping occasions, and show a strong taste for variety and assortment, as they tend to shop at larger stores for larger bundles.

The model is suitable for a wide range of policy experiments, and allows to assess the welfare effects of a change in the choice set across different groups of consumers. The model can also be applied to the study of local competition and store location decisions, as it allows to study the effect on store revenues and customer visits under realistic substitution patterns. The two counterfactual experiments conducted exemplify a large number of questions that can be answered within this framework. Supply side considerations could be added to the model in the study of larger markets. Higher variation in store characteristics allows estimation of price elasticity of demand at a store, leading to a better understanding of competition between (chain) stores and of the effect of mergers in local spatially differentiated markets.

Tables and Figures

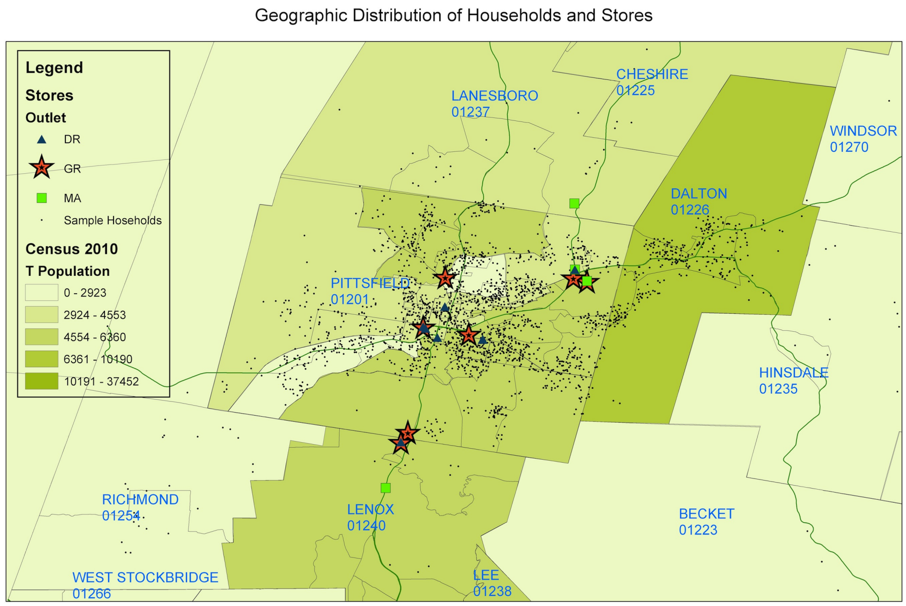

Figure 1.1: Distribution of Households and Stores

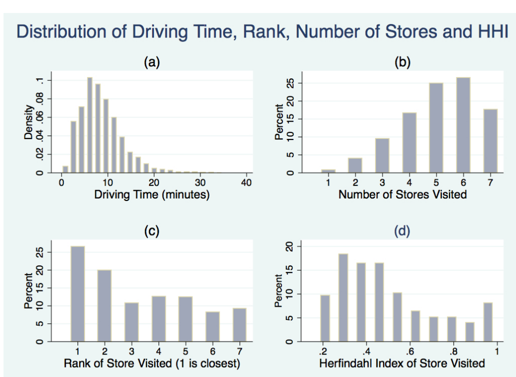

Figure 1.2: Distribution of Driving Time to a store (a), Number (b), Rank (c) and Herfindahl Index (d) of Stores visited by households in a year

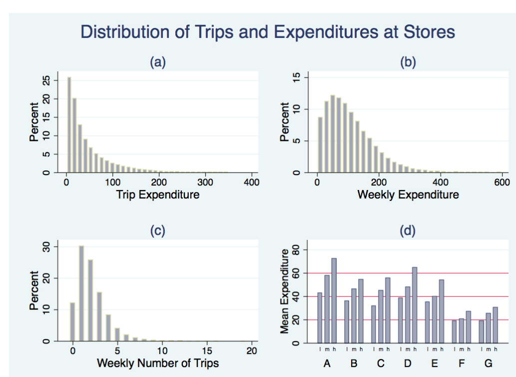

Figure 1.3: Distribution of Trip (a) and Weekly (b) Expenditures, Weekly Trips (c) and Average bundle expenditure (d) at stores by Income (low, medium, high)

Table 1.1: Pittsfield Sample Descriptive Statistics

	Obs	Median	Mean	S.d.	Min	Max
Weekly Trips	122,824	2	2.11	1.67	0	19
Weekly Expenditure	122,824	79.12	94.27	81.30	0	1268.73
Trip Expenditure	259,415	27.07	44.63	48.76	0.5	982.02
Distance (miles)	259,415	2.76	3.07	2.07	0	20.30
Driving (minutes)	259,415	7.88	8.59	4.55	0	41.55
Family Size*	2362	2	2.53	1.26	1	6
Income**	2362	40	48.78	31.19	5	125
Married	2362	1	0.63	0.48	0	1
Children	2362	0	0.26	0.44	0	1
Over 55	2362	1	0.51	0.5	0	1
Under 35	2362	0	0.05	0.21	0	1
Alone	2362	0	0.21	0.41	0	1

Notes: The table provides descriptive statistics for the larger sample of 2,362 households.
[*] Family Size takes value of 6 when there are 6 or more members in the family.
[**] Households report an income band they belong to; while bands are "tight" for low income, the highest income band reports all households with an income higher than 100,000 dollars. To compute the values reported here I use the mean value of a band and 125 thousand dollars for the highest band.

Table 1.2: Reduced Form Regressions

Dep. Variable	(Avg. Expenditure)	(Avg. Trips)
Constant	34.323**	1.958**
	(4.63)	(0.19)
Income	0.120**	-0.002
	(0.02)	0
Family Size	4.966**	0.104**
	(0.67)	(0.03)
Child	1.756	-0.333**
	(1.84)	(0.08)
House	3.804*	-0.144*
	(1.6)	(0.07)
Married	3.166*	0.176**
	(1.47)	(0.06)
Over55	-4.472**	0.086
	(1.62)	(0.07)
Under35	-5.651*	-0.059
	(2.77)	(0.12)
Edu 1	0.711	-0.098
	(1.22)	(0.05)
Edu 2	11.097*	-0.359
	(4.47)	(0.19)
Work 1	6.015**	-0.191**
	(1.66)	(0.07)
Work 2	6.469**	-0.266**
	(1.69)	(0.07)
Work 3	-7.644	0.527
	(7.54)	(0.32)
Location Index	-11.340**	0.358**
	(2.59)	(0.11)
R-squared	0.2	0.05
N	2362	

Notes: results from reduced form the reduced form regression of household average trip expenditure and number of weekly trips on consumer demographics and a Location Index. Income and Family Size are numerical variables. The other demographic variables are dummies that take value of 1 if a household belongs to the specific class. Edu refers to a households education (0=low, 1=high, 2=unkown); Work refers to a household's workload (0=low, 1-med, 2=high, 3=unknown). The Location index used is $L_h = \sum_s 1/r_s d_s^{-1/2}$ and denotes a household's proximity to supermarkets, a higher value indicating a better location.

Table 1.3: Estimation Results: Multinomial Logit Models (1)-(5)

model	(1)	(2)	**(3)**	(4)[a]	(5)
b1	-0.002 (0.000)	-0.002 (0.000)	**-0.002** (0.000)	-0.002 (0.000)	-0.001 (0.000)
b2	-0.003 (0.000)	-0.003 (0.000)	**-0.003** (0.000)	-0.003 (0.000)	-0.002 (0.000)
b3	-0.002 (0.000)	-0.002 (0.000)	**-0.002** (0.000)	-0.002 (0.000)	-0.002 (0.000)
b4	-0.002 (0.000)	-0.002 (0.000)	**-0.002** (0.000)	-0.002 (0.000)	-0.002 (0.000)
ξ_A	1.050 (0.025)	1.399 (0.026)	**1.274** (0.026)	1.282 (0.040)	-2.371 (0.028)
ξ_B	0.480 (0.024)	0.774 (0.025)	**0.737** (0.025)	0.748 (0.039)	-2.409 (0.026)
ξ_C	0.529 (0.024)	0.850 (0.025)	**1.011** (0.025)	1.022 (0.039)	-2.193 (0.026)
ξ_D	0.240 (0.025)	0.574 (0.026)	**0.549** (0.026)	0.559 (0.043)	-2.403 (0.027)
ξ_E	-0.446 (0.026)	-0.118 (0.027)	**0.089** (0.027)	0.099 (0.047)	-2.437 (0.028)
ξ_F	-0.755 (0.025)	-0.497 (0.025)	**-0.272** (0.025)	-0.261 (0.043)	-2.620 (0.026)
ξ_G	-1.001 (0.025)	-0.742 (0.025)	**-0.540** (0.025)	-0.528 (0.046)	-2.750 (0.026)
γ_0 (d)	-0.091 (0.000)	-0.119 (0.001)	**-0.117** (0.001)	-0.117 (0.001)	-0.041 (0.001)
γ_1 (d^2)		0.001 (0.000)	**0.001** (0.000)	0.001 (0.000)	0.000 (0.000)
γ_Y		-0.064 (0.007)	**-0.062** (0.007)	-0.062 (0.007)	-0.047 (0.008)
γ_{FS}		-0.006 (0.000)	**-0.006** (0.000)	-0.009 (0.001)	-0.004 (0.000)
γ_{CH1}		0.009 (0.001)	**0.009** (0.001)	0.009 (0.001)	0.003 (0.001)
γ_{CH2}		0.009 (0.001)	**0.010** (0.001)	0.010 (0.001)	0.004 (0.001)
γ_{AGE1}		0.014 (0.001)	**0.014** (0.001)	0.014 (0.001)	0.001 (0.001)
γ_{AGE2}		-0.006 (0.001)	**-0.006** (0.001)	-0.006 (0.001)	-0.002 (0.001)
γ_{WR1}		0.003 (0.001)	**0.002** (0.001)	0.002 (0.001)	0.003 (0.001)
γ_{WR2}		0.004 (0.001)	**0.003** (0.001)	0.003 (0.001)	0.004 (0.001)
γ_{WR3}		-0.013 (0.003)	**-0.012** (0.003)	-0.012 (0.003)	-0.020 (0.003)
γ_{EDU1}		-0.001 (0.001)	**-0.001** (0.001)	-0.001 (0.001)	-0.000 (0.001)
γ_{EDU2}		-0.000 (0.002)	**-0.002** (0.002)	-0.002 (0.002)	0.004 (0.002)
γ_{MRD1}		0.005 (0.001)	**0.005** (0.001)	0.005 (0.001)	0.003 (0.001)
$\beta_q(\times 1000)$			**0.150** (0.002)	0.150 (0.002)	0.158 (0.002)
α	0.005 (0.000)	0.004 (0.000)	**-0.004** (0.000)	-0.004 (0.000)	-0.005 (0.000)
Experience					2.930 (0.008)
Loglikelihood	-454974	-452932	**-450453**	-449,741	-349509
N			**274254**		
store G to A (minutes)[b]	23.6	18.8	**19.5**	19.5	-

Notes: MNL models (1)-(5). γs are the driving time parameters, interacted with the inverse of Income (γ_Y), family size (γ_{FS}) and a set of demographic dummies. $CH1$ denotes having one young child, $CH2$ having two or more children. ξs denote store fixed effects. β_q interacts real quantity of the bundle with size of a store; coefficient and standard error are multiplied by 1000. b_{1-4} reflect the effect of normalized previous weeks purchases on utility from the outside option. Experience is the parameter on a dummy taking value of 1 if a households has visited the store at least 10 times. α is the coefficient on expenditure.

[a] In model (8) store effects are time varying. Mean Values are reported here.

[b] "store G to A" is the average additional time a person would drive to visit store A rather than store G. When a quadratic term is included the value is reported using store G as a starting location (moving away from it would decrease the measure).

Table 1.4: Estimation Results: Moment Inequalities

	lb.		ub.	
α	**-1**	-	**-1**	-
γ	**-1.85**	(-1.87)	**-1.75**	(-1.74)
ξ_A	0	-	0	-
ξ_B	**-25.56**	(-25.64)	**-8.60**	(-8.52)
ξ_C	**-13.70**	(-13.73)	**-13.08**	(-13.04)
ξ_D	**-7.39**	(-7.44)	**7.06**	(7.11)
ξ_E	**-9.80**	(-9.86)	**4.64**	(4.69)
ξ_F	**-29.07**	(-29.24)	**-19.27**	(-19.14)
ξ_G	**-34.35**	(-34.53)	**-13.87**	(-13.72)
Avg. n		142,873		
store G to A (minutes)[a]	19		8	

Notes: Estimation results from the Inequality model. I construct moments by comparing to each store all alternative stores, for a total of 7-by-6, 42 moments. γ is the travel cost parameter. ξs denote store fixed effects. Disutility from expenditure, α, is normalized to -1. Avg. n denotes the average number of observations per moment.

[a] "store G to A" $(= (\xi_A - \xi_G)/\gamma)$ is the implied additional time a person would drive to visit store A rather than store G.

Table 1.5: Welfare Change from (a) Store Removal and (b) Replacement

Counterfactual: store removed/replaced		A	B	C	(a) D	E	F	G	(b) B-B'
Income									
	low	-37.3	-48.9	-40.3	-12.1	-7.2	-13.3	-9.5	-35.0
	medium	-43.8	-44.7	-40.3	-14.2	-8.0	-12.7	-8.3	-27.1
	high	-51.4	-38.7	-40.3	-14.6	-7.8	-11.9	-6.3	-17.6
Family Size									
	1	-36.6	-42.7	-38.3	-12.1	-7.1	-12.5	-7.9	-27.7
	2	-43.0	-43.3	-41.4	-13.5	-7.8	-13.0	-8.2	-26.9
	3	-45.5	-44.9	-40.3	-14.5	-7.8	-12.3	-8.2	-26.1
	>3	-51.7	-47.9	-40.2	-15.2	-8.0	-12.5	-8.4	-26.9
Children									
	no children	-43.8	-46.9	-42.4	-14.2	-8.2	-13.6	-8.8	-29.6
	1 child	-40.5	-37.2	-32.8	-12.0	-6.1	-9.8	-6.0	-21.1
	more than 1 child	-56.3	-38.6	-40.9	-13.2	-7.1	-10.9	-7.4	-14.1
Age									
	adult	-43.7	-41.0	-36.2	-12.7	-6.6	-11.0	-7.2	-22.8
	young	-30.2	-28.3	-26.6	-9.9	-4.9	-8.3	-5.4	-16.5
	senior	-45.9	-49.7	-45.9	-15.2	-9.0	-14.8	-9.4	-32.2
Work									
	low	-43.1	-49.4	-45.6	-14.2	-8.5	-14.6	-10.3	-33.4
	medium	-46.4	-45.3	-38.2	-15.2	-8.2	-12.4	-7.1	-26.0
	high	-42.3	-38.7	-36.5	-12.1	-6.4	-10.8	-6.7	-21.5
Education									
	low	-39.7	-44.8	-39.3	-12.5	-7.0	-12.1	-8.8	-29.5
	high	-48.6	-43.7	-41.1	-15.3	-8.5	-13.4	-7.1	-23.9
Married									
	not married	-37.4	-46.1	-38.4	-12.6	-7.2	-12.9	-8.9	-31.3
	married	-47.7	-43.4	-41.4	-14.4	-8.0	-12.5	-7.7	-24.4
Overall		-43.9	-44.4	-40.3	-13.7	-7.7	-12.7	-8.2	-26.9

Notes: Summary table of the welfare changes from (a) store removal and (b) replacement. The welfare change is measured as the percentage change in expected consumer surplus. Columns A through G report the effect of the removal of the respective store. Column B-B' the effect of the re-location of store B to a new location. Measures are reported for the overall population and for demographic groups.

Table 1.6: Expenditures

		A	B	C	D	E	F	G
Expenditures								
	Actual	3,882,571	2,895,632	2,606,166	1,150,249	545,196	405,906	533,707
	Model	3,774,935	3,135,673	2,355,431	1,053,234	504,803	734,383	460,945
Counterfactuals								
	A	0	4,138,630	3,666,426	1,389,507	651,077	978,652	594,977
	B	4,769,574	0	3,015,069	1,534,446	717,272	1,023,389	669,636
	C	4,887,552	3,742,639	0	1,248,218	606,098	913,454	567,639
	D	4,048,936	3,503,647	2,530,655	0	585,752	812,856	503,828
	E	3,884,236	3,279,722	2,435,344	1,121,004	0	772,713	480,051
	F	3,972,293	3,356,660	2,508,124	1,128,919	547,215	0	494,993
	G	3,893,646	3,301,146	2,452,067	1,097,917	527,292	770,874	0

Notes: The table reports observed (actual) expenditure and model predicted expenditure at stores A through G in the first two rows. Rows labeled A through G report counterfactual predicted expenditures after the removal of the respective store.

Table 1.7: Trips

		A	B	C	D	E	F	G	Total
Trips									
	Actual	63,649	64,672	60,537	22,643	13,052	21,157	13,705	259,415
	Model	63,648	64,671	60,536	22,643	13,052	21,157	13,704	259,411
Counterfactuals									
	A	0	80,103	86,763	28,116	16,088	26,816	16,929	254,816
	B	79,542	0	76,784	32,312	18,273	29,062	19,593	255,566
	C	88,037	79,744	0	27,693	16,081	27,127	17,361	256,043
	D	68,128	71,883	64,889	0	15,049	23,308	14,928	258,184
	E	66,063	68,389	63,030	24,479	0	22,459	14,375	258,795
	F	68,241	70,627	65,522	24,758	14,405	0	14,940	258,493
	G	66,299	68,963	63,537	23,845	13,742	22,390	0	258,776

Notes: The table reports observed (actual) visits and model predicted visits at stores A through G in the first two rows. Rows labeled A through G report counterfactual predicted visits after the removal of the respective store.

Table 1.8: Change in trips share

		A	B	C	D	E	F	G
Trips Share								
	actual	24.5	24.9	23.3	8.7	5.0	8.2	5.3
	model	24.5	24.9	23.3	8.7	5.0	8.2	5.3
Δ in Trips Share								
	A	-24.5	6.5	10.7	2.3	1.3	2.4	1.4
	B	6.6	-24.9	6.7	3.9	2.1	3.2	2.4
	C	9.8	6.2	-23.3	2.1	1.2	2.4	1.5
	D	1.9	2.9	1.8	-8.7	0.8	0.9	0.5
	E	1.0	1.5	1.0	0.7	-5.0	0.5	0.3
	F	1.9	2.4	2.0	0.8	0.5	-8.2	0.5
	G	1.1	1.7	1.2	0.5	0.3	0.5	-5.3

Notes: The table reports the change in the share of visits under the different counterfactuals. Observed (actual) shares and model predicted shares by stores A through G are reported in the first two rows. Rows A through G under counterfactuals denote the store that is being removed. Reading along a row will express how a store's share is substituted to other stores; reading along a column expresses how the removal of a competitor affects the share of a particular store.

Table 1.9: Change in Expenditures Share

		A	B	C	D	E	F	G
Expenditure Share								
	actual	32.3	24.1	21.7	9.6	4.5	3.4	4.4
	model	31.4	26.1	19.6	8.8	4.2	6.1	3.8
Δ in Expenditure Share								
	A	-31.4	10.2	12.5	3.4	1.5	2.5	1.4
	B	9.3	-26.1	6.1	4.3	1.9	2.6	1.9
	C	9.4	5.2	-19.6	1.7	0.9	1.5	0.9
	D	2.4	3.1	1.5	-8.8	0.7	0.7	0.4
	E	1.0	1.3	0.7	0.6	-4.2	0.3	0.2
	F	1.7	1.9	1.3	0.6	0.4	-6.1	0.3
	G	0.9	1.3	0.8	0.4	0.2	0.3	-3.8

Notes: The table reports the change in the revenue share of stores under different counterfactuals. Observed (actual) shares and model predicted shares by stores A through G are reported in the first two rows. Rows A through G under counterfactuals denote the store that is being removed. Reading along a row will express how a store's revenue share is substituted to other stores; reading along a column expresses how the removal of a competitor affects the share of a particular store.

Table 1.10: Store B re-placement: Trips, Expenditures and Shares

		A	B(B')	C	D	E	F	G
Trips*								
	Actual	636	647	605	226	131	212	137
	Model	636	647	605	226	131	212	137
	Counterfactual	693	272	685	301	170	270	183
Expenditures**								
	Actual	3883	2896	2606	1150	545	406	534
	Model	3775	3136	2355	1053	505	734	461
	Counterfactual	4084	1484	2654	1413	663	943	619
Share of Trips								
	Actual	24.5	24.9	23.3	8.7	5.0	8.2	5.3
	Model	24.5	24.9	23.3	8.7	5.0	8.2	5.3
	Counterfactual	26.9	10.6	26.6	11.7	6.6	10.5	7.1
	Change	2.4	-14.4	3.3	2.9	1.6	2.3	1.8
Share of Expenditure								
	Actual	32.3	24.1	21.7	9.6	4.5	3.4	4.4
	Model	31.4	26.1	19.6	8.8	4.2	6.1	3.8
	Counterfactual	34.4	12.5	22.4	11.9	5.6	8.0	5.2
	Change	3.0	-13.6	2.8	3.2	1.4	1.8	1.4

Notes: summary table of the effect on visits and expenditures at stores after the replacement of store B to a new location. Values are reported for actual, model and counterfactual number of trips, expenditure, share of trips and revenue shares at stores A through G. Change in the share of trips and revenues are also reported.
* Trips in hundreds of trips.
** Expenditures in thousands of dollars.

Table 1.11: Model Comparison: Welfare Change from store removal

		A	B	C	D	E	F	G	B-B'
1. Overall									
Model (1)		-40.8	-41.1	-37.3	-12.9	-7.2	-11.8	-7.6	-25.4
Model (3)		-43.9	-44.4	-40.3	-13.7	-7.7	-12.7	-8.2	-26.9
2. By Income									
	low	-35.7	-42.2	-34.3	-11.5	-6.3	-11.2	-7.9	-29.5
Model (1)	medium	-41.1	-42.4	-37.9	-13.4	-7.5	-12.0	-7.9	-26.6
	high	-46.5	-37.1	-39.9	-13.4	-7.6	-11.9	-6.5	-18.0
	low	-37.3	-48.9	-40.3	-12.1	-7.2	-13.3	-9.5	-35.0
Model (3)	medium	-43.8	-44.7	-40.3	-14.2	-8.0	-12.7	-8.3	-27.1
	high	-51.4	-38.7	-40.3	-14.6	-7.8	-11.9	-6.3	-17.6
3. By Family Size									
	1.0	-38.1	-41.9	-35.7	-12.2	-6.7	-11.6	-7.5	-27.5
Model (1)	2.0	-41.4	-41.4	-38.1	-12.9	-7.3	-12.1	-7.7	-26.4
	3.0	-41.1	-41.1	-38.0	-13.5	-7.3	-11.7	-7.7	-24.3
	>3	-41.9	-39.8	-36.7	-13.0	-7.2	-11.3	-7.3	-22.6
	1.0	-36.6	-42.7	-38.3	-12.1	-7.1	-12.5	-7.9	-27.7
Model (3)	2.0	-43.0	-43.3	-41.4	-13.5	-7.8	-13.0	-8.2	-26.9
	3.0	-45.5	-44.9	-40.3	-14.5	-7.8	-12.3	-8.2	-26.1
	>3	-51.7	-47.9	-40.2	-15.2	-8.0	-12.5	-8.4	-26.9

Notes: Model comparison table, comparing the welfare changes from store removal and replacement using the estimates of MNL specifications (1) and (3) from table 5. Model (1) assumes homogeneous linear travel costs, model (3) assumes heterogeneous consumer travel costs and includes a quadratic term in the disutility from driving time as well as a control for the bundle size (i.e. β_q).ps. The welfare change is measured as the percentage change in expected consumer surplus.

Table 1.12: Model Comparison: Model's Fit and Substitution Patterns

		A	B	C	D	E	F	G
Trip Shares								
	Actual	24.5	24.9	23.3	8.7	5.0	8.2	5.3
	Model (1)	24.5	24.9	23.3	8.7	5.0	8.2	5.3
	Model (3)	24.5	24.9	23.3	8.7	5.0	8.2	5.3
Expenditure Shares								
	Actual	32.3	24.1	21.7	9.6	4.5	3.4	4.4
	Model (1)	27.1	24.5	22.8	8.2	5.1	7.6	4.6
	Model (3)	31.4	26.1	19.6	8.8	4.2	6.1	3.8
Δ Expenditure Shares^a								
Model (1)								
	A	-1	0.27	0.45	0.08	0.05	0.10	0.05
	B	0.28	-1	0.27	0.15	0.09	0.12	0.08
	C	0.46	0.26	-1	0.08	0.05	0.09	0.05
	D	0.21	0.35	0.20	-1	0.10	0.10	0.05
	E	0.21	0.31	0.19	0.14	-1	0.10	0.05
	F	0.26	0.29	0.23	0.10	0.07	-1	0.05
	G	0.23	0.33	0.22	0.09	0.05	0.09	-1
Model (3)								
	A	-1	0.32	0.40	0.11	0.05	0.08	0.04
	B	0.35	-1	0.23	0.17	0.07	0.10	0.07
	C	0.48	0.26	-1	0.09	0.04	0.08	0.05
	D	0.27	0.36	0.17	-1	0.08	0.08	0.04
	E	0.25	0.31	0.18	0.14	-1	0.08	0.04
	F	0.27	0.31	0.21	0.10	0.06	-1	0.05
	G	0.24	0.34	0.20	0.09	0.05	0.08	-1

Notes: Model comparison table, comparing MNL specifications (1) and (3) from table 5 in terms the fit of trip and expenditure shares at different stores, and substitution patterns in expenditure shares under counterfactual store removal. Model (1) assumes homogeneous linear travel costs, model (3) assumes heterogeneous consumer travel costs and includes a quadratic term in the disutility from driving time as well as a control for the bundle size (i.e. β_q).ps. Rows labeling A through G refers to the removal of the respective store.

[a] The change in expenditure share, Δ Expenditure Shares^a, has been normalized to allow comparison between the two models. Normalization was done by dividing the percentage change in expenditure share at a store by the model's predicted share at the store being removed.

Chapter 2

Retail Pricing with Attractor Goods

2.1 Introduction

When going on a shopping trip, consumers are most often interested in buying a specific subset of goods and specific needs will likely guide their decision. Moreover, once in a store, consumers often end up buying substitute and complementary products, or even additional products which were not used to formulate their store choice. The assumption of rationality in retail shopping decisions appears very problematic when stores sell thousands of products and frequently vary their assortments and prices. Consumers are typically uncertain about prices at different stores and for a consumer to consider the entire distribution of bundles and prices might be a far too complex decision process. Furthermore, models with rational consumers are incapable of fully explaining important features of retail markets such as price dispersion, advertising and leader pricing. Conversely, acknowledging the limited ability of consumers to make such complicated decisions might improve our understanding of retail pricing and competition.

In this paper I attempt to characterize optimal pricing by multi-product retailers when imperfectly informed consumers buy more than one product. The distinctive feature of the model is that there are two relevant moments to all purchase decisions. First, the choice to go on a shopping trip and of the store to visit, and second, the choice of the items to purchase. These two moments are separated in time, and while consumers might *rationally* choose a store to best meet their specific needs and desires, the choice of the items to purchase is made only once in a store. Whether guided by impulse, contingent and unforeseen needs or in-store learning about a product, consumers often end up buying additional products which can generate higher profits for the stores. In this sense consumers are not fully rational, because they do not anticipate the surplus they will get from buying additional products.

To examine the implications on retail pricing of this kind of behavior, I depart from a standard rational setup and introduce the concept of *attractor goods*. Using an an approach similar to that found in Osborne and Rubinstein (1998) and Spiegler (2006) I consider boundedly rational consumers whose choice between stores is based solely and entirely on the price and valuation of a subset of goods, the attractors, while the decision about what goods to purchase is made only once a consumer is in the store and is driven by a standard willingness to pay approach. The notion of attractor good captures perhaps several concepts found in the literature. For example, Hess and Gerstner (1987) use the notion of *shopping good* (as opposed to *impulse goods*) to address the retailing strategy of *Loss Leader Pricing*: people visit a store to buy the shopping good and end up buying also impulse goods.[1]

[1]Hess and Gerstner though, and most of the literature on Loss Leader Pricing, fail to recognize why a particular good is a leader in the first place, focusing only on whether a profit maximizing firm will price that good below marginal cost. The notion of attractor good generalizes that of shopping good in that consumers are heterogeneous in what good(s) attract them, and the distinction between goods is continuous rather than dichotomous (leader/non leader).

I build a three-stage game in which stores sells K products to a continuum of consumers with heterogeneous tastes over goods. Each consumer makes her decision whether to go on a shopping trip based only on the price of an attractor which assigned by nature according to a distribution that is known to the stores. Once in a store a consumer will purchase all the items for which her valuation is higher than the price. As spatial differentiation in traditional retail markets implies that stores possess some degree of market power in the presence of consumers travel cost, I first consider the optimal pricing by a multi-product monopolist. The monopoly case can be a useful benchmark to assess the implications for retail margins and profits even in competitive markets. I then extend the framework to the case of two duopolists at the endpoints of a segment competing over consumers with linear travel costs.

The notion of attractor goods improves on the existing literature in that it shows a pattern of prices that cannot be explained only in terms of price elasticities and positioning. By giving up a fully rational consumer, using the notion of attractor goods, I am able to characterize a very important aspect of retail pricing: a good's price exerts a market size externality on the rest of the goods (by enlarging the pool of customers), and the relative strength of such externality depends on the relative attractiveness of a good; stores are typically more informed (on aggregate) about consumer tastes and needs and their pricing strategies must take into account such externality to be optimal. The optimal pricing schedule will be a decreasing function of the goods' attractiveness, and pricing below marginal cost might be optimal for some goods. The model provides a rationale for the strategy of loss leader pricing and offers an intuitive explanation to countercyclical markups.

The remainder of the paper is organized as follows. In section 2.2 I discuss some of the existing literature. In section 2.3 I describe the model. In section 2.4 I discuss the results in connection with empirical evidence found in the literature. In section

2.5 I conclude.

2.2 Background

Modeling multi-product stores and store choice by consumers is essential to the analysis of retail competition and pricing strategies. Not surprisingly, a large number of papers in the fields of Marketing and Industrial Organization have focused their attention to the analysis of pricing decisions by multi-product stores. Unfortunately though, analyzing market equilibrium and price competition with multi-product firms is intrinsically difficult,[2] and no unified treatment of market equilibrium exists in such an environment. Researchers have instead used distinct approaches to analyze different features of retail pricing.

As Varian (1980) clearly stated in an early paper, economists have long recognized that "the law of one price is no law at all." Retail stores compete along multiple dimensions, including price, quality, services and location, and enact a large number of strategies to increase their profitability. The theoretical prediction that all transactions will take place at marginal cost rested on the underlying assumptions of identical firms, perfect information and no travel costs, and economist have tried to explain the conflicting empirical evidence by relaxing these assumptions.

Large part of the literature on sales and price dispersion has focused on imperfect information on prices. Varian (1980), for example, shows that in the presence of search costs firms might use sales to discriminate between informed and uninformed consumers. Sobel (1984) however, notes that sales are sometimes so "traditional and

[2] As Anderson and De Palma (2006) point out, "to characterize profit-maximizing prices for a firm selling m products requires simultaneously solving m first-order conditions, [...] (and) to find the profit-maximizing product range for a firm necessitates finding not only the direct effect on profit from an additional product, but also the equilibrium pricing response of all other firms for all other products."

well publicized" that it is unlikely for them to occur in order to discriminate between informed and uninformed consumers. The author instead suggests that sales might occur to inter-temporally discriminate between high and low value customers. In their rationalization of loss leader pricing, the strategy of setting a price below marginal cost to increase store traffic and earn higher profits on other goods, Lal and Matutes (1994) show that in the presence of travel costs, if consumers are uninformed about the firms' prices unless advertised, firms might advertise loss leaders to compete for store traffic. As the authors suggest, the result holds even when consumers are allowed to visit more than one store and when consumers correctly anticipate higher prices on the other goods.

While numerous papers have shown how imperfect information, product differentiation, travel costs and search costs can help explain many of the strategies practiced by retailers in the presence of rational consumers, their results often have limited scope in their application, and are incapable of fully explaining important features of retail markets such as price dispersion and leader pricing. For example, Baye and Morgan (2004) note that price dispersion for identical products occurs even on internet price-search engines, where it is implausible for high search costs to exist. Similarly, low markups and loss leader pricing sometimes occur even on the regular price of a product and without extensive advertisement, making implausible the assumption that consumers are uninformed unless advertised. Furthermore, Weinstein and Ambrus (2006) show that loss leader pricing cannot occur with rational informed consumers even if they are constrained to purchase from a single store.

Unease with the assumption of consumer rationality has led many researchers to explore alternative explanations to retail strategies, often specifying simple rules for consumer behavior. Examples closely related to the current work include Feichtinger, Luhmer, and Sorger (1988), who argue that consumers' store choice is governed by aggregate information on price levels without much stress on comparison and choice, or Hess and Gerstner (1987), who explain loss leader pricing by distinguishing between *impulse goods*, goods bought on sight without price comparison across stores, and *shopping goods*, those used to determine which store to visit. Furthermore, as Ellison (2006) points out, even papers will fully rational agents could be easily revisited as an instance of bounded rationality.[3]

The notion of attractor goods in retail shopping decisions introduces a form of bounded rationality very similar to that found in Spiegler (2006) and Osborne and Rubinstein (1998). Spiegler (2006) considers markets in which products have multiple dimensions (a bank account, an insurance policy, ecc.), and firms set a price along each of these dimensions. Consumers, who are limited in their ability to understand such complicated pricing schemes use a sampling procedure to choose a product: they randomly pick one dimension and choose according to the firms' prices along that dimension. Firms on the other hand are rational, and randomize their prices to make their product really attractive along certain dimensions and more profitable in others. The sampling procedure was first developed in Osborne and Rubinstein (1998) who introduced a class of games with procedurally rational agents.

2.3 The model

Consider a three-stage game in which stores sells K products to a continuum of consumers with heterogeneous tastes over goods. At the beginning of the game, each consumer is assigned by nature with an attractor, and her decision whether to visit a store, and eventually which store to visit, will rely solely and entirely on the price and

[3]For example, Ellison (2006) notes that the high search cost consumers who choose at random in Varian (1980), or the infinitely impatient consumers in Sobel (1984), could be cast as boundedly rational consumers, and that for both papers sales could be interpreted as an attempt to price discriminate between more and less sophisticated consumers.

valuation of the attractor. Once in a store a consumer observes the prices and learns her valuations for the remaining products, and purchases all products for which her willingness to pay is higher than the price.

The peculiar assumption that consumers decision to go on a shopping trip only involves the price and valuation for the attractor is grounded on a few considerations. Stores sell a very large number of products, and consumers are typically imperfectly informed about both products and prices. While it is plausible that consumers *rationally* anticipate prices, or otherwise form expectations on the potential surplus from a shopping trip, these prices are totally uninformative when consumers are unaware of the existence of a product, or when the actual valuation for a product can be observed only once in the store. Furthermore, consumers shopping decisions are often guided by specific needs and desires. By assuming perfect information on the attractor's value and price, the attractor can be interpreted as the set of goods the consumer is both aware of and plans to buy (with respect to which he is indeed rational). The actual purchase decisions, however, are made only once in a store, and, whether guided by impulse, unforeseen needs or in-store learning about a product, consumers often end up buying more than originally planned.

In the following section I first introduce the case of a monopolist facing a mass one of heterogeneous consumers. As in the presence of consumers travel cost and spatial differentiation retail stores possess some degree of market power, a model of monopoly can be a useful benchmark to assess the implications for retail margins and profits even in competitive markets. In this case, the choice of a consumer in the second stage reduces to the choice between visiting or not visiting the store. I then extend the framework to the case of two competing duopolist at the endpoints of a segment.

2.3.1 The Monopolist Case

Assume that a Monopolist sells K products, and that the marginal cost is equal zero for all the goods. Further assume that consumers' preferences are identically and independently distributed across goods, and their willingness to pay are a draw from the uniform (0,1). The setup of the game is the following. In stage 0, nature assigns to each consumer a vector of valuations $v = (v_1, ..., v_K)$, were $v_k \sim U(0,1)$ for $k = 1, .., K$, and an attractor $j \in \{1, 2, ..., K\}$ according to a distribution $f(j) = q_j$. The distribution $f(\cdot)$ is known to the monopolist. In stage 1 the monopolist sets the prices to maximize expected profits given $f(\cdot)$. In stage 2 consumers learn their attractor j, their valuation for it v_j, and its price p_j, and decide to go on a shopping trip whenever $v_j \geq p_j$. In stage 3 consumers who visit the store make their consumption decision in order to maximize their surplus: they will purchase any product k for which $v_k \geq p_k$.

Consider a consumer who's attractor is good j; since her valuation v_j is a draw from a uniform (0,1), the probability she will visit the store is equal to

$$P(v_j \geq p_j) = \begin{cases} 0 & \text{for } p_j > 1 \\ 1 - p_j & \text{for } 0 \leq p_j \leq 1 \\ 1 & \text{otherwise} \end{cases}$$

Similarly, once in the store, the consumer will buy good $l \neq j$ with probability $1 - p_l$ when $0 \leq p_l \leq 1$, and, respectively, with probability 1 or 0 for p_l smaller than 0 or greater than 1.

In my setup, the lower bound for the willingness to pay coincides with the marginal cost and is set equal to zero; this makes it trivial to show that setting a negative price is never convenient for the monopolist, as setting a negative price has no beneficial effect on the number of consumers who visit the store[4]. Moreover, for any price $p_j \geq 1$, the demand for good j will be equal zero. Given these considerations, we can restrict the domain of prices to the compact set $[0,1]^K$.

As the total mass of consumers is equal to 1, the mass of consumers for which good j is the attractor is exactly equal to q_j. Given that the relevant domain of prices is the compact set $[0,1]^K$, I can express the profit function for the monopolist as a weighted sum of revenues, over the distribution of attractors, from people who decide to go to the store; formally,

$$\Pi(p) = \sum_{j=1}^{K} q_j(1-p_j) \cdot \left(p_j + \sum_{l \neq j} p_l(1-p_l) \right). \tag{2.1}$$

Note that Π is a continuous function in $p \in \mathbb{R}^K$, and thus attains a global maximum in $[0,1]^K$, compact subset of $\mathbb{R}^K$.

For the characterization of the maximum we can start considering the K first order conditions (FOCs) of the problem:

$$\frac{\partial \Pi(p)}{\partial p_j} \leq 0 \quad \text{and} \quad p_j \frac{\partial \Pi(p)}{\partial p_j} = 0 \tag{2.2}$$

for $j = 1, ..., K$, where

$$\frac{\partial \Pi(p)}{\partial p_j} = (1-2p_j)\left(q_j + \sum_{l \neq j} q_l(1-p_l) \right) - q_j \sum_{l \neq j} p_l(1-p_l) \tag{2.3}$$

The following propositions further restrict the relevant domain to the interval $[0,1/2]^K$.

Proposition 2.3.1. *Assume $1 > q_j > 0$ for some $j \in \{1,..,K\}$. A monopolist selling K goods will never set a price $p_j \geq 1/2$ for any p_{-j}, where p_{-j} denotes the price of all other commodities.*

[4]Note that this is not necessarily the case when there is more than one firm.

Proof. For all values of $p_{-j} \geq 0$ the derivative of the profit function w.r.t. p_j

$$\left.\frac{\partial \Pi(p)}{\partial p_j}\right|_{p_j \geq 1/2} = (1-2p_j)\Big(q_j + \sum_{l \neq j} q_l(1-p_l)\Big) - q_j \sum_{l \neq j} p_l(1-p_l) \leq 0$$

since both the first and second term are smaller or equal to zero. Moreover, if for some $l \neq j$, $p_l > 0$ then the disequality is strict.

Now, consider the case where $p = (p_j, p_{-j}) = (1/2, 0)$. Since $1 > q_j > 0$, there exists $k \neq j$ such that $q_k > 0$; but

$$\left.\frac{\partial \Pi(p)}{\partial p_k}\right|_{p} = \Big(q_k + \sum_{l \neq j} q_l(1-p_l)\Big) - q_k p_j(1-p_j) > 0.$$

$\square$

Proposition 2.3.2. *A monopolist selling K goods will set a price $p_j = 1/2$ whenever $q_j = 0$ for some $j \in \{1,..,K\}$.*

Proof. Given $q_j = 0$, for any p_{-j}, $p_j = 1/2$ is the only solution to the FOCs

$$\frac{\partial \Pi(p)}{\partial p_j} = (1-2p_j)\sum_{l \neq j} q_l(1-p_l) = 0.$$

$\square$

Propositions 2.3.1 and 2.3.2 restrict the domain of the profit maximizing price vector to the $[0, 1/2]$ interval. A price equal to $1/2$ is the price that a single product monopoly store would set; with attractor goods the multi-product monopolist has to take into account the market size externality of setting a lower price; he will set price equal to $1/2$ only on those goods for which this externality is equal to zero, those which never attract customers in the store. On the other hand, we cannot rule out that, setting a price equal to zero is optimal for some goods[5]; in fact the sign of the derivative of Π at $p_j = 0$ can be either positive or negative, depending on the weights q_i for $i = 1, .., k$:

$$\left.\frac{\partial \Pi(p)}{\partial p_j}\right|_{p_j=0} = q_j \left(1 - \sum_{l \neq j} p_l(1-p_l)\right) + \sum_{l \neq j} q_l(1-p_l).$$

For the solution to the FOCs to be a global maximum one possibility is to prove is concavity of the profit function. A sufficient condition for concavity of Π is for the Hessian of Π, H_Π, to be negative definite in $(0, 1/2)^K$. Unfortunately, as we will see, the result is not necessarily true for $K > 2$.

Consider the second own and cross partial derivatives of Π_i with respect to p_i, p_j. We have that, for all i, j, and for all p_i, $p_j \in (0, 1/2)$,

$$\frac{\partial^2 \overline{\Pi}_i(p)}{\partial p_i \partial p_i} = -2\big(q_i + \sum_{l \neq i} q_l(1-p_l)\big) < -2\big(q_i + \sum_{l \neq i} \frac{1}{2} q_l\big) < -1 \tag{2.4}$$

$$\frac{\partial^2 \overline{\Pi}_i(p)}{\partial p_i \partial p_j} = -\big(q_i(1-2p_j) + q_j(1-2p_i)\big) > -(q_i + q_j > -1. \tag{2.5}$$

By equations (2.4) and (2.5) we have that the diagonal elements of H_Π are strictly greater than one in modulus, while the off diagonal elements are strictly smaller.

Proposition 2.3.3. *For $K = 2$, the profit function $\Pi(p)$ is concave in p in the interval $(0, 1/2)^2 \subset \mathbb{R}^2$.*

Proof. For $K = 2$, as the diagonal elements are strictly greater than the off diagonal elements, $H_{11} < 0$ and $H_{11}H_{22} - H_{12}^2 > 0$; therefore H_Π is negative definite, and Π is concave. □

[5] In the present setup a corner solution at which price for good j is set equal to zero implies that pricing below marginal cost can be optimal when marginal cost is greater than zero.

Unfortunately, matters get more complicated when $K \geq 3$. Concavity of H_Π would require that, for all $x = (x_1, x_2, .., x_K)$

$$x'Hx \leq 0$$

$$-2\sum_i x_i^2\big(q_i + \sum_{l \neq i} q_l(1-p_l)\big) - 2\sum_{i,j>i} x_i x_j\big(q_i(1-2p_j) + q_j(1-2p_i)\big) \leq 0. \qquad (2.6)$$

While an explicit closed form solution the the FOCs cannot be found, and we are not guaranteed that such a solution would imply a global maximum, I can characterize an optimal pricing scheme in terms of the probabilities q_js. By proposition 2.3.4, the optimal pricing schedule is a decreasing function of goods attractiveness. This is due to the market size externality of setting a low price: the more attractive a good the larger this externality and thus the lower the price. A useful corollary to proposition 2.3.4 is that goods with the same probability of being an attractor must be priced symmetrically.

Proposition 2.3.4. *Let $q_1 \leq q_2 \leq \cdots \leq q_K$. Suppose $p = (p_1, p_2, .., p_K)$ maximizes Πp; then it must be that $p_1 \geq p_2 \geq .. \geq p_K \geq 0$. Moreover, if $p_j > 0$ and $q_j \neq q_{j+1}$ it must be that $p_j > p_{j+1}$ (Price Decreasing in q), and whenever $q_i = q_j$ for some i, j, it must be that $p_i = p_j$, and if (Symmetry Within Groups).*

Proof. See Appendix A.3. □

In now analyze two particular cases and provide some numerical solutions. I first illustrate the "symmetric" case in which all goods have the same probability of being an attractor; using symmetry, I explicitly solve for the optimal price as a function of the number of products K. I then consider the more general case of two sets of goods, and numerically solve some examples.

The Symmetric Case

Assume that for all j, $q_j = 1/K$. The profit function for the monopolist simplifies to

$$\Pi(p) = \frac{1}{K}\sum_{j=1}^{K}(1-p_j)\left(p_j + \sum_{l\neq j} p_l(1-p_l)\right).$$

Consider the first order conditions:

$$\frac{\partial \Pi(p)}{\partial p_j} = \frac{1}{K}\cdot\left((1-2p_j)\big(1+\sum_{l\neq j}(1-p_l)\big) - \sum_{l\neq j} p_l(1-p_l)\right) = 0 \text{ for } j = 1, ..., K.$$

Assuming symmetry, we can solve for the optimal price p^* as a function of K:

$$p^* = \frac{2K-1-\sqrt{(2K-1)^2-3K(K-1)}}{3(K-1)}$$

By proposition 2.3.4 (*Symmetry Within Groups*), p^* is a global maximizer since it is the unique symmetric solution to the set of first order conditions.

The relevant result here is that as K gets larger the monopolist will charge lower prices[6] (see Figure 2.1). While in a standard setup the monopolist would charge monopoly price for all products, our behavioral setup introduces a market size externality on prices. An increase (decrease) in the price of commodity j, besides the direct effect on the market for product j, will decrease (increase) the market size for all the other products.

A surprising result is that, a K gets large, the $\lim_{K\to\infty} p^* = 1/3$. The intuition behind this is that, at optimum, the marginal benefit (cost) in the market for j of increasing (lowering) p_j, must equal to the marginal cost (benefit) caused in the market for the other goods. More formally, if we move p_j away from p^*, keeping all other prices fixed at p^*, as K gets large, the marginal effect in the market for j

[6] In fact $\partial p^*/\partial K < 0$

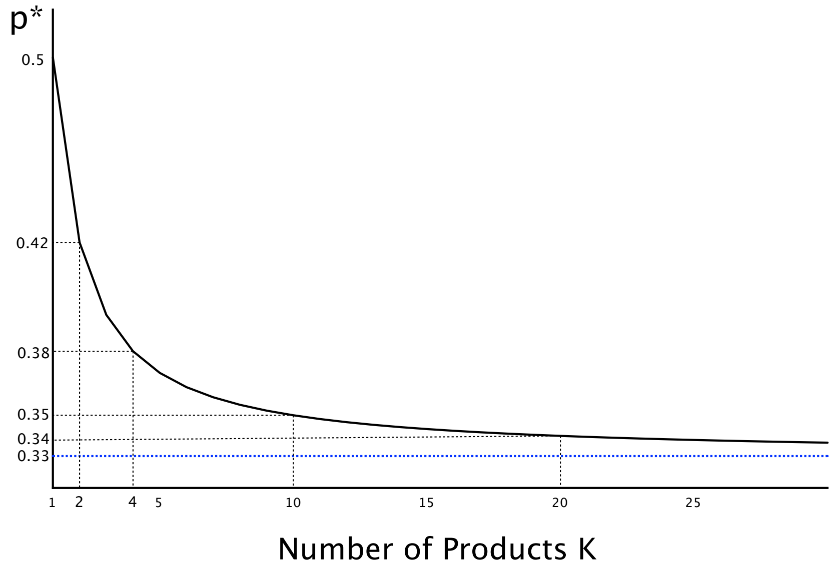

Figure 2.1: The Symmetric Case: Profit Maximizing Price as a Function of the Number of Goods

will be equal to $2/3(1 - 2p_j)$, while that on the market for other goods is equal to $p^*(1 - p^*) = 2/9$; for the two effects to be equal we need $p_j = 1/3$. Alternatively you can consider the limit as K gets large of the profit function restricted to symmetric prices; we have that

$$\lim_{K \to \infty} \Pi_K = (1 - p)^2 p,$$

which indeed attains a maximum at $p = 1/3$.

Two Sets of Goods

Consider now the more general case where $q_i \neq q_j$ for some i, j. As no general closed form solution exists even in the two goods case, I explore a specific setup to offer some qualitative insights on the pattern of optimal prices as the number of products and the relative probabilities change.

Assume there are two sets of goods, K_1 and K_2, such that for all $i \in \{1, .., k_1\}$,

$q_i = q_1$, and for all $j \in \{k_1+1, .., K\}$, $q_j = q_2$; also, $k_2 = K - k_1$, and $0 < k_1 < K$. By proposition 2.3.4 (*symmetry within groups*) we know that if $\overline{p}$ is maximizes Π, then all goods in one group must have the same price. I consider two simple cases.

Only Two Goods Suppose $K_1 = K_2 = 1$. The FOCs in this case are equal to:

$$(1-2p_j)\big(q_j + q_i(1-p_i)\big) - q_j p_i(1-p_i) = 0 \text{ for } i, j = 1, 2,\ i \neq j.$$

By concavity of $\Pi(p_1, p_2)$ (proposition 2.3.3), the unique solution to the FOCs $p^* = (p_1^*, p_2^*)$ is a global maximum. I numerically solve the maximization problem for different values of q and plot the results in Figure The black (blu) line represents p_1^* (p_2^*) as a function of q_1 (q_2); note that $p_1^* = 1/2$ when $q_1 = 0$, as we already know by proposition 2.3.1, and decreases smoothly to 0.375 when $q = 1$. The two lines intersect when $q_1 = q_2 = 1/2$ at $p = 0.422$. 2.2.

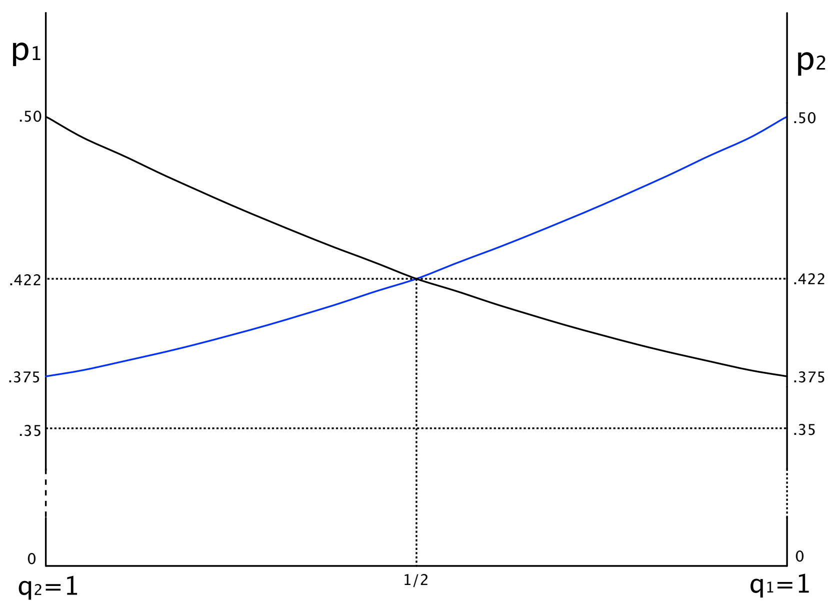

Figure 2.2: Profit Maximizing Prices with Two Goods as a function of q_1 and q_2

Few Attractors, Many Goods Consider now the special case in which $K_1 \cdot q_1 = 1$, that is, only one subset of goods K_1 can be attractors with probability $q_1 = 1/K_1$. By proposition 3.2 we know that for all other goods the price will be set to 1/2. I consider the optimal price p_1 as a function of both K_1 and K_2.

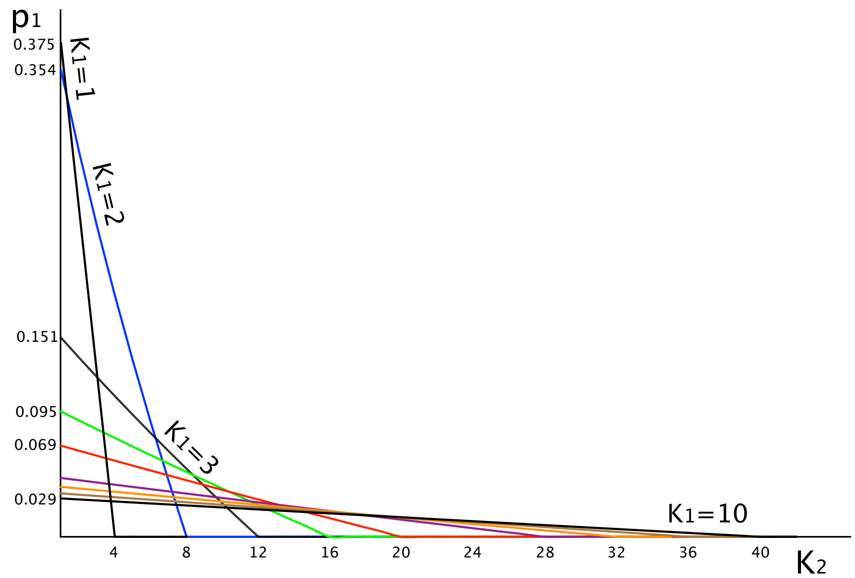

Figure 2.3: Few Attractors, Many Goods: the optimal price p_1 as a function of K_2

Let's start considering the case were $K_1 = 1$; in this case $q_1 = 1$, and given that all other prices will be set equal to 1/2, we can express the profit for the monopolist as a function of p_1 only:

$$\hat{\Pi}(p_1) = (1 - p_1)\left(p_1 + \frac{K_2}{4}\right).$$

It is easy to verify that

$$p_1^* = \begin{cases} \frac{4-K_2}{8} & \text{for } K_2 \leq 4 \\ 0 & \text{for } K_2 > 4 \end{cases}.$$

In words price decreases from 1/2 to 0 in steps of 1/8 by adding one product at the time.

Now let K_1 vary; again, we know that all goods in the second set will have a price equal to 1/2. So consider the monopolist's profit:

$$\hat{\Pi}(p_1) = (1 - p_1)\left(p_1 + (K_1 - 1)p_1(1 - p_1) + \frac{K_2}{4}\right).$$

It is easy to verify that

$$p_1^* = \begin{cases} \frac{(2K_1-1)-\sqrt{(2K_1-1)^2-3(k_1-1)(k_1-\frac{K_2}{4})}}{3(K_1-1)} & \text{for } K_2 \leq 4K_1 \\ 0 & \text{for } K_2 > 4K_1 \end{cases}.$$

In Figure 2.3 I plot the optimal price p_1 as a function of K_2 for $K_1 = 1, 2, ..., 10$.

2.3.2 The Duopoly Case

Consider now competition between two stores selling K products, A and B, located at the endpoints of a segment of length 1. As for the monopolist case, I assume that the marginal cost of supplying the products is equal zero. I further assume that a mass one of consumers is distributed uniformly along the segment, and consumers' willingness to pay ν_{ij} for good $j \in \{1, ..., K\}$ is i.i.d. across both consumers and goods. Consumers face linear travel costs of visiting a store; if we denote by x_i the distance of consumer i from store A, where $x_i \sim U(0, 1)$, the cost of visiting store A is thus equal to x_i.

The setup of the game is very similar to the monopolist setup. In stage 0, nature assigns to each consumer i a vector of valuations $v_i = (v_{i1}, ..., v_{iK})$, were $v_{ik} \sim U(0, 1)$ for $k = 1, .., K$, and an attractor $j \in \{1, 2, ..., K\}$ according to a distribution $f(j) = q_j$.

The distribution $f(\cdot)$ is known to the stores. In stage 1 the two stores simultaneously set prices to maximize expected profits given $f(\cdot)$. In stage 2 consumers learn their attractor j, their valuation for it v_{ij}, and its price p_j and decide to visit a store A whenever $v_{ij} \geq p_{Aj} + x_i$ and $p_{Aj} + x_i \geq p_{Bj} + 1 - x_i$. Alternatively they decide to visit store B if $v_{ij} \geq p_{Bj} + (1 - x_i)$ and $p_{Bj} + 1 - x_i \geq p_{Aj} + x_i$, or decide not to visit a store if $v_{ij} \leq \min\{p_{Aj} + x_i \ , \ p_{Bj} + (1 - x_i)\}$. In stage 3 consumers, once in a store, learn their valuations for the other products and purchase any product k for which $v_{ik} \geq p_{Sk}$.

Let p_{Aj}, p_{Bj} for $j = 1, ..., K$ denote the prices charged by the two firms, and let $k_1 \leq K$ denote the (sub-)set of attractors (i.e. $\{j : q_j > 0\}$). The profit function of firm A is a weighted sum of revenues, over the distribution of attractors j, from people who decide to visit the store. Note that both firms will never charge a price greater than one for any of the products, as in this case demand would be zero. Conversely, it is possible for prices to be negative, as pricing below marginal cost can be supported in equilibrium when the higher revenues due to the increased traffic at a store outweigh the loss deriving from such a low price. The profit of store A is thus equal to

$$\Pi_A(p_A, p_B) = \sum_{j=1}^{k_1} q_j \left(p_{Aj} + \sum_{l \neq j} p_l (1 - \max\{0, p_l\}) \right) \overline{D}_{Aj}(p_{Aj}, p_{Bj}), \tag{2.7}$$

where $\overline{D}_{Aj}(p_{Aj}, p_{Bj})$ denotes the fraction of consumers who's attractor is j that decide to fit store A, and $(1 - \max\{0, p_l\})$ denotes the probability a consumer will buy good l once in a store.

Denote $\overline{d}_j$ the distance for which a consumer who's attractor good is j would be indifferent between visiting the two stores, that is

$$\overline{d}_j = \{x : p_{Aj} + x = p_{Bj} + (1 - x)\} = \max\left\{\frac{1 - p_{Aj} + p_{Bj}}{2}\ ,0\right\}.$$

The probability that a consumer will visit store A will thus be positive if and only if $x_i \leq \min\{\overline{d}_j\ ,\ 1 - p_{Aj}\}$. The fraction of consumers $\overline{D}_{Aj}$ that will visit store A is therefore equal to:

$$\overline{D}_{Aj}(p_{Aj}, p_{Bj}) = \begin{cases} \int_0^{\overline{d}_j} Pr(\nu_{ij} \geq p_{Aj} + x)dx & \text{if } p_{Aj} \leq 1 - p_{Bj} \text{ and } p_{Aj} \leq 1 + p_{Bj} \\ \int_0^{1-p_{Aj}} Pr(\nu_{ij} \geq p_{Aj} + x)dx & \text{if } 1 + p_{Bj} \geq p_{Aj} \geq 1 - p_{Bj} \\ 0 \quad \text{if } p_{Aj} \geq 1 + p_{Bj} \end{cases}$$

Propositions 2.3.5 and 2.3.6 further restrict prices to be smaller or equal to 1/2.

Proposition 2.3.5. *A duopolist will set a price $p_{Sj} = 1/2$ whenever $q_j = 0$ for some $j \in 1, ..., K$*

Proof. W.l.g. consider, p_{Aj}. Given $q_j = 0$, for any $(p_A, p_B)_{-j}$, $p_{Aj} = 1/2$ is the only solution to the FOCs $\partial\Pi_A/\partial p_{Aj} = 0$, where

$$\frac{\partial\Pi_A}{\partial p_{Aj}} = \begin{cases} 1 & \text{for } p_{Aj} \leq 0 \\ (1 - 2p_{Aj}) \sum_{l \neq j} q_l \overline{D_l}(p_{Al}, p_{Bl}) & \text{otherwise} \end{cases}$$

□

Proposition 2.3.6. *Assume $1 > q_j > 0$ for some j. A duopolist S will never charge a price $p_{Sj} \geq 1/2$ for any prices p_{Sk}, $k \neq j$ and any prices offered by the competitor $p_{S'}$.*

Proof. See Appendix A.4. □

The case of One Attractor

In the simplest version of the model the duopolists compete over consumers when only one good is an attractor and k additional products are sold. Under these assumptions the model is very similar to a Hoteling model with the complication that consumers are heterogeneous in their willingness to pay.

Let p_A, p_B denote the price charged by the two stores for the attractor. By proposition 2.3.5 we know that the k non attractors will be charged a price equal to 1/2, and by propositions 2.3.6 we know that $p_S \leq 1/2$ for $S = A, B$. Furthermore, the following proposition holds:

Proposition 2.3.7. *A set of prices* (p_A, p_B) *such that, either* $p_A - p_B > 1$ *or* $p_B - p_A > 1$ *will never support an equilibrium.*

Proof. Suppose $p_A - p_B > 1$ ($p_B - p_A > 1$); then nobody is visiting store A (B) and store B (A) can strictly increase profits by setting a higher price. □

Under propositions 2.3.5-2.3.7 we can write the profit of firm A as be equal to:

$$\begin{aligned} \Pi_A(p_A) &= (p_A + k/4) \int_0^{\hat{d}} Pr(\nu_i \geq p_A + x) dx \\ &= \frac{1}{8}(p_A + k/4)(1 - p_A + p_B)(3 - 3p_A - p_B) \end{aligned}$$

Now consider the FOCs of the problem for firm A:

$$\frac{\partial \Pi_A(p_A)}{\partial p_A} = 9p_A^2 - (12 + 4p_B)p_A + (-p_B^2 + 2p_B + 3) + K/4(6p_A - 6 - 2p_B) = 0 \quad (2.8)$$

The following proposition shows that there exist a unique symmetric equilibrium $p_A = p_B = p^*$, that is decreasing in k.

Proposition 2.3.8. *There exist a unique symmetric Nash equilibrium in which both stores charge the same prices. Furthermore, the symmetric equilibrium price is decreasing in K and will be negative for $k > 2$.*

Proof. $\Pi_A(p_A, p_B)$ is a third order polynomial in p_A. If we have an interior solution the maximum of $\Pi_A(\cdot)$ corresponds to the smallest solution of the FOC, so that the best response of store A is continuous in p_B. Once we impose symmetry, a symmetric equilibrium price is found at the smallest solution to the FOC

$$4p^2 - (10 + K)p + (3 - 3K/2) = 0,$$

that is

$$p^* = \frac{(10 - K) - \sqrt{(10 - K)^2 - 16(3 - 3/2K)}}{8} \tag{2.9}$$

Furthermore, the largest solution is always found at $p > 1$, so that the symmetric equilibrium is unique. If we totally differentiate the expression above, by the implicit function theorem we have that

$$\frac{dp}{dK} = -\frac{3/2 - p}{10 - 8p - k} < 0,$$

as the numerator is positive by proposition 2.3.6, and, by (2.9) a) for $K = 2$, $p^* = 0$ and b) for $K \neq 2$, $p^* < (10 - K)/8$. $\square$

Figure 2.4 plots the symmetric equilibrium price as a function of k.

The case of Two Attractors

I now consider the more general case in which the two stores compete over consumers with two attractor goods and sell k additional goods. By propositions 2.3.5 Profits of firm A are given by:

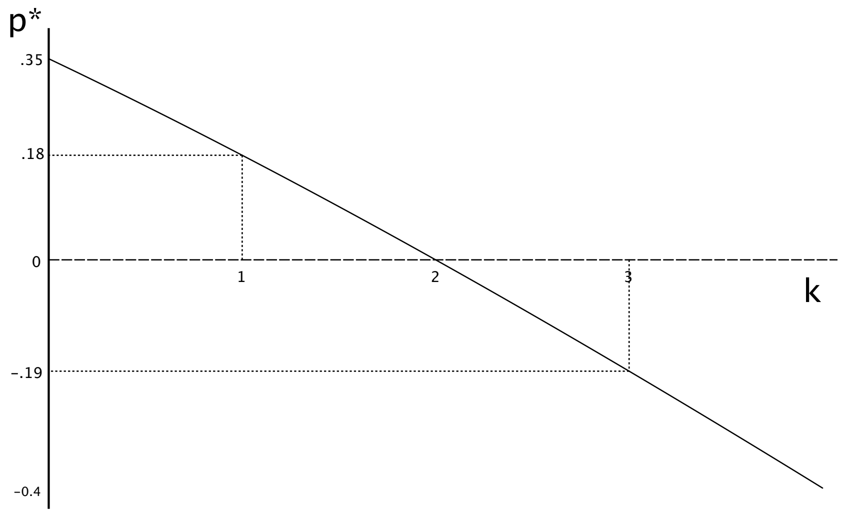

Figure 2.4: The Duopoly Case with linear Travel Costs, 1 attractor and K non attractors: symmetric equilibrium prices p^* as a function of K

$$\Pi_A(p_A, p_B) = \sum_{j=1,2} q_j(p_{Aj} + p_{A-j}(1 - \max\{0, p_{Aj}\})) + k/4) \int_0^{\bar{d}_j} Pr(\nu_i \geq p_{Aj} + x)dx$$

where $j, -j = \{1, 2\}$ and $\bar{d}_j = (1 - p_{Aj} + p_{Bj})/2$. Suppose firms set prices such that they attract a strictly positive mass of consumers for each of the two goods, then the profit of firm A is equal to

$$\begin{aligned} \Pi_A = \quad & \frac{q_1}{8}\left[(p_{A1} + p_{A2}(1 - \max\{0, p_{A2}\}) + k/4\right]\left[1 - p_{A1} + p_{B1}\right]\left[3 - 3p_{A1} - p_{B1}\right] + \\ & \frac{q_2}{8}\left[(p_{A2} + p_{A1}(1 - \max\{0, p_{A1}\}) + k/4\right]\left[1 - p_{A2} + p_{B2}\right]\left[3 - 3p_{A2} - p_{B2}\right] \end{aligned}$$

For ease of notation, consider the case in which prices are non-negative (it is straightforward to consider the case when prices are negative). Consider the set of first order conditions for the two firms. A set of equilibrium prices at an interior solution must satisfy the following set of four quadratic equations in four unknowns:

$$\begin{aligned}\frac{\partial \Pi_S}{\partial p_j} = \quad & q_j \Big[9p_{Sj}^2 - (12 + 4p_{S'j})p_{Sj} + (-p_{S'j}^2 + 2p_{S'j} + 3) + \\ & \big(6p_{Sj} - (6 + 2p_{S'j})\big)\big(p_{Sl}(1 - p_{Sl}) + K/4\big)\Big] + \\ & q_l(1 - 2p_{Sj})\Big[3p_{Sl}^2 - (6 + 2p_{S'l})p_{Sl} + (-p_{S'l}^2 + 2p_{S'l} + 3)\Big] = 0\end{aligned}$$

for $S, S' = A, B$ and $j, l = 1, 2$.

I focus here on symmetric equilibria. As the profit function is a truncated third order polynomial in p_{Sj} for $S = A, B$ and $j = 1, 2$, the maximum at an interior solution is found at the smallest solution of the quadratic first order conditions; for a given value of (q_1, q_2) the symmetric equilibrium is therefore the unique solution to the following set of two equations in the two unknowns $p_1,\ p_2$:

$$q_j \Big[\big(4p_j^2 - 10p_j + 3\big) + (4p_j - 6)\big(p_k(1 - p_k) + k/4\big) \Big] + q_k(1 - 2p_j)\big[-4p_k + 3\big] = 0$$

for $j = 1, 2$. In figure 2.5 I plot the symmetric equilibrium prices when $k = 0, 1, 2$. The plot was constructed by numerically solving the set of equations. As it appears clear from the figure price is a decreasing function a good's attractiveness, and negative pricing can be supported in equilibrium. Conversely, given a distribution q_j, for $j = 1, 2$, the symmetric equilibrium prices are not always decreasing in K. The latter however holds for the good with the highest probability of being an attractor.

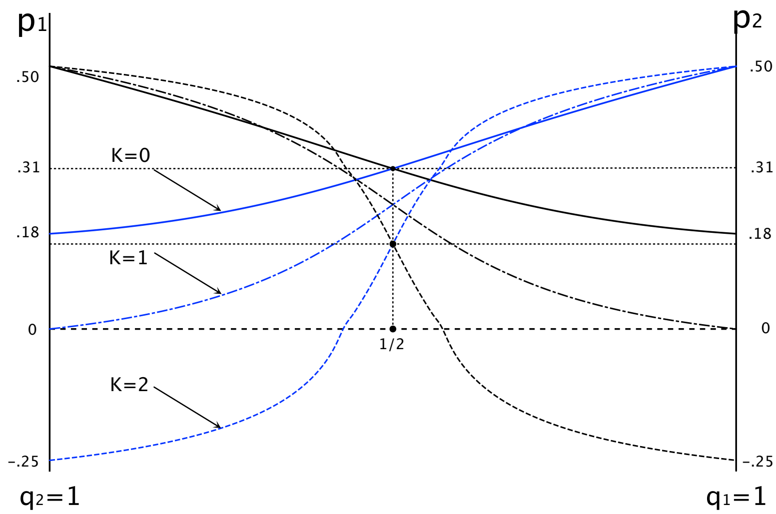

Figure 2.5: The Duopoly Case with linear Travel Costs, 2 Attractors and K Goods: symmetric equilibrium prices p_1 and p_2 as a function of q_1 and q_2

2.4 Attractor Goods: Model Predictions and Empirical Evidence

Retail stores typically sell a large number of products and frequently vary their assortments and prices. Consumers are typically uncertain about prices at different stores and for a consumer to consider the entire distribution of bundles and prices might be a far too complex decision process. While it is plausible that consumers *rationally* anticipate prices when planning a shopping trip, or otherwise form expectations on the potential surplus from a shopping trip, consumers make their purchase decision only once in a store and often end up buying substitute and complementary products, or even additional products which were not used to formulate their store choice. Attractor goods can so be interpreted as the set of goods the consumer is both aware of and rationally plans to buy, and by introducing this form form of bounded rationality I can characterize the effect on prices of consumers tendency to buy more than originally planned.

According to the model, the optimal pricing involves pricing goods lower and lower the higher the probability of being an attractor. This result naturally follows from the market size externality of setting a low price: the more attractive a good the larger this externality. In particular, goods which are never attractors should be priced as in a standard single-product monopoly case by equating a product's marginal revenue to marginal cost, while attractor goods should be priced taking into account also the marginal revenue on the other goods, and might be priced below marginal cost when the latter effect is large. The examples below suggest that the model's predictions are consistent with the empirical evidence.

Barsky, Bergen, Dutta, and Levy (2001) focus on marginal cost and the estimation of markups; using the price gap between branded and private label products they estimate markups on selected product categories in the groceries industry using data from a major Chicago supermarket chain. In particular the build three measures of markup: the full markup, the retailer markup on the national brand, and the retailer markup on the private label. To the extent that national brands attract more than private labels, and given that the national brand manufacturer does not internalize the externality for the retailer, the attractor good hypothesis would imply lower retail margins on national brands than on private labels. Conversely, we would expect higher markups on products that do not fall in the main categories sold at the store. Consistently with the model, Barsky et al. find that in all of the 19 categories analyzed markups are significantly higher on private labels than on national brands. Furthermore, the highest markups are found in categories of products for which consumers only seldomly anticipate a purchase.[7]

[7]The five categories with the highest markups are toothbrush, soft drinks, crackers, grooming

A second inference that can originate from the model has to do with the cyclicality of demand for certain goods. Using the terminology of the paper, some goods attractiveness peaks in different times of the year: a perfect example would be turkey for thanksgiving. According to the model we expect very low prices and margins during demand peaks in order to attract a larger pool of customers. Empirical evidence points in this prediction as Chevalier, Kashyap, and Rossi (2003) showed in a recent paper. The authors compare three theories of imperfect competition which can produce countercyclical prices; by examining retail and wholesale prices they show that prices and margins for specific goods fall during demand peaks, even if these periods do not coincide with aggregate demand peaks for the retailer; they suggest their findings are consistent with *loss leader models* rather than cyclical demand elasticities or cyclical firm conduct. The attractor good hypothesis generalizes model of loss leader pricing in that low markups and loss leaders can arise even on the regular price of a product and without extensive advertisement when a good's probability of being an attractor is high.

While the examples above are limited to the supermarket industry, the notion of attractor goods should not be confined to, nor is most important in, the analysis of supermarkets. It can provide a powerful explanation to a large amount of price dispersion where the assumption of rationality fails, and to the extent that consumers tend to purchase more than what they plan before visiting a store, the qualitative predictions of the model should be observed in the pricing strategies of retailers.

products and analgesics.

2.5 Conclusions

Retailers' pricing decisions have to take into account not only the direct effect of prices on a product's demand but also the effect on the demand for the other products sold in the store. A low price on some goods will attract more traffic into the store, which turns into higher profits if the other goods are sold at higher margins. Pricing strategies by retailers typically involve varying margins and markups across products and over time and several models give different explanations to such differences. Loss leader models, for example, show that it is sometimes convenient to advertise low prices on some goods to attract uninformed consumers who, once in the store, and because of transportation costs, will decide to shop also other goods on which retailers set higher margins.

While numerous papers have explained many retail strategies within a rational framework, their results often have limited scope in their application, and unease with the assumption of consumer rationality has led many researchers to explore alternative explanations. By giving up a fully rational consumer, using the notion of attractor goods, I am able to characterize a very important aspect of retail pricing: a goods price exerts a market size externality on the rest of the goods, and the relative strength of such externality depends on the relative attractiveness of a good. If stores take into account such externality in their pricing decisions markups will be a decreasing function of the goods' attractiveness. The model provides a generalization to the strategy of loss leader pricing as low markups can arise even on the regular price of a product and without extensive advertisement. Furthermore, the attractor goods hypothesis offers an intuitive explanation to countercyclical markups that is consistent with the empirical evidence.

Appendix A

Appendices

A.1 Calculation of GMS tests and confidence intervals for Moment Inequalities

Andrews and Soares (2010) introduced a generalized moment selection (GMS) procedure which provides confidence sets (CS) that have correct asymptotic size in a uniform sense and are not asymptotically conservative. For the inequalities models in the present paper I adapt one of selection criteria suggested by the authors to the case in which moments have a different number of observations, and introduce a way to characterize a $1 - \alpha$ confidence interval by means of a grid search around the vertices of the estimated set. I here describe how I proceed with the actual calculations; more details on the GMS procedure can be found in Andrews and Soares (2010).

Consider the moment inequality model. The value of the true parameters $\theta_0 \in R^d$ satisfies the moment conditions:

$$E_{F_0} m_j(y_i, \theta_0) \geq 0 \quad \text{for } j = 1, ..., p \tag{A.1}$$

where $\{m_j(\cdot,\theta): \quad j=1,...,p\}$ are p known real-valued functions and $\{y_i : i \geq 1\}$ are i.i.d random vectors with distribution F_0. Let $T_n(\theta)$ denote a test statistic for testing $H_0 : \theta = \theta_0$; Andrews and Soares (2010) propose we construct CS by inversion of the test, that is finding

$$CS_n = \{\theta \in \Theta : T_n(\theta) \leq c_{1-\alpha}(\theta)\} \tag{A.2}$$

where $c_{1-\alpha}(\theta)$ is a critical value that depends on the value of θ.

The observed sample is a subset of $\{y_i : i \leq n\}$, where for each moments we have observations $i_j = 1,...,n_j$, $n_j \leq n$. Consider the empirical moment functions:

$$\overline{m}_j(y,\theta) = \frac{1}{n_j}\sum_{i}^{N_j} m_j(y_i,\theta) \tag{A.3}$$

Denote $\overline{m}_n(\theta) = (\overline{m}_1(\theta),...,\overline{m}_p(\theta))'$ and let Σ denote the asymptotic covariance of $n^{1/2}\overline{m}_n(\theta)$. The test statistic $T_n(\theta)$ is defined to be

$$T_n(\theta) = S(n^{1/2}\overline{m}_n(\theta), \hat{\Sigma}_n(\theta))$$

where S is one of the real functions suggested in Andrews and Soares (2010), and $\hat{\Sigma}_n(\theta)$ is a consistent estimator of Σ.

As in my setup moments have a different number of observations, I consistently estimate Σ, the asymptotic covariance of $n^{1/2}\overline{m}_n(\theta)$, by bootstrapping 1000 times the means $n_j^{1/2}\overline{m}_j(\theta)$ for $j = 1,...,p$. I then compute T_n using $\mathbf{n}^{1/2}\overline{m}_n = (n_1^{1/2}\overline{m}_1,...,n_p^{1/2}\overline{m}_p)$

As a choice of S I consider $S = S_1$ defined by

$$S_1 = \sum_{j=1}^{p}(m_j/\sigma_j)_-^2,$$

where σ_j^2 is the j-th diagonal element of Σ.

Now consider the asymptotic null distribution of $T_n(\theta)$. By central limit theory this is given by

$$S(\Omega_0^{1/2}Z^* + h, \Omega_0)$$

where $Z^* \sim N_p(0, I)$, and the p-vector $h \in R^p_{+\infty}$ is the limit of $n^{1/2}E_{F_n}(m(y_i, \theta_0))$. As h cannot be consistently estimated, Andrews and Soares (2010) propose using a data-dependent version of $S(\Omega_0^{1/2}Z^* + h, \Omega_0)$ that replaces h with with a vector $\phi(\cdot)$ that depends on the "slackness" of the moment inequalities. As a measure of the degree of slackness for moment j, the authors suggest using

$$\xi_n^j(\theta) = \kappa_{n_j}^{-1} n_j^{1/2} \hat{\sigma}_j^{-1}(\theta)\overline{m}_j(\theta),$$

for a divergent sequence of constants $\{\kappa_n : n \geq 1\}$. The desired replacement for h then will be zero when $\xi_n^j(\theta)$ is zero or close to zero, it will $+\infty$ when $\xi_{nj}(\theta)$ is large. My choice for the constant is $\kappa_n = (\ln n)^{1/2}$, and as a replacement for h is use

$$\phi_j(\xi^j, \Omega) = \begin{cases} 0, & \text{if } \xi^j \leq 1 \\ +\infty, & \text{if } \xi^j > 1 \end{cases}$$

Given a choice of $\phi(\xi, \Omega)$, the GMS critical value $c_n(\theta_0, 1-\alpha)$ is the $1-\alpha$ quantile of

$$L_n(\theta_0, Z^*) = S(\hat{\Omega}_n^{1/2}(\theta_0)Z^* + \phi_n(\xi_n(\theta_0)), \hat{\Omega}_n(\theta_0)) \tag{A.4}$$

where $Z^* \sim N_p(0, I)$. That is

$$c_n(\theta_0, 1-\alpha) = \inf\{x \in R : P(L_n(\theta_0, Z^*) < x) \geq 1-\alpha\} \tag{A.5}$$

Construction of CS by inverting the test can be very demanding computationally, if feasible at all; the search of points in the CS over the parameter space Θ involves estimating Σ and constructing c_n at any parameter value. Additionally, characterization of the CS can be very hard when the estimates are a set. For these reasons I take a simple approach to characterize the CS around the vertices of the set estimate by means of a grid search.

When the model is linear, the identified set is a $d-$dimensional polytope, and minima and maxima for each of the parameters are found at (some of) the vertices of the set. Minima (and maxima) are readily found by solving one dimension at the time the linear programming problem

$$\min \theta_l \ (\max \theta_l) \text{ subject to } \overline{m}_n(y, \theta) \geq 0$$

Let $\theta = (\theta_l, \theta_{-l})$, and denote $\hat{\theta}$ one of the bounds solving the linear programming problem. To characterize the CS I proceed as follows:

1. Estimate $\hat{\Sigma}_{\hat{\theta}} = \hat{\Sigma}(\hat{\theta})$
2. For dimensions $l = 1, ..., d$, for an increasing sequence δ_n, let

$$\hat{\theta}_n = \begin{cases} (\hat{\theta}_l - \delta_n, \hat{\theta}_{-l}) \text{ if } \hat{\theta} \text{ is a lower bound} \\ (\hat{\theta}_l + \delta_n, \hat{\theta}_{-l}) \text{ if } \hat{\theta} \text{ is an upper bound} \end{cases}$$

3. Start from $n = 1$; construct $T_n(\hat{\theta}_n)$ and $c_n(\hat{\theta}_n)$. Continue while $T_n < c_n$.
4. Repeat 1-3 at all vertices of the identified set.

Table 13 provides an example for the set estimates of inequality model (2). The upper-left block of the table reports the highest and lowest values for each of the parameters as well as a 95% confidence bounds moving away from the vertex. The other blocks report the value of the other parameters at the specific vertices, as well as a 95% confidence bounds moving away from the vertex along each of the parameters' dimension.

Table A.1: Characterizing CS for set estimates using GMS

Vertex of serch	All lb.	All ub.	γ_0 lb.	γ_0 ub.	ξ_B lb.	ξ_B ub.	ξ_C lb.	ξ_C ub.
γ_0	**-1.85**	**−1.75**	**-1.85**	**−1.75**	-1.85	-1.85	-1.85	-1.74
	-1.87	-1.74	-1.87	-1.74	-1.87	-1.84	-1.87	-1.74
ξ_A	0	0	0	0	0	0	0	0
ξ_B	**-25.56**	**−8.60**	-16.57	-16.32	**-25.56**	**−8.60**	-18.86	-16.51
	-25.64	-8.52	-19.91	-12.63	-25.64	-8.52	-21.81	-12.75
ξ_C	**-13.70**	**−13.08**	-13.70	-13.08	-13.68	-13.64	**-13.70**	**−13.08**
	-13.73	-13.04	-13.74	-13.04	-13.75	-13.59	-13.73	-13.04
ξ_D	**-7.39**	**7.06**	1.29	0.66	-5.56	6.38	-0.92	0.51
	-7.44	7.11	1.24	0.72	-5.62	6.42	-0.96	0.56
ξ_E	**-9.80**	**4.64**	-1.13	-1.49	-7.98	3.96	-3.34	-1.67
	-9.86	4.69	-1.19	-1.45	-8.03	4.01	-3.38	-1.59
ξ_F	**-29.07**	**−19.27**	-23.35	-22.82	-29.07	-19.70	-25.26	-22.95
	-29.24	-19.14	-25.59	-20.07	-29.21	-19.59	-27.76	-20.20
ξ_G	**-34.35**	**−13.87**	-22.22	-22.10	-30.22	-17.23	-25.04	-22.36
	-34.53	-13.72	-28.00	-17.89	-34.53	-13.76	-30.30	-18.08

Vertax of search	ξ_D lb.	ξ_D ub.	ξ_E lb.	ξ_E ub.	ξ_F lb.	ξ_F ub.	ξ_G lb.	ξ_G ub.
γ_0	-1.85	-1.85	-1.85	-1.85	-1.85	-1.74	-1.85	-1.85
	-1.87	-1.84	-1.87	-1.84	-1.87	-1.74	-1.87	-1.84
ξ_A	0	0	0	0	0	0	0	0
ξ_B	-23.56	-12.66	-23.61	-12.56	-22.33	-13.09	-23.92	-12.32
	-25.63	-8.50	-25.63	-8.51	-25.65	-9.15	-24.68	-9.49
ξ_C	-13.70	-13.63	-13.69	-13.67	-13.66	-13.08	-13.70	-13.65
	-13.75	-13.59	-13.74	-13.59	-13.75	-13.04	-13.73	-13.58
ξ_D	**-7.39**	**-7.06**	-7.39	7.06	-4.55	4.13	-4.67	4.51
	-7.44	7.11	-7.44	7.11	-4.61	4.19	-4.72	4.56
ξ_E	-9.80	-4.64	**-9.80**	**-4.64**	-6.97	1.96	-7.09	2.09
	-9.87	4.68	-9.86	4.69	-7.02	2.02	-7.15	2.14
ξ_F	-29.07	-19.70	-29.07	-19.70	**-29.07**	**−19.27**	-28.13	-20.68
	-29.28	-19.58	-29.21	-19.58	-29.24	-19.14	-29.22	-19.58
ξ_G	-29.04	-19.86	-29.20	-19.66	-28.28	-19.68	**-34.35**	**−13.87**
	-34.58	-13.92	-34.55	-13.82	-33.77	-14.66	-34.53	-13.72
Avg. n	142,873							

Notes: Characterization of the Confidence Set (CS) for the set estimates from the inequality model. I use the GMS procedure and perform a grid search around the vertices of the set estimate. γ_0 denotes disutility from driving time and ξs denote store fixed effects. α, disutility from expenditure is normalized to -1. The upper-left block of the table reports the highest and lowest values as well as a 95% confidence bounds moving away from the vertex of the identified set where they are located. The other blocks of table (5) report the value of the other parameters at the specific vertices, as well as a 95% confidence bounds moving away from the vertex along each of the parameters' dimension.

A.2 Price Indices

In the present paper price indices are used in the determination of expenditure on a bundle at an alternative store. As expenditure on the specific bundle affects the utility of a consumer on top and beyond the "price level" at a store, key in the construction of indices is to capture the consumer's expected difference in expenditure for each specific bundle. In general households are extremely different both in the category expenditure mix and in the within category product mix, and the construction of trip and household specific price indices seems most appropriate. The adequacy of one particular index ultimately depends on the nature of the stores in the market, whether they are close substitutes, whether they carry the same or entirely different sets of product, and the degree of price variation across stores and over time.

In the market I analyze supermarkets are fairly homogeneous, and products accounting for over 95% of expenditure are sold in more than one chain. I observe both aggregate and consumer level data on purchases in 31 packaged goods categories for the years 2001-2007. In my main model specifications I use category level price indices using aggregate data on sales. I consider aggregation both at a weekly level and at a larger period of aggregation; as my aim is to capture expected difference in expenditure I use periods of 5 weeks (for a total of 73 periods in 365 weeks). The use of a larger time span is convenient in that it smooths out the effect of temporary promotional sales; this is important because a weekly index would make the store offering a promotional sale relatively cheaper than another, even when the consumers expect it to be more expensive. Additionally, as not all items available in a store are sold in a given week, a larger time period increases significantly the number of items effectively observed.

I construct category price indices as a weighted average of log-deviations of items' prices at the store from a reference average price. To avoid that an average price is influenced by own price at the store, I construct average prices using national data from the 50 markets. For each store, I construct category indices including 1) all items sold at the store, 2) only items that are sold in more than one store and 3) only items that are sold in more than one chain. Although conceptually very different, the three indices are very similar in their values. When I construct alternative expenditures on trip bundles using the three indices, not only is the relative ranking of stores in terms of price the same for almost all of the trips, but also the resulting difference in expenditure is very close. This should not come to a surprise, as products available in more than one chain cover over 95% of expenditure.

The actual construction of the indices is as follows. Consider all the products i sold in category j at time t in the market, and construct market expenditure weights w_{ijt}. For all items i construct average price p_{ijt}; similarly, if the item is sold at time t in supermarket s, construct average price at store s, p^s_{ijt}. Let s_j denote the set of products sold at time t at store s. The price index for catagory j in time t at store s is given by

$$p_{sjt} = \frac{\sum_i \mathbf{1}(i \in s_j) w_{ijt}(1 + \log(p^s_{ijt}/p_{ijt}))}{\sum_i \mathbf{1}(i \in s_j) w_{ijt}}$$

I further construct an overall index p_{st} for the price level at the store as a category-expenditure weighted average of category price indices, that is

$$p_{st} = \sum_j w_{jt} p_{sjt},$$

where w_{jt} are market expenditure weights for category j. As I do not observe products and prices in for the residual category, I use the overall price level at the store p_{st}. Consider a trip (h, s, b, t) and let $e_j(b, p_{st})$ denote the observed expenditure in category

j for bindle b; for all alternative stores $s' \neq s$ I construct expenditure $e_{bs't}$ as

$$\tilde{e}_{bs't} = \sum_j e_j(b, p_{st})\big(p_{js't}/p_{jst}\big).$$

Results are presented using 5-week indices based only on items that are sold in more than one chain.

A.3 Proof of Proposition 2.3.4

By Proposition 2.3.2, $p_j = 1/2$ whenever $q_j = 0$. Let $i = \min_j \{j : q_j > 0\}$; by Proposition 2.3.1 $p_i < 1/2$. Let $j \geq i$.

- Suppose $p_j > 0$; then it must be the case that

$$\frac{\partial \Pi(p)}{\partial p_j} = 0 \quad \text{and} \quad \frac{\partial \Pi(p)}{\partial p_{j+1}} \leq 0 \tag{A.6}$$

  Let $q_{j+1} = q_j + \Delta q$, and $p_{j+1} = p_j + \Delta p$; substitute q_{j+1} and p_{j+1} in (3). We get:

$$\frac{\partial \Pi(p)}{\partial p_{j+1}} = C_1 - 2\Delta p\Big(q_j(2-p_j) + \Delta q + C_2\Big) + \Delta q\Big(1 - 2p_j + C_3 - p_j(1-p_j)\Big)$$

$$\frac{\partial \Pi(p)}{\partial p_j} = C_1 + \Delta p\Big(-2q_j(1-2p_j) + q_j p_j\Big) + (\Delta p)^2 q_j + \Delta q\big(1 - p_j - \Delta p\big)(1-2p_j)$$

  where C_2 and C_3 are positive functions of prices independent from p_j, p_{j+1}, q_j and q_{j+1}.

We need to show that $\Delta p \leq 0$.

- Suppose $\Delta q > 0$; for (A.6) to hold we have that

$$\frac{\partial \Pi(p)}{\partial p_{j+1}} - \frac{\partial \Pi(p)}{\partial p_j} \leq 0 \tag{A.7}$$

  Rearranging terms we get

$$\begin{aligned}\frac{\partial \Pi(p)}{\partial p_{j+1}} - \frac{\partial \Pi(p)}{\partial p_j} = \quad & (\Delta p)^2 \underbrace{q_j}_{>0} + \Delta p \Delta q \underbrace{(3 - 2p_j)}_{>0} + \Delta q \underbrace{(p_j^2 + C_3)}_{>0} \\ & + \Delta p \underbrace{(2C_2 + 2q_j(1 + p_j))}_{>0} \leq 0;\end{aligned}$$

  therefore $\Delta p < 0$.

- Suppose instead $\Delta q = 0$. Let $p' = (p_1, ..., p_{j-1}, 0, 0, p_{j+2}, ...)$; since $q_j = q_{j+1}$ and $\partial \Pi(p)/\partial p_j$ and $\partial \Pi(p)/\partial p_{j+1}$ are decreasing in all prices, and the optimal p_j is positive, it must be the case that

$$\left.\frac{\partial \Pi(p)}{\partial p_{j+1}}\right|_{p'} > 0,$$

  and therefore, $\partial \Pi(p)/\partial p_{j+1}$ must equal zero at optimum. But then equation (A.7) must hold with equality, and since $\Delta q = 0$ it reduces to

$$\frac{\partial \Pi(p)}{\partial p_{j+1}} - \frac{\partial \Pi(p)}{\partial p_j} = (\Delta p)^2 \underbrace{q_j}_{>0} + \Delta p \underbrace{(2C_2 + 2q_j(1 + p_j))}_{>0} = 0;$$

  so $p_j = p_{j+1}$.

- Suppose $p_j = 0$; it must be that

$$\frac{\partial \Pi(p)}{\partial p_j} = \big(q_j + q_{j+1}(1 - p_{j+1}) + C_2)\big) - q_j p_{j+1}(1 - p_{j+1}) - q_j C_3 \leq 0,$$

where C_2 and C_3 are defined as before. Now consider the FOC for p_{j+1}:

$$\frac{\partial \Pi(p)}{\partial p_{j+1}} = (1-2p_j)\big(q_j + q_{j+1} + C_2)\big) - q_{j+1}C_3$$
$$= \frac{\partial \Pi(p)}{\partial p_j} - \underbrace{q_j(p_{j+1} + p_{j+1}^2)}_{\geq 0} - \underbrace{\Delta q C_3}_{\geq 0} - \underbrace{p_{j+1}(q_{j+1} + 2C_2)}_{\geq 0} \leq \frac{\partial \Pi(p)}{\partial p_j}.$$

If either $\Delta q > 0$ or $p_{j+1} > 0$ the disequality will be strict,, hence $p_{j+1} = 0$.

A.4 Proof of Proposition 2.3.6

W.l.g. consider the case of firm A. Suppose for some $q_j > 0$, $p_{Aj} \geq 1/2$. The proposition holds if $\partial \Pi_A / \partial p_{Aj} \leq 0$ for all $p_{Aj} \geq 1/2$ and all p_{Bj}.

- Suppose $p_{Aj} > 1 - p_{Bj}$. Then

$$\begin{aligned} \frac{\partial \Pi_A}{\partial p_{Aj}} &= q_j \frac{\partial}{\partial p_{Aj}} \left(p_{Aj} \overline{D_j} \right) + (1 - 2p_{Aj}) \sum_{l \neq j} q_l \overline{D_l} \\ &\leq q_j \big(3p_{Aj}^2 - 4p_{Aj} + 1) \leq 0 \end{aligned}$$

- Suppose instead $p_{Aj} \leq 1 - p_{Bj}$. Then $p_{Bj} \leq 1/2$
  - If $p_{Aj} \geq 1/2$ and $p_{Bj} \geq -1/2$

$$\begin{aligned} \frac{\partial \Pi_A}{\partial p_{Aj}} &= q_j \frac{\partial}{\partial p_{Aj}} \left(p_{Aj} \overline{D_j} \right) + (1 - 2p_{Aj}) \sum_{l \neq j} q_l \overline{D_l} \\ &\leq 9p_{Aj}^2 - 12p_{Aj} - 4p_{Aj}p_{Bj} + 3 + 2p_{Bj} - p_{Bj}^2 \\ &\leq 0 \end{aligned}$$

– If $p_{Aj} \geq 1/2$ and $p_{Bj} \leq -1/2$

$$\frac{\partial \Pi_A}{\partial p_{Aj}} = (1 - 2p_{Aj}) \sum_{l \neq j} q_l \overline{D_l} < 0.$$

Bibliography

Anderson, S. P., and A. De Palma (2006): "Market Performance with Multi-product Firms," *Journal of Industrial Economics*, 54(1), 95–124.

Andrews, D., S. Berry, and P. Jia (2004): "Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location," Discussion paper, Cowles Foundation, Yale University.

Andrews, D., and X. Shi (2012): "Inference Based on Conditional Moment Inequalities," Discussion paper, Cowles Foundation, Yale University.

Andrews, D. W. K., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78(1), 119–157.

Barsky, R., M. Bergen, S. Dutta, and D. Levy (2001): "What Can the Price Gap between Branded and Private Label Products Tell Us about Markups?," NBER Working Papers 8426, National Bureau of Economic Research, Inc.

Baye, M. R., and J. Morgan (2004): "Price Dispersion in the Lab and on the Internet: Theory and Evidence," *RAND Journal of Economics*, pp. 449–466.

Beresteanu, A., and P. Ellickson (2006): "The Dynamics of Retail Oligopoly," *Duke University.*

Beresteanu, A., and F. Molinari (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, 76(4), 763–814.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–90.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela (2008): "Database Paper—The IRI Marketing Data Set," *Marketing Science*, 27, 745–748.

Burnett, P. (1973): "The Dimensions of Alternatives in Spatial Choice Processes," *Geographical Analysis*, 5(3), 181–204.

Carpenter, J. M., and M. Moore (2006): "Consumer demographics, store attributes, and retail format choice in the US grocery market," *International Journal of Retail and Distribution Management*, 34(6), 434–452.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

Chevalier, J. A., A. K. Kashyap, and P. E. Rossi (2003): "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93(1), 15–37.

Crawford, G., and A. Yurukoglu (2011): "The welfare effects of bundling in multichannel television markets," .

Davis, P. (2006): "Spatial competition in retail markets: movie theaters," *The RAND Journal of Economics*, 37(4), 964–982.

Ellickson, P. B. (2007): "Does Sutton apply to supermarkets?," *The RAND Journal of Economics*, 38(1), 43–59.

Ellickson, P. B., and S. Misra (2008): "Supermarket Pricing Strategies," *Marketing Science*, 27(5), 811–828.

Ellison, G. (2006): "Bounded rationality in industrial organization," *Econometric Society Monographs*.

Feichtinger, G., A. Luhmer, and G. Sorger (1988): "Optimal Price and Advertising Policy for a Convenience Goods Retailer," *Marketing Science*, 7(2), pp. 187–201.

Griffith, R., and H. Harmgart (2008): "Supermarkets and Planning Regulation," Discussion Paper 6713, C.E.P.R. Discussion Papers.

Griffith, R., E. Leibtag, A. Leicester, and A. Nevo (2008): "Timing and Quantity of Consumer Purchases and the Consumer Price Index," NBER Working Papers 14433, National Bureau of Economic Research, Inc.

Hausman, J., and E. Leibtag (2009): "CPI Bias from Supercenters: Does the BLS Know that Wal-Mart Exists?," in *Price Index Concepts and Measurement*, NBER Chapters, pp. 203–231. National Bureau of Economic Research, Inc.

Hess, J. D., and E. Gerstner (1987): "Loss Leader Pricing and Rain Check Policy," *Marketing Science*, 6(4), pp. 358–374.

Ho, K., J. Ho, and J. H. Mortimer (2012): "The Use of Full-Line Forcing Contracts in the Video Rental Industry," *American Economic Review*, 102(2), 686–719.

Ishii, J. (2005): "Compatibility, competition, and investment in network industries: ATM networks in the banking industry," *Unpublished working paper*.

Katz, M. (2007): "Estimating supermarket choice using moment inequalities," *Harvard University, Mimeo.*

Keng, K. A., and A. S. C. Ehrenberg (1984): "Patterns of Store Choice," *Journal of Marketing Research*, 21(4), pp. 399–409.

Lal, R., and C. Matutes (1994): "Retail Pricing and Advertising Strategies," *The Journal of Business*, 67(3), pp. 345–370.

Leszczyc, P. T. L. P., A. Sinha, and H. J. P. Timmermans (2000): "Consumer store choice dynamics: an analysis of the competitive market structure for grocery stores," *Journal of Retailing*, 76(3), 323 – 345.

Osborne, M. J., and A. Rubinstein (1998): "Games with Procedurally Rational Players," *The American Economic Review*, 88(4), pp. 834–847.

Pakes, A. (2010): "Alternative Models for Moment Inequalities," *Econometrica*, 78(6), 1783–1822.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2011): "Moment inequalities and their application," Cemmap working papers.

Reutterer, T., and C. Teller (2009): "Store format choice and shopping trip types," *International Journal of Retail and Distribution Management*, 37(8), 695 – 710.

Rhee, H., and D. R. Bell (2002): "The inter-store mobility of supermarket shoppers," *Journal of Retailing*, 78, 225–237.

Rosen, H. S., and K. A. Small (1981): "Applied Welfare Economics with Discrete Choice Models," NBER Working Papers 0319, National Bureau of Economic Research, Inc.

Smith, H. (2004): "Supermarket Choice and Supermarket Competition in Market Equilibrium," *Review of Economic Studies*, 71(1), 235–263.

Sobel, J. (1984): "The Timing of Sales," *The Review of Economic Studies*, 51(3), 353–368.

Spiegler, R. (2006): "Competition over agents with boundedly rational expectations," Open Access publications from University College London http://discovery.ucl.ac.u, University College London.

Train, K. (2009): *Discrete Choice Methods with Simulation*. Cambridge University Press.

Varian, H. R. (1980): "A model of sales," *The American Economic Review*, 70(4), 651–659.

Weinstein, J., and A. Ambrus (2006): "Price Dispersion and Loss Leaders," *Theoretical Economics*.

Wrigley, N., and R. Dunn (1984): "Stochastic panel-data models of urban shopping behaviour: 2. Mulistore purchasing patterns and the Dirichlet model," *Environment and Planning A*, 16(6), 759–778.

Wrigley, N., and R. Dunn (1985): "Stochastic panel-data models of urban shopping behaviour: 4. Incorporting independent variables into the NBD and Dirichlet models," *Environment and Planning A*, 17(3), 319–331.