THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par:
Université de Toulouse I (UT1 Capitole)

Discipline ou spécialité:
Sciences Economiques

Présentée et soutenue par

Jorge Florez Acosta

le: 21 juillet 2015

Titre:
Essays on the Empirical Analysis of Grocery Retailing and Consumer Shopping behavior

---

École doctorale :
Toulouse School of Economics

Unité de recherche:
GREMAQ - TSE

Directeurs de thèse:
Pierre Dubois (Professeur, Université Toulouse I)
Bruno Jullien (Directeur de recherche, CNRS-TSE)

Rapporteurs:
Helena Perrone (Professeur, Universitat Pompeu Fabra)
Andrea Pozzi (Professeur, Einaudi Institute for Economics and Finance)

Membres du jury:
Philippe Bontems (Professeur, INRA-TSE)
Bruno Jullien (Directeur de recherche, CNRS-TSE)

L'Université n'entend ni approuver, ni désapprouver les opinions particulières du candidat.

Toulouse School of Economics,
Université Toulouse I - Capitole,
21 allée de Brienne,
31015 Toulouse, Cedex 6, France.

A mis padres Norma y Jorge Hernán.
A mi hermano Carlos Andrés.

"—I've been working on this poem for 12 years!
–*Really?*
—There's a lot of expectation. I don't wanna disappoint my fans.
–*May I?*

> *The grass so green*
> *Skies so blue*
> *Spectre is really great!*

–*It's only three lines long...*
—This is why you should never show a work in progress."

Big Fish, Tim Burton, 2003

Acknowledgments

There are many people who got involved over the course of this dissertation by giving me support and encouragement, through academic advise and stimulating discussions or by helping in nonacademic formalities and processing that is always necessary. I want to thank all of them.

First, I thank my supervisors Bruno Jullien and Pierre Dubois for their invaluable advice and guidance. Bruno always challenged me with his questions and gave me a good sense of what is academic research. He would never underestimate a research question or idea without first carefully analyzing its economic sense and potential impact with his outstanding capacity of understanding a wide range of topics in economics. Perhaps one of his best advises was that I should learn to work wit data given that economic research was increasingly turning to empirical analyses and testing of theory results with data. Pierre paired thorough direction in the work with data and the empirical implementation of economic questions and structural models, with pointed questions and critiques. His great support, encouragement and time were key during the preparation of my Job Market talks. It was a challenging but fascinating experience to work with both a theorist and an econometrician. Trying to reconcile the two perspectives was not always an easy task but it was certainly exciting and key to learn how to do rigorous structural modeling yet empirically tractable.

I benefited greatly from conversations with Yinghua He, Fabian Bergès, Patrick Rey, Otto Toivanem, Allan Collard-Wexler, Michelle Sovinsky, Liran Einav and Rachel Griffith. I also benefited from interesting discussions with fellow students Laura Lasio, Daniel Herrera, Margaret Leighton, Milena Petrova, Shagun Khare, Yann Kervinio, Ananya Sen and Carlos Cañón; as well as students not associated with TSE: Morten Saethre, Martin O'Conell and Alessandro Iaria. I thank the faculty and students in charge of the Applied Microeconomics Workshop at the TSE over the course of this dissertation. There, I presented the three papers of this thesis in very early stages and obtained important feedback. I also thank Aude Schloesing and Laurance Delorme for their invaluable administrative work and support to make PhD students life easier.

I am grateful to David Tobón, Jorge Barrientos, Piedad Restrepo, Germán Valencia and Mauricio Alviar for their encouragement and support when I was conceiving the idea of making a PhD in economics. They not only believed I could undertake a PhD formation of high quality but were also very enthusiastic on helping with recommendation letters and other formalities. This project would not have been possible without the generous financial support of Banco de la República and Colciencias. I acknowledge the *Institut National de la Recherche Agronomique*, INRA, for a three-year grant and for giving me access to the data used in this dissertation.

I am deeply grateful to Andrea Medina, who carefully gave me support, encouragement and plenty of life advice that allowed me to go through the hardest moments of this process with resolution and strength. Karoll Gómez has been generous and supportive over the course of this dissertation. In the last years she became a great support to successfully finish my dissertation. I was lucky to have a good web of friends. In particular I thank Aude Renard and Sebastian Olarte.

Finally, I am indebted to my family. My brother Carlos Andrés, a PhD candidate in Physics, passionate for science, literature and other intellectual matters since he was very young, has always been a key support and a great motivation to pursue my projects. My parents Jorge Hernán and Norma did a wonderful job giving us the foundation that allowed us to take on challenging projects. With very limited economic resources and a lot of effort, dedication and love they raised us convinced that education was the most important wealth we should acquire in our lives. They expected from us to obtain bachelor's degrees, but they never expected that we both got this far in our education attainment. I dedicate this thesis to them.

Abstract

The food and grocery retail industry is of great importance for its significant contribution to a country's GDP, the key role it plays in the supply chain, and the complex structure in which it has evolved: few differentiated great scale multiproduct retail chains that concentrate a big share of domestic retail sales, sophisticated logistics and distribution systems, and a multiplicity of store formats ranging from downtown convenience to big-box stores. In addition, this industry is characterized by the promotion of private labels and other non-price strategies to induce customer loyalty. The emergence and consolidation of hard-discount stores is as well an important characteristic of the modern retail sector. Retail chains are no longer simple intermediaries between manufacturers and consumers as they used to be in the past, they can be rivals of upstream firms as well. Even though markets are dominated by such large retailers, smaller competitors such as independent convenience stores, specialized stores, among others, still survive with a lesser share generally offering either narrower product lines of higher quality or by specializing in the supply of a specific category of products.

For all these reasons, the food and grocery retail industry has been of central attention for policy makers and economic research. A vast literature has covered a wide variety of topics from both theoretical and empirical perspectives. Yet, many interesting questions remain to be addressed. This dissertation aims at studying three topics concerning the grocery retailing industry using empirical methods. In particular, the emphasis has been put on three questions that have special relevance in the context of powerful retailers. Each of which is the subject of one of the three chapters of this thesis.

The first chapter focuses on supermarket loyalty programs and the impact on the demand for private labels. It is motivated is motivated by the observation that supermarkets often link loyalty rewards to private label purchases. Why supermarkets would do such thing? is the question this paper addresses using structural methods of demand estimation and data on french households' supermarket purchases. Loyalty programs (*LPs*) are by now a predominant short-run nonprice strategy in retailing markets. In France, for instance, supermarkets give loyalty rewards almost exclusively on private labels which are cheaper than their branded counterparts. Why do profit-maximizing retailers give rebates on their lower-price own brands? This paper empirically examines the link between supermarket LPs and private labels. Using scanner data on grocery purchases of French households and discrete-choice methods, I estimate brand-level demand accounting for household membership to loyalty programs. Results are consistent with the general wisdom: private labels are less valued products by all consumers relative to quality-equivalent national brands. However, members of loyalty programs have a larger valuation of private labels than non-members. Moreover, the more prone to subscribe to LPs a customer is, the larger her sensitivity to a price increase and the weaker the expected effects on the demand for private labels.

The second chapter, joint with Daniel Herrera Araujo, is inspired by a number of theory papers showing that when shopping costs, that rationalize the observed heterogeneity in consumer shopping patterns, are introduced in economic analysis, policy conclusions might change dramatically. We structurally identify consumer shopping costs —real or perceived costs of dealing with a store— using scanner data on grocery purchases of French households. We present a model of demand for multiple stores and products consisting of an optimal stopping problem in terms of individual shopping costs. This rule determines whether to visit one or multiple stores at a shopping period. We then estimate the parameters of the model and recover the distribution of shopping costs. We quantify the total shopping cost in 18.7 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 1.53 € and mean total transport cost of 17.1 € per trip. We show that consumers able to source three or more grocery stores have zero shopping costs, which rationalizes the low proportion of three-stop shoppers observed in our data. Theory predicts that when shopping costs are taken into account in economic analysis, some seemingly pro-competitive practices can be welfare reducing and motivate policy intervention. Such striking findings remain empirically untested. This paper is a first step towards filling this gap.

The third chapter, I empirically examine the role of nonlinear contracts between manufacturers and retail stores, and Resale Price Maintenance (RPM) on nominal price stability. It is widely accepted in the literature that the incomplete transmission of costs shocks into retail prices is explained by the existence of markup adjustment and price adjustment costs. The vertical conduct of the industry and the existence of vertical restraints such as RPM might introduce further price stickiness or reinforce it. I present a structural model of vertical relations between manufacturers and retailers allowing for nonlinear contracts and vertical restraints, and accounting explicitly for retail price rigidity by including fixed costs of price adjustment in retailer's profit function. Using micro data on sales of ready-to-eat breakfast cereals from a large supermarket chain in Chicago, I estimate demand, retrieve upstream and downstream markups, and compute bounds of retail price adjustment costs. Results show that the total costs the retailer bears for adjusting prices of its products in a year lie between 1.6% and 3% of its total revenue, on average.

Résumé

La grande distribution alimentaire est actuellement d'une telle importance, et cela grâce à sa contribution au PIB d'un pays, son rôle essentiel au sein de la structure verticale, et l'organisation assez complexe dont elle est devenue au cours des dernières décennies: peut de détaillants différenciés offrant un grand assortiment de produits disponibles à grande échelle et qui concentrent une partie considerable du marché domestique, des systèmes de distribution et de logistique bien complexe, ainsi qu'une variété de formats de magasin dont les magasins à proximité, les supermarchés et les hypermarchés sont les plus fréquents. De plus, le développement et promotion des marques de distributeur (MDD) ainsi que la mise en place des strategies visant la fidélisation des consommateurs sont devenues des caractéristiques marquantes dans cette industrie. Par ailleurs, l'émergence et consolidation des magasins hard-discount est également l'une des caractéristiques remarquables de la grande distribution à ce jour. Les détaillants ne sont plus des simples intermédiaires entre producteurs et consommateurs. Aujourd'hui, par contre, avec les MDD les détaillants sont devenus concurrents de ses fournisseurs. Même si les marchés sont dominés par des si grands détaillants, des petits concurrents, tels que des magasins à proximité independents et des magasins spécialisés (boulangeries, boucherie, etc.), sont toujours présents avec une petite part de marché mais avec un assortiment des produits de qualité supérieure en général.

C'est pourquoi la grande distribution est de ce fait l'objet de nombreuses interventions de politique et de la recherche économique. En effet, une importante littérature aussi bien théorique qu'empirique a abordé une grande variété des sujets. Cependant, il reste un nombre considerable des questions à être l'objet de recherche. Cette thèse vise à étudier trois questions concernant la grande distribution et le comportement d'achat des consommateurs, dans le cadre de détaillants avec pouvoir de marché, en utilisant des méthodes empiriques.

Le premier chapitre est dedié à l'étude des programmes de fidélité des supermarchés et leur impact sur la demande de marques de distributeur (MDD). Souvent les supermarchés lient les avantages fidélité à l'achat en marques de distributeurs, quelles sont les motivations des supermarchés à faire cela? C'est la question que cet étude cherche à répondre d'un point de vue empirique. Je travail sur des donnés extraites d'un panel représentatif des consommateurs concernant les achats des ménages français, et l'utilisation d'une méthode structurelle d'estimation de demande. Les résultats sont conformes aux faits: les MDD sont des produits moins préférés vis-à-vis les marques nationales (MN) de même qualité. Cependant, la carte fidélité a, en effet, un impact positive sur le choix du consommateur: ceux qui portent une carte fidélité ont une probabilité supérieur de choisir des MDD que ceux qui ne l'ont pas. Par ailleurs, l'impact d'un programme de fidélité sur la demande des MDD est moins importants chez les détenteurs de plusieurs cartes.

Le deuxième chapitre, co-écrit avec Daniel Herrera Araujo, vise à mesurer les coûts d'achat des consommateurs à partir des données de panel concernant les achats des ménages en France. Quand l'analyse économique tient compte des coûts d'achat, que rationalisent l'hétérogénéité observée en nombre des enseignes visitées par les consommateurs, les conclusions de politique publique peuvent changer remarquablement. Nous identifions les coûts d'achat du consommateur dans le cadre d'un modèle structurel de demande en plusieurs enseignes ainsi qu'en plusieurs produits, qui lie le choix optimale du nombre de supermarchés à visiter (un seule ou plusieurs) aux coûts d'achat. Nous estimons les paramètres du modèle et mesurons le coût d'achat total moyen en 18,7 € par enseigne visitée. Deux quantités y sot compris: le coût fixe moyen, 1,53 € et le coût de transport moyen 17,1 € par visite.

Le troisième chapitre porte sur l'analyse empirique du rôle des tarifs binômes et la fixation du prix de vente (RPM, d'après l'expression anglo-saxonne Resale Price Maintenance) dans la stabilité des prix de vente. Il est largement reconnu dans la littérature économique que la transmission incomplète des chocs de coûts aux prix de vente est expliquée par l'ajustement des marges ainsi que les coûts d'ajustement des prix. Les relations entre fournisseurs et distributeurs et le RPM peuvent renforcer la rigidité des prix. Je présente un modèle structurel de relations verticales dont des tarifs binômes peuvent être adoptées ainsi que le RPM. Ce modèle tient compte de la rigidité des prix de vente à travers des coûts fixes d'ajustement des prix qui sont ajoutés au profit du détaillant. En utilisant des données concernant les ventes de marques de céréales pour le petit déjeuner dans une grande chaîne des supermarchés au Chicago, j'estime la demande, récupère les marges et calcule les limites supérieure et inférieure de l'intervalle que contienne les vrais coûts d'ajustement. Les résultats obtenus montrent que ces coûts représentent en moyenne entre 1.6% et 3% des revenus totales du distributeur par an.

Contents

List of figures

List of tables

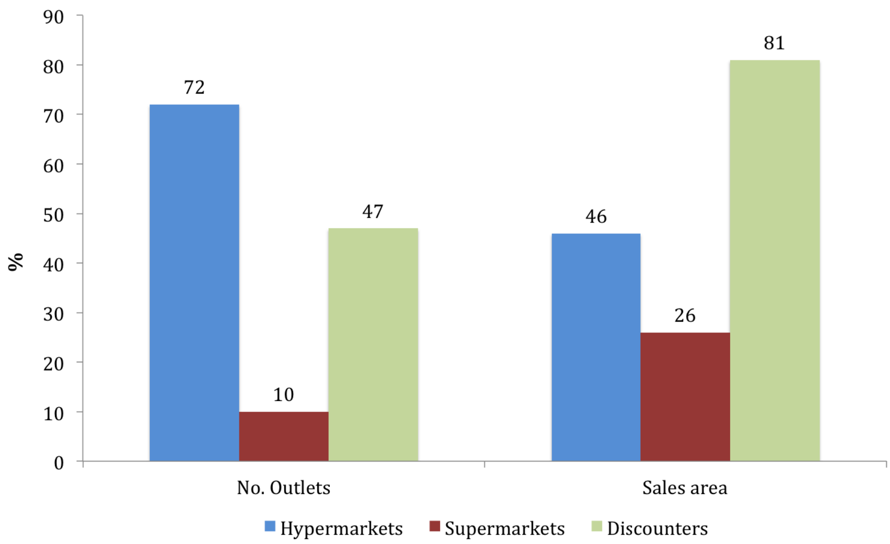
General Introduction

The food and grocery retail industry is of great importance for its significant contribution to a country's GDP, the key role it plays in the supply chain, and the complex structure in which it has evolved. In a world in which vertical relations are often more complex than simple linear pricing contracts, in the last decades retailing chains have shown to be very interested in increasing their market power and bargaining position vis-à-vis manufacturers. In line with this, in most modern economies we can list as main characteristics of retail sectors, the following: high concentration of domestic retail sales by a few differentiated great scale multiproduct retail chains; sophisticated logistics and distribution systems; the proliferation of great scale superstores, that can cover a sales area of 20,000 square meters or more and carry more than 200,000 different brands at the same time, offering parallel services such as shopping malls, beauty salons, restaurants, car wash, gas stations, and recreation grounds for kids, as an attempt to attract customers through one-stop shopping; and the promotion of private labels and other non-price strategies to induce customer loyalty. The emergence and consolidation of hard-discount stores, that provide an alternative for all those consumers looking for lower prices mostly on basic products, is as well an important characteristic of the modern retail sector.

In the European Union, for instance, the retail sector accounts for 4.3% of the Gross Value Added in the EU economy, around 8% of employment and the retail sales represent 54% of the edible grocery sales (European Commission, 2014). In the United Kingdom, the grocery sales account for 51.3% of UK's retail sales, 14% of all UK employment (IDG, 2015). On the other hand, the share of household expenditure on food and non-alcoholic beverages in the EU including the UK was on average 13% of total expenditure in 2012. In 13 countries of the European Union (among which we have France, the UK, Germany, Sweden, Finland, and the Netherlands), the leading 5 retailers of each country concentrated a market share greater than 60% (European Commission, 2014). Such phenomenon has been possible due to the following reasons: *i*) zoning regulations limiting internal growth and favoring mergers, acquisitions and organization of retailers in alliances and corporations of international scope (Chen and Rey, 2012; European Commission, 2014). *ii*) the increase in the constitution of buying groups, which are groups of retailers that get together to improve their bargaining power and get better purchasing conditions through higher volumes, to reduce purchasing costs of NBs and PLs. Finally, *iii*) consumer habits have been that depending on their time availability, transportation costs and taste for shopping tend to favor one-stop shopping for a broad range of products (European Commission, 2014). By carrying a large range of different products and brands, stores make the shopping experience less costly to consumers.

In the last decade, larger store formats experienced a remarkable proliferation. According to the European Commission (2014), Hypermarkets increased the number of outlets in 72% between 2000 and 2011 and the sales area in 46% in the same period. The number of supermarket outlets increased by 10% and the sales area of this format raised by 26%. On the other hand, discounters experienced the greatest growth in sales area in the same period with 81% while the number of outlets of this retail form increased by 47% (see Figure 1).

Figure 1: Increase of number of outlets and sales area by store format (in percentage), 2000-2011



Source: European Commission, 2014.

Retail chains are no longer simple intermediaries between manufacturers and consumers as they used to be in the past, but act as rivals of upstream firms as well. Private labels (PL), i.e. products distributed exclusively by a retail chain under its own brand, have evolved from low-quality generic products that captured demand from low willingness-to-pay consumers to more attractive products of comparable quality to that of manufacturer branded (known as National Brands, NB) and of lower price. In EU countries, in particular, PLs have steadily increased their market share and European consumers are changing their perception of PLs from rough substitutes of NBs provided they are a similar-quality alternative with lower prices.[1] In fact, penetration of private labels exceeds 40% in some European countries such as

[1] Evidence shows that PLs are supplied at a 20% lower price on average relative to NBs (Berges-Sennou et al., 2009).

Switzerland and the UK and in general is high in the region (European Commission, 2014). A traditional motive claimed by retailers to promote private labels is that they are more profitable than reselling NBs. Other important motives are: they are key to increase customer loyalty, because they are exclusive products. And though there are other supermarkets with private labels on the same products that can be seen as close substitutes, the significant differentiation across retailers is enough to make those products considerably differentiated. By building customer loyalty through private labels and loyalty programs, retailers may be increasing their market power to be used against upstream as well as downstream markets.

Yet, smaller competitors such as independent convenience stores, specialized stores, among others, still survive in a market dominated by large chains with a lesser share generally offering either narrower product lines of higher quality or by specializing in the supply of a specific category of products (such as wine shops, vegetables markets, butcheries, etc.).

For all these reasons, the food and grocery retail industry has been of central attention for policy makers and economic research. A vast literature has covered a wide variety of topics (such as retailer competition, vertical relations, consumer product and supermarket choice, consumer behavior in the presence of switching, search or shopping costs, the effects of new product introduction, the economics of private labels, etc.) from both theoretical and empirical perspectives by Marketing and Economics literature. Yet, many interesting questions remain to be addressed. This dissertation aims at studying three topics concerning the grocery retailing industry using empirical methods. In particular, the emphasis has been put on three questions that have special relevance in the context of powerful retailers: one is related to supermarkets nonprice strategies to make customers loyal and the effects on private labels; a second is concerns consumer choice between one- or multi-stop shopping in the presence of shipping costs; and a third relates to the problem of retail price rigidity in the presence of price adjustment costs in a context of vertical restraints. I do this in three chapters that I describe in the following.

In the first chapter I focus on the effects of loyalty programs on the demand for PLs. This products have become a key element of supermarkets strategic interaction with rivals at the horizontal dimension and suppliers (also rivals in some sense) at the vertical dimension. There are a number of nonprice competition strategies that have been used by supermarkets presumably to promote PLs. This is the case of supermarket loyalty programs that have become a predominant strategy for supermarkets. In fact, there are cases in which retail chains give loyalty rewards almost exclusively on PLs purchases, as in France for instance. Why do profit-maximizing retailers give rebates on their lower-price own-brands? One might think of many reasons as to why retailers might be putting extra efforts on PLs promotion: they are more profitable than NBs, along with loyalty programs PLs may help increase customers' feelings of store as well as brand loyalty; as a consequence, the share of one-stop shoppers may as well increase as LPs raise artificial switching costs, reducing customers' incentives to patronize other stores. LPs may even induce what Armstrong and Vickers (2010) call *excessive loyalty*, which makes customers make less preferred choices so as to enjoy rewards. Last, retailers might be using loyalty programs on PLs as a way to strengthen their bargaining position vis-à-vis manufacturers.[2]

Yet, key questions remain to be answered: Why do LPs put emphasis on private labels rather than on the full product range of the store? Are those LPs an effective nonprice strategy to boost PL demand? The objective of this paper is to empirically examine the effects of loyalty programs on the demand for PLs. My empirical strategy consists of using discrete-choice methods to estimate brand-level demand accounting for household membership to supermarket loyalty programs. In particular, I estimate a random coefficients Logit model following the standard literature (Berry, 1994; Berry, Levinsohn and Pakes (BLP), 1995; and Nevo,2000a, 2001). I use a three-dimensional panel of quantities and prices for up to 13 brands of plain yogurt, purchased from the 6 largest supermarket chains in up to 94 departments of France, weekly in 2006. Household membership to loyalty programs is given by a household-supermarket specific indicator variable taking on 1 if the household is a member of a given supermarket LP. I deal with the correlation between membership to LPs indicator and the unobserved supermarket attributes using loyalty program characteristics.

Results confirm that private labels are, on average, less valued products relative to national brands. However, I find that the marginal valuation of PL products increases with subscription to the supermarket LP, which supports the believe that LPs might be used as a way to boost store-brands demand. Moreover, when customers hold simultaneously subscriptions to LPs of competing retailers, the expected effects become weaker, i.e. the marginal valuation of PLs decreases with the number of subscriptions and customers are more sensitive to price changes. Demand-side counterfactuals suggest that loyal customers benefit from when the market is not fully covered, as supermarkets might make efforts to improve loyalty rewards so as to attract nonmembers. On the other hand, making LPs prohibitively costly may harm consumers.

In the second chapter, joint with Daniel Herrera Araujo, we focus on supermarket and product choice in the presence of consumer shopping costs. As discussed previously, supermarkets have been putting great effort in increasing their sales area and number of superstores that integrate a variety of additional services to attract customers through one-stop shopping. Yet, evidence shows that an important proportion of consumers engage in the so-called *multi-stop* shopping, i.e. visit several different retailers within a week to buy different products. This heterogeneity might be explained by several factors such as preferences, demographics, geographic

[2] See Scott-Morton and Zettelmeyer (2004), Bergès-Sennou (2006), and Meza and Sudhir (2010).

location, information frictions, differentiated retailers, and time availability for shopping activities. Previous literature has introduced a concept that accounts for some (or most) determinants: *shopping costs*, defined as all real or perceived costs a consumer incurs when sourcing a grocery store (Klemperer, 1992, Klemperer and Padilla, 1997, Armstrong and Vickers, 2010, and Chen and Rey, 2012, 2013). Economic theory shows that in a context of multiproduct retailing and consumer shopping costs, several practices that would otherwise be considered competitive and good from a social welfare perspective can be less competitive. Is it possible to quantify shopping costs from observed consumer shopping behavior?

This research work provides a framework to assess the role of shopping costs in explaining heterogeneous shopping patterns. To do so we develop a structural model in the spirit of the main theoretical contributions on the topic. Consumer optimal shopping behavior is given by a threshold strategy where the choice between one- or multistop shopping depends on the size of individual shopping costs. We are able to take the model to data through parametric specifications of consumer utility and shopping cost along with some distributional assumptions on the unobserved shocks. We use scanner data on household grocery purchases in France in 2005, which is representative of French households and contains information on a wide product range and household demographics. As for grocery stores, an additional data set allows us to observe store characteristics and location.

By solving the implied optimal stopping problem of a consumer who needs to decide how many stores to source, we are able to recover the distribution of shopping costs. We quantify the total shopping cost in 18.7 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 1.53 € and the total transport cost of 17.1 €per trip to a given store. Moreover, we are able to compute the transport and total costs of shopping by store format. Transport and total costs of shopping are decreasing in the size of the stores, on average, as smaller formats are closer to downtowns. The largest total shopping cost, 24.7 €, are incurred by consumers who source big-box stores, because they are farther away from downtown. Sourcing a supermarket or a hard-discounter implies total costs of shopping of 14.3 € and 13.4 € per trip, respectively. Finally, the costs of sourcing a convenience store, 4.8 € per trip, are the lowest provided that they are located in downtown. We find that individuals who source three or more stores in a week have zero shopping costs. This might be an indicator that those households actually visiting more than two separate stores a week should have a strong preference for shopping. Finally, the predicted proportions of shoppers by number of stops are 90.1% of one-stop shoppers, 9.7% of two-stop shoppers and only 0.26% do three-stop shopping.

Finally, the third chapter addresses the question of nominal price rigidity and its determinants. We are mainly used to hear that price stickiness is a concern of Macroeconomics. However, a widely held view is that aggregate price inertia is determined by individual goods prices rigidity (Midrigan, 2011; Kehoe and Midrigan, 2012), and there is a lot microeconomic analysis can contribute to the understanding of the determinants of such phenomenon. The low degree of retail price responsiveness to costs shocks is best known as *incomplete pass-through*. Does the structure of the industry matters when it comes to explain individual goods price stickiness? Engel (2002) points out three sources of such incomplete transmission that are related to the industry structure and firms' strategic behavior: the existence of local costs, markup adjustment either by retailers or manufacturers or both, and nominal price rigidity. A growing empirical Industrial Organization (IO) literature has made important advances in understanding incomplete pass-through by looking at how vertical relations and vertical restraints such as Resale Price Maintenance (RPM), a practice through which manufacturers determine prices retailers charge to consumers, shape markup adjustment. Less attention has been put on price rigidity. We know from microeconomic theory that RPM makes prices less responsive to costs shocks (Jullien and Rey, 2007). In this chapter I empirically examine the role of nonlinear contracts between manufacturers and retail stores, and RPM on nominal price stability. My focus is, however, on a totally different sector: the ready-to-eat (RTE) breakfast cereal industry.

The empirical IO literature has study the main sources of price rigidity. I set out a modeling framework that closely relates to two of those contributions. Goldberg and Hellerstein (2013) address the question of the incomplete transmission of exchange rate shocks to local currency prices of imported beer in the U.S. They set out a structural model in which linear tariffs characterize the industry's vertical conduct. Unless previous research, they control explicitly for price rigidity by including fixed costs of price adjustment in the profit functions of both manufacturers and retailer in a static framework. On the other hand, Bonnet et al. (2013) are the first to empirically investigate the role of nonlinear pricing and RPM on incomplete pass-through of costs into retail prices, focusing on how the vertical structure of the industry affects strategic markup adjustment.

My empirical strategy relies on the consistent estimation of demand, Using data from *Dominik's Finer Foods*, a large supermarket chain in Chicago, that contains among other things information on weekly prices and quantities sold at the universal product code (UPC) level for 224 weeks. I set out three structural models of supply (linear pricing, simple two-part tariffs, and two-part tariffs with RPM) in a context of static Nash-Bertrand oligopolistic competition with several competing manufacturers and a single retailer that carries all products. The retailer faces fixed costs of repricing whenever it decides to adjust the price of a product. At each period, the retailer weighs the costs and the benefits of changing the price of each product and makes one of two decisions: if benefits exceed costs, it sets a new price that maximizes current period profit; otherwise, it keeps the same price from previous period which implies a deviation from first order conditions of the static profit maximization problem.

I find that the linear pricing specification of the supply side gives biased results for price adjustment costs relative to two-part tariffs with RPM. In fact, under linear pricing I obtain that the retail chain is willing to change the price of a product if it obtains, on average, an extra profit of at least US \$109, whereas under two-part tariffs with RPM this amount is US \$98.84. On the other hand, I obtain that adjustment costs are bounded above by US \$447 on average under either supply conduct. Simple two-part tariffs (i.e. without RPM) give similar results to those of linear pricing. To have an idea of the relative importance of these magnitudes, I compute the share of the sum of each lower and upper bounds of adjustment costs on retailer's total revenue for the entire period considered here (224 weeks). I find that the share of adjustment costs on total revenue is 7.27% for lower bounds and 10.12% for upper bounds in the linear tariffs case, and to 6.5% and 10.14% respectively, in the two-part tariffs case.

Chapter 1

An empirical analysis of loyalty programs and the demand for private labels

Abstract: Loyalty programs (*LPs*) are by now a predominant short-run nonprice strategy in retailing markets. In some countries, such as France, supermarkets give loyalty rewards almost exclusively on private labels which are cheaper than their branded counterparts. Why do profit-maximizing retailers give rebates on their lower-price own brands? This paper empirically examines the link between supermarket LPs and private labels. Using scanner data on grocery purchases of French households and discrete-choice methods, I estimate brand-level demand accounting for household membership to loyalty programs. Results are consistent with the general wisdom: private labels are less valued products by all consumers relative to quality-equivalent national brands. However, members of loyalty programs have a larger valuation of private labels than non-members. Moreover, the more prone to subscribe to LPs a customer is, the larger her sensitivity to a price increase and the weaker the expected effects on the demand for private labels.

JEL codes: D12, L13, L22, L81.

Keywords: Grocery retailing, supermarket chains, loyalty programs, private labels, oligopolistic competition, discrete choice models, random coefficients.

1.1 Introduction

Store brands, also known as private labels (hereafter, *PLs*),[1] have become a key element of supermarkets strategic interaction with rivals at the horizontal dimension and suppliers (also rivals in some sense) at the vertical dimension. Originally perceived as rough substitutes of national brands (*NBs*), they are in most cases as attractive for consumers as NBs, provided they are a similar-quality alternative with lower prices.[2] In fact, PLs market share has been increasing steadily in the last decade, attaining important levels in European countries such as the United Kingdom (46%), Germany (35%) and Spain (33%) –Steenkamp and Geyskens (2013), as well as in the United States (19.1%) –Turcic et al., 2013.

Economics and marketing literature have widely studied and empirically supported the positives and the negatives of PLs.[3] However, less attention has been put on nonprice competition strategies and their interaction with private labels. This is the case of supermarket loyalty programs that have become a predominant strategy for supermarkets and a way to promote PLs. In fact, there are cases in which retail chains give loyalty rewards almost exclusively on PLs purchases, as in France for instance.[4] Why do profit-maximizing retailers give rebates on their lower-price own-brands? One might think of many reasons as to why retailers might be putting extra efforts on PLs promotion: they are more profitable than NBs, along with loyalty programs PLs may help increase customers' feelings of store as well as brand loyalty; as a consequence, the share of one-stop shoppers may as well increase as LPs raise artificial switching costs, reducing customers' incentives to patronize other stores. LPs may even induce what Armstrong and Vickers (2010) call *excessive loyalty*, which makes customers make less preferred choices so as to enjoy rewards. Last, retailers might be using loyalty programs on PLs as a way to strengthen their bargaining position vis-à-vis manufacturers.[5]

[1] They are retailers' own-branded products supplied exclusively in their stores as opposed to national brands, which are the regular highly advertised and generally everywhere-available products.

[2] Evidence shows that PLs are supplied at a 20% lower price on average relative to a quality-equivalent NBs (Bergès-Sennou, 2009).

[3] They increase the range of products available for consumers, intensify intra-brand competition and may stimulate upstream competition, among others. On the other hand, they may help retailers increase market and buyer power.

[4] Most retailing chains give lagged rebates based on current purchases of selected PL products, conditional on the previous subscription to the program. Rebates are accumulated in customer's account and after a given time/money threshold is crossed, the acquired amount of money is given back to customers as a purchase coupons to be spent in any of the retailer's stores. Some programs, such as that of Casino's, work slightly different as they give "miles" to customers according to a predetermined exchange rate, and have a catalog where members can pick a gift according to the accumulated number of miles.

[5] See Scott-Morton and Zettelmeyer (2004), Bergès-Sennou (2006), and Meza and Sudhir (2010).

Yet, key questions remain to be answered: Why do LPs put emphasis on private labels rather than on the full product range of the store? Are those LPs an effective nonprice strategy to boost PL demand? The objective of this paper is to empirically examine the effects of loyalty programs on the demand for PLs.

Previous research has provided several explanations as to why retailers offer such costly programs, instead of just competing on prices. They can be summarized in two categories, namely, consumer retention and the exercise of market power. On the former, LPs allow retailers to retain customers and induce repurchase, as it is a way to impose artificial switching costs on customers and they are more likely to come back when there is a promised price reduction. Cremer (1984) finds that, as long as consumers have elastic participation, the monopolist optimal strategy is to charge a lower price to repeat buyers instead of precommitting to a second-period price. Klemperer (1987a, 1987b) lists repeat-purchase coupons and "frequent-flyer" programs (FPPs) as examples of artificial or contractual switching costs that make rational customers prone to display brand loyalty as demand becomes more inelastic and firm's market power increases. Caminal and Matutes (1990) provide the rational behind the advantages of loyalty discounts relative to price reductions. Using a similar framework as Klemperer (1987b), they endogenize switching costs to show that coupons valid for next period purchases perform better than price precommitment as they allow firms to get higher overall profits and reduce the intensity of competition. In line with these results, Chen and Pearcy (2010) find that for those markets in which customers do not have a strong preference for a particular brand (because good substitutes are available or preferences are likely to change for future choices), firms tend to enroll them in loyalty programs to reward repeat purchases and discourage brand switching.

Concerning price discrimination, LPs can be thought of as an explicit discriminatory device as only those customers who join the program and present a card at each purchase occasion are able to enjoy the benefits. When firms reward repeat buying, they are at the same time setting a differentiated future price schedule that charges new customers the full tariff whereas repeat buyers get a lower one if they redeem the coupon they have got in a previous period. A similar conclusion can be obtained from behavior-based price discrimination (see, for example, Caillaud and De Nijs, 2011).[6] Although most contributions on this topic conclude that firms should offer lower prices to new customers, Caillaud and De Nijs (2011) get the opposite result under the assumption that some firms cannot distinguish between old and new customers. Hence, they reward loyalty by offering lower tariffs to previous customers and charge full prices to the new ones.

My empirical strategy consists of using discrete-choice methods to estimate brand-level demand accounting for household membership to supermarket loyalty programs. In particular, I estimate a random coefficients Logit model following

[6] It consists of offering different prices to different consumers depending on their past purchases.

the standard literature (Berry, 1994; Berry, Levinsohn and Pakes (BLP), 1995; and Nevo,2000a, 2001). I use a three-dimensional panel of quantities and prices for up to 13 brands of plain yogurt, purchased from the 6 largest supermarket chains in up to 94 departments of France, weekly in 2006. I also observe demographics and other household characteristics. Household membership to loyalty programs is given by a household-supermarket specific indicator variable taking on 1 if the household is a member of a given supermarket LP. In addition to the well documented challenges faced when estimating demand, I deal with the correlation between membership to LPs indicator and the unobserved supermarket attributes. In a first stage of the estimation, I try some standard sets of instruments (BLP, 1995 and Nevo, 2001) to treat the endogeneity of prices and LP membership. Yet, due to a poor performance of those sets I circumvent endogeneity issues by computing optimal instruments based on Chamberlain (1987), using empirical methods developed by BLP (1999) and Reynaert and Verboven (2014).

Results confirm that private labels are, on average, less valued products relative to national brands. However, I find that the marginal valuation of PL products increases with subscription to the supermarket LP, which supports the believe that LPs might be used as a way to boost store-brands demand. Moreover, when customers hold simultaneously subscriptions to LPs of competing retailers, the expected effects become weaker, i.e. the marginal valuation of PLs decreases with the number of subscriptions and customers are more sensitive to price changes. Demand-side counterfactuals suggest that loyal customers benefit from when the market is not fully covered, as supermarkets might make efforts to improve loyalty rewards so as to attract nonmembers. On the other hand, making LPs prohibitively costly may harm consumers.

Related literature

There is vast literature on topics related to private labels, brand- and store-loyalty and loyalty programs. However, to my knowledge this is the first paper to provide empirical evidence on the link between supermarket loyalty programs and private label demand. Table 1.1 presents some contributions by topic.[7]

This article relates to Lewis (2004), who makes an evaluation of the effects of loyalty programs based on the idea that they are addressed to "enhance [customer] retention". Bonfrer and Chintagunta (2004) study the effects of the introduction of PL on retailers' profits taking into account that consumers can be store- and brand-loyals. They propose measures for store- and brand-loyalty based on the number of trips to the same store and the average "share of wallet" of a brand relative to the total expenditure on that category of goods. They find a significant negative

[7] For a complete survey about the theoretical and empirical literature on this theme, see Berges-Sennou et al. (2009).

Table 1.1: Contributions to the literature on Loyalty programs and PLs

Topic	Theory	Empirics
Introduction of PLs	Raju et al. (1995)	Bonfrer & Chintagunta (2004)
	Chintagunta et al. (2002)	
PLs demand determinants	Berges et al. (2009)	
Competition & vertical	Soberman & Parker (2006)	Bonnet & Dubois (2010)
relationships	Bonnet & Dubois (2010)	
Store & brand loyalty	Berges (2006)	
Loyalty programs	Lal & Bell (2003)	Bolton et al. (2000)
		Lal & Bell (2003)
		Lewis (2004)
		Lederman (2007)
Coupons	Caminal & Matutes (1989)	Nevo & Wolfram (2002)
	Cremer(1989)	

correlation between both types of loyalty. Bolton, Kannan and Bramlett (2000) argue that loyalty programs members are more likely to do repeat buying, as they weigh less than others the best outside alternative. They conclude that LPs members are less sensitive to both quality changes and lower prices offered by competitors. Lal and Bell (2003) claim that there are two reasons explaining the "success" of loyalty programs: (*i*) reduced price competition and therefore higher profits due to switching costs, and (*ii*) reduced marketing expenses by focusing attention on retaining loyal (and known) customers. This is in line with the marketing idea that promotions should be addressed to customers that are more likely to stay. All these papers share a common basic question: What are the determinants of customer retention and repeat buying? This article asks a rather different question, as the interest is focused on the effects of loyalty programs on private label demand.

In particular, this paper closely relates to Nevo and Wolfram (2002). They provide empirical support on coupons issuing strategies by manufactures. Their objective is to describe manufacturers' motivations for issuing coupons. They evaluate some hypotheses such as price discrimination, dynamic demand effects and retailers' pricing strategies using data on breakfast cereal. The key difference with my article is that I focus on retailers' rather than manufactures' case with the particularity that the former give customers personalized "checks" that can be expended in any set of products in stock in the supermarket.

The remainder of this paper is structured as follows. Section 3.2 outlines the data and a preliminary analysis. Section 1.3 describes the empirical framework, the estimation procedure and the identification strategy. Section 2.1 presents and discusses the results. Section 1.5 presents the results of two simulated experiments on how demand responds to changes in LP membership. Finally, Section 3.6 concludes and discusses directions for further research.

1.2 Preliminary analysis: customer profile and the grocery retailing industry

This Section aims at giving an overview of the supermarket industry in France and the customer profiles, through an exploratory analysis of the data. As loyalty programs are a supermarket-level rather than a product-level marketing strategy, I analyze a wide range of products (about 350 different food products) purchased by the households in the sample during 2006 to provide some descriptive evidence on customers behavior in the presence of loyalty programs offered by competing supermarkets. In Section 4 I will focus on a single product to assess the effects of LPs on product choice.

1.2.1 The data

This study uses the TNS Worldpanel data base provided by the TNS-Sofres Institute. It is homescan data on grocery purchases made by 14,529 households in France during 2006. Household members collect the data with the help of scanning devices provided by TNS. The sample is representative of the french population. It was originally randomly selected in 1998 and keeps most households over the years. Those households rarely reporting data are dropped from the original sample and replaced by new randomly selected participants. Furthermore, increases the sample size every year.

The database contains information on 352 different grocery products from around 90 retailers including hyper- and supermarkets, convenience stores, hard-discounters and specialized stores. The data is reported at the purchase level, so we observe product characteristics such as total quantity, total expenditure, the retailer where it was purchased from, brand, etc. In addition, the data include household characteristics such as household size, number of children, location, income, number of cars, internet access, storage capacity etc. The 2006 database contains as well information on household membership to retailers' LPs. It is an indicator variable taking on 1 if the household is member of the retailer's LP and zero otherwise. Unfortunately, no further information such as loyalty coupons issuing or redemption rates is available.

1.2.2 Customer profile

Table 1.2 displays summary statistics on household demographics, purchases, and information on store loyalty. The survey includes people aged between 19 and 75 years old. On average, a household consists of two to three members and has an income of around 2300 € per month. 75% of the households included in the data set lives in urban areas of France. 34% of products purchased by a household are PLs, which corresponds to 24.75% of its total expenditure. Moreover, the average French household members are one-stop shoppers, as they only visit one store a week. Finally, 85% of households are members to at least one supermarket loyalty program and, on average, they are subscribed to two separate programs.

Table 1.2: Summary statistics on household characteristics and purchases

Variable	Mean	Median	Sd	Min	Max
Demographics					
Size of household	2.63	2.00	1.39	1	9
Live in city	0.75	1	0.43	0	1
Income (€/month)	2337	2100	1175	150	7000
No. of cars	1.47	1.00	0.82	0.00	9.00
Purchases					
Private label purchases	0.34	0.32	0.17	0	1
Total expenditure (€ /day)	39.10	28.56	35.40	0.01	2,221
PL share (% total exp.)	24.75	15.40	38.6	0	63.63
NB share (% total exp.)	75.24	72.49	83.17	0	100
Store-related information					
No. different stores visited the same day	1.13	1.00	0.38	1.00	7.00
Duration (days) between visits	8.10	6.63	6.17	1	126
LP membership	0.85	1.00	0.36	0	1
No. of different memberships	2.60	2	1.48	1	12

In a month households visit, on average, two different retailers and only 16.65% of the times they go to the store owning the LP to which the household is a member. Using this information along with the average duration of 1.89 months a household takes to return to the same retailer, we get seven weeks as the average duration a LP member takes to go back to its "patronized" supermarket, which is not so frequent taking into account that household members go shopping at least once a week on average (see Table 1.3).

Table 1.3: Summary statistics of monthly visits to stores

Variable	Mean	Median	Sd	Min	Max
# different stores visited	2.27	2.08	0.98	1	9.25
# visited if loyalty subscription	0.36	0	0.50	0	3
Loyalty ratio*	16.83	0	25.40	0	100
Duration (months) between visits	1.89	1.33	1.48	1	12

*Computed as the number of visits to stores where customer is subscribed over the total of stores visited.

"Loyal" vs. "non loyal" customers

A simple exploratory analysis of the data by subgroup of population gives no evidence to support the hypothesis of loyalty programs as a discriminatory device. In effect, some descriptive statistics (see Table 1.4) suggest there are no important differences between subscribers to a LP relative to non subscribers.

One can say that the fact of being a member of a LP does not say much about customer types as long as LPs are available to everybody and subscription costs might be negligible.[8] Table 1.4 shows that these two groups not only have similar characteristics, but also similar consumption patterns. In fact, the portion of PL products purchased is similar on average (34%) as well as the total expenditure per trip to the supermarket, the allocation of this expenditure between PLs and NBs, and the average duration in days between trips to a supermarket (around eight days).

Table 1.4: Loyals and non-loyals average characteristics

Variable	**LP members**	**Non-members**
% on total	85	15
Size of household	2.63	2.62
Income	2333	2361
# of cars	1.47	1.48
Private label purchases	0.34	0.35
Total expenditure (€/day)	39.09	39.19
PL share (% total exp.)	25.87	26.36
NB share (% total exp.)	74.13	73.64
# different stores visited the same day	1.12	1.12
Duration (in days) between visits	8.04	8.41

I regressed consumer weekly expenditure per retailer on the LP membership dummy and household and supermarket characteristics. Table 2.4 displays the results. The coefficient for membership is positive and significant indicating that a customer tends to spend a larger share of his income in those supermarkets where he is a LP member than in those where he is not. Estimates for the number of LP subscriptions and number of stores visited are both significant and negative, which indicates that multi-homing as well as multi-stop shopping affect negatively the total expenditure on each supermarket.

[8] Although it is argued that subscription to a LP is free and thus subscription should be the rational behavior, loyalty programs can actually be costly as members must invest time an effort to subscribe, get a clear idea about how the program works and how discounts are given to them, in addition to give detailed personal information, bear larger amounts of advertising (even personalized), e-mail spam, etc. It is actually an empirical fact that not all consumers subscribe to LPs: 15% of French households who frequent supermarkets do not have any LP card whatsoever.

Table 1.5: Results for HH weekly expenditure per supermarket

	Log of total expenditure per store	
Variable	OLS	Fixed effects
LP member	0.372	0.499
	(0.003)	(0.004)
# of LP subscriptions	-0.0477	-0.0578
	(0.001)	(0.001)
# visits to	-0.123	-0.112
different stores	(0.001)	(0.001)
Log of Income	0.202	0.205
	(0.003)	(0.003)
Log of age	0.109	0.104
	(0.004)	(0.004)
hh size	0.155	0.151
	(0.001)	(0.001)
Hypermarket	0.138	
	(0.003)	
Minimarket	-0.252	
	(0.006)	
Hard-discount	0.0211	
	(0.004)	
Constant	1.099	1.151
	(0.027)	(0.027)
adj. R^2	0.145	0.179

Notes: Regressions are based on 658,866 observations. The two regressions include time dummy variables. Asymptotically robust s.e. are reported in parenthesis.

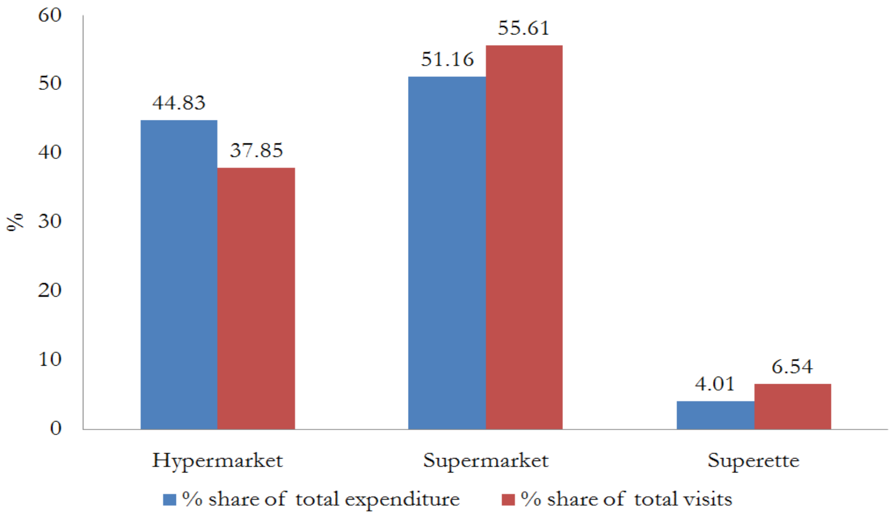
1.2.3 The grocery retailing industry

The preferred store format by french households is the supermarket: the average market share of supermarkets in terms of total consumer expenditure per day is 51.16% against 44.83% of hypermarkets and 4% of convenience stores.[9] In terms of daily number of visitors, supermarkets also appear leading: 55.61% of the total number of customers went to supermarkets whereas 37.85% went to hypermarkets and 6.57% to convenience stores (see Figure 1.1). As compared to regular stores,

[9] According to the French law, a grocery retailing store is considered a Hypermarket when, among other characteristics, its surface is greater or equal to $2,500\text{m}^2$, a Supermarket if the surface is in the interval $[400, 2500)$ m^2 and a convenience store if the surface is in $[120, 400)$ m^2. Hard-discount stores are also included in these three categories as they also have shops of all sizes.

Hard discounters have a share of 14% on total daily household expenditure in groceries and 16% share on total visits per day.

Figure 1.1: Daily market share indicators by store format, 2006



In terms of LPs subscribers, the largest market share is 21% suggesting that the market is not very concentrated.[10] Still, around 75% of LPs subscribers are concentrated by the five leading retailers (see Figure 1.2). Moreover, around 62% of the households have multiple LPs subscriptions (in some cases 8 in total), reason why it is necessary to look for other indicators such as the percentage of members visiting the store in a day. Finally, the proportion of "loyal" customers visiting the store owning the LP is 20.6%, a very low proportion considering the high concentration of the sector in both market shares and LP members.[11]

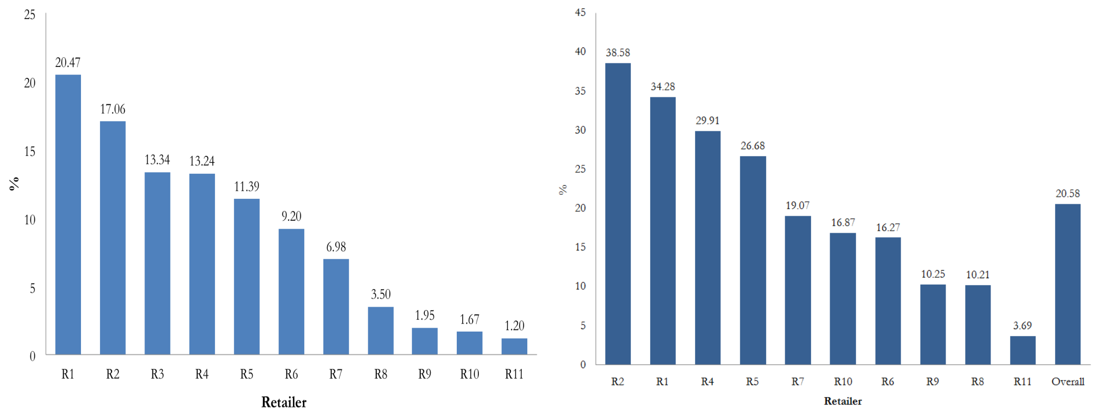
1.3 Empirical framework and estimation

In the empirical analysis I follow the discrete-choice literature and estimate two models: a multinomial Logit and a random-coefficients Logit (I use random-coefficients Logit, mixed Logit and the full model exchangeably). The reason for conducting these two estimations is that the Logit model is useful as a diagnostic tool as it is easy to estimate and gives important preliminary information about the explanatory power of the variables of interest. However, as it is well known,

[10] There are cases of retailers that have several cards that serve to the same end like, for example, Carrefour that has both "Carte fidelité" and "Carte Pass". In such cases, I aggregate up subscribers under the same label, e.g., Carrefour loyalty program.

[11] The actual names of the retailers are hidden at the request of the TNS-Sofres, the institution providing the data.

Figure 1.2: Share of subscribers to LPs by retailer (left) and Average percentage of 'loyal' visitors on the total visitors per day, 2006



this model has some limitations due to its restrictive assumptions, in particular, it gives unrealistic substitution patterns. The random coefficients model provides more reliable results.

1.3.1 The empirical framework

Assume we observe $t = 1, 2, \ldots, T$ markets and $i = 1, 2, \ldots, I_t$ consumers per market. I define a market as a week-Department[12] combination where consumer purchases are observed. Every time a consumer goes shopping to a given supermarket $s = 1, 2, ..., S$, he faces a multiple-choice decision among J brands. The conditional indirect utility of consumer i from choosing product j at supermarket s in market t writes as

$$u_{ijst} = x_j\beta_i + x_s\lambda - \alpha_i p_{jst} + \varphi_i M_{is} + \eta_i M_{is} \times PL_{js} + \xi_j + \xi_s + \Delta\xi_{jt} + \Delta\xi_{st} + \epsilon_{ijst} \quad (1.1)$$

where x_j and x_s are K- and R-dimensional (row) vectors of observable product j and supermarket s characteristics, respectively;[13] p_{jst} is the unit price of product j in supermarket s, M_{is} is a dummy taking on 1 if individual i is a member of supermarket s's loyalty program and zero otherwise, and PL_{js} is a dummy taking on 1 if j is a private label of retailer s. ξ_j and ξ_s are mean (across individuals and markets) valuations of the unobserved (by the econometrician) product and supermarket characteristics and $\Delta\xi_{jt} = \xi_{jt} - \xi_j$ and $\Delta\xi_{st} = \xi_{st} - \xi_s$ are market-specific deviations from the respective means under the assumption that in each market

[12] In France, a Department (or *Departement* in French) makes part of the administrative division of the national territory being the second level of the government at the local area, after Administrative Regions which group departments.

[13] Unlike product characteristics, I do not allow supermarket characteristics to interact with household characteristics, i.e. λ is a fixed coefficient.

people value differently those characteristics. Finally, individual heterogeneity enters the model through the set of $K+3$ individual-specific parameters $(\alpha_i, \beta_i, \varphi_i, \eta_i)$ and an additive separable mean-zero random shock ϵ_{ijst}.

Following Nevo (2000a, 2001), I model consumer taste parameters as a function of observed and unobserved household characteristics and assume that the latter are normally distributed

$$\begin{pmatrix} \alpha_i \\ \beta_i \\ \varphi_i \\ \eta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \varphi \\ \eta \end{pmatrix} + \Pi D_i + \Sigma v_i, \qquad v_i \sim N(0, I_{K+3})$$

where D_i is a $d \times 1$ vector of demographic and household characteristics, Π is a $(K+3) \times d$ matrix of coefficients measuring the change in tastes with household characteristics, Σ is a $(K+3) \times (K+3)$ scaling matrix and v_i captures all those unobserved demographic characteristics that influence consumer choice but that are generally not included in surveys.

Define the "outside good" as any other brand or type of yogurt or any other product that is an alternative to those considered in this analysis. It too accounts for the no purchase option. Normalizing the mean utility to zero, the indirect utility derived from the outside option writes as $u_{i0t} = \epsilon_{i0t}$.

Let $\theta = (\theta_1, \theta_2)$ be a vector containing all the parameters of the model ($\theta_1 = (\alpha, \beta, \varphi, \eta, \lambda)$ contains the linear parameters and $\theta_2 = (\Pi, \Sigma)$ the nonlinear ones). We can rewrite the indirect utility that consumer i derives from purchasing brand j at supermarket s in market t as the sum of two components: a mean utility common to all consumers of the same type (defined here by the subscript $m = \{0, 1\}$ of "loyals" and "nonloyals", respectively)

$$\delta_{mjst} = x_j\beta + x_s\lambda - \alpha p_{jst} + \varphi M_{ms} + \eta M_{ms} \times PL_{js} + \xi_j + \xi_s + \Delta\xi_{jt} + \Delta\xi_{st}, \quad (1.2)$$

and a mean-zero heteroscedastic deviation, $\mu_{ijst} + \epsilon_{ijst}$ with

$$\mu_{ijst} = [p_{jst}, x_j, x_s, M_{ms}, M_{ms} \times PL_{js}]' * (\Pi D_i + \Sigma v_i),$$

The utility can then be written as

$$u_{ijst} = \delta_{mjst}(\cdot; \theta_1) + \mu_{ijst}(\cdot; \theta_2) + \epsilon_{ijst}$$

A key assumption of this model is that consumers choose at most one unit of the brand that gives the highest utility. Let the following be the set of observed and unobserved variables determining the preference for brand j at store s

$$A_{mjst}(x, p_{.t}, \delta_{.t}; \theta_2) = \{(D_i, v_i, \epsilon_{ijst}) | u_{ijst} \geq u_{ilkt}, \ \forall\, l = 0, 1, \ldots, J; k = 0, 1, \ldots, S\}$$

where x are the characteristics of all products and supermarkets, and $p_{.t}$ and $\delta_{.t}$ are $J \times S$ matrices of prices and mean utilities of the J products available in the S supermarkets, respectively. Assuming ties occur with zero probability, the market share of the jth brand purchased from supermarket s in market t as a function of the mean utility levels of all the $J+1$ products, given the parameters, by group of population m is

$$s_{mjst}(x, p_{.t}, \delta_{.t}; \theta_2) = \int_{A_{mjst}} dP(D, v, \epsilon) = \int_{A_{msjt}} dP(\epsilon)dP(v)dP(D) \tag{1.3}$$

where $P(\cdot)$ denotes population distribution function. The last equality is a result of the independence assumption of D, v and ϵ. Market shares in (3.3.23) can be computed by making assumptions on the distribution of each of the individual variables $(D_i, v_i, \epsilon_{i.t})$.

1.3.2 The yogurt data

Among all the products in the data described in Section 3.2, yogurt seems to be suitable to estimate the effects of LPs subscription on the demand for PLs, as it matches pretty well the assumptions of the Logit setup. First, it is a product of regular consumption by French households. In fact, according to the National Statistics Institute of France (*Institut National de la Statistique et des Etudes Economiques* —INSEE), in 2006 yogurt consumption was, on average, 415.4g per week, which assuming a serving of 125g (the size most frequently consumed), is equivalent to 3.3 servings a week. Second, although it can be stored in the fridge for about a week, its regular consumption indicates that it does not last very long in customers' fridges, i.e. stockpiling is not a concern in practice. Finally, it can be considered a good of unit demand in the sense that individuals do not generally consume more than one serving at a time,[14] which is key for the empirical framework I use.

To avoid dimensionality problems due to the large availability of yogurt varieties, I use purchases of plain yogurt, which is the best consumed variety of yogurt in France with a 33% of the national market share in 2006.[15] I include flavored yogurts in the

[14] It is true that people can buy several varieties of yogurt in the same shopping trip in order to have multiple choices at home (different flavors, fruit contents, thickness, etc.). However, I claim that in general, an individual consumes one serving at a time, so that the choice is discrete in this sense. Of course there could be cases in which some people consume more than one brand of yogurt at a time. In such cases, the assumption should be seen as an approximation to the real demand problem.

[15] The original database of yogurt contains purchases of 174 varieties of yogurt sold by an average of 20 separate retailers in the 94 metropolitan departments of France. In addition, different flavors are branded under the same label, which increases the dimension of brand varieties. In the French market there are around 144 different flavors available, being 5 the average number of flavors by brand.

outside good. Second, I keep only the 6 leading grocery retailing chains with a loyalty program. Third, I select the 13 leading brands based on the national market shares on total sales in 2006. The aggregate market share of the selected brands is of 66.5%, six of them being PLs, which guarantees the representativeness of the subsample and still keeps a reasonable variation in the data. Summary statistics on price and market shares of the selected brands are displayed in Table 1.6. Moreover, I consider the supermarket as an additional brand characteristic, i.e. I define each brand variety as a supermarket-brand combination. This exercise results in more than 120 varieties, most with a very low market share. Based on this, I keep a final sample with the leading 31 varieties of brands with market shares varying between 0.9% to 5.8% and an aggregate share of 72.9%, purchased from the largest 6 supermarket chains in France in 2006.[16]

Table 1.6: Summary statistics for price and market shares of brands in sample

Variable	Mean	Median	sd	Min	Max
Total					
Price (€/125g)	0.229	0.187	0.091	0.151	0.463
Market share (% on total sales)	5.113	4.092	4.386	1.816	17.960
Private labels					
Price (€/125g)	0.174	0.170	0.019	0.151	0.203
Market share (% on total sales)	3.750	4.077	1.120	2.200	4.788
National brands					
Price (€/125g)	0.277	0.246	0.103	0.165	0.463
Market share (% on total sales)	6.282	4.092	5.828	1.816	17.960

1.3.3 Variables description

The data used for the estimation of the models previously described were aggregated to the brand level and contain information on total sales, unit price (per 125g serving), product and store characteristics and the distribution of the household characteristics. In particular, the following variables were constructed:

– *Brand market share* (S_{mjst}): It was computed per subgroup of population of LPs members ($m = 1$) and non LPs members ($m = 0$), as the percentage of 125g servings sold in a market (in this paper a Department-week combination) on the total potential number of 125g portions that could have been consumed

[16]The lack of randomness of the final sample considered in this paper does not lead to inconsistent estimates of the parameter as long as I include brand-supermarket dummy variables in the estimation. See Manski and Lerman (1977) and Bierlaire, Bolduc and McFadden (2008) for a detailed discussion on consistent estimation of choice probabilities from choice-based samples.

in that market.[17] The serving was determined by converting volume sales originally in kg into a 125g unit which is the size best sold in France. The potential volume sales per market was computed by multiplying the average national plain yogurt consumption of 1.14 125g servings per person per week in 2006, that I obtained from the database of yogurt purchases, and the total population in a department.[18]

- *Market share of the outside good* (S_{0t}): It was defined as the difference between one and the sum of the inside products market shares.

- *Price €/125g* (p_{jst}): It was generated by dividing the total expenditure on yogurt products over the total number of servings purchased.

- *LP membership* × *PL dummy* ($M_{ms} \times PL_{js}$): It is an interaction variable between the LP membership indicator and a dummy variable taking on the value 1 if the brand variety is a private label and zero otherwise. It aims at capturing the marginal effect of LP membership on private label demand.

- *Other interactions of interest*: the regressions include other interactions between household and product characteristics such as: *#Subscriptions* × *PL dummy*, which combines information on the number of separate LP cards held by a household and a dummy for private label, and *Price* × *#Subscriptions* which will be useful to see the marginal effects of a price change on loyalty programs' members.

1.3.4 Estimation

The estimation of the model described previously was conducted following the standard methods —Berry, Levinsohn and Pakes (1995), Nevo (2000a, 2001). I exploit the panel structure of my data to control for brand and supermarket fixed-effects. This implies guarantees that demand is identified without the need to put structure on the supply side.[19]

Estimation relies on the population moment conditions given by $E[h(z)'\rho(x, \theta_o)] = 0$, where $z_1, ..., z_M$ are a set of exogenous variables to be used as instruments; ρ is

[17] I computed the market shares by subgroup of people in order to preserve the meaning of the membership indicator in a context of aggregate data.

[18] The data on average consumption on food products from the National Accounts by *Institut National de la Statistique et des Etudes Economiques* (INSEE, Comptes nacionaux, base 2000) confirms this is a good measure of potential market. The reported average quantity consumption of 125g servings of yogurt per person per week was 3.32, and provided that around 34% of that number is plain yogurt, the average consumption of this variety per person per week is of 1.29 servings, which is similar to the one reported in my database.

[19] For a detailed discussion of the estimation algorithm and the differences with BLP(1995) procedure, see Nevo (2000a, 2001).

a function of the parameters of the model and θ_o is the vector containing the true value of the parameters (see subsection 1.3.5 for a discussion on the instruments used for identification). A generalized method of moments estimator is obtained by solving the problem

$$\min_{\theta} \rho(\theta)' h(z) \hat{\Lambda}^{-1} h(z)' \rho(\theta), \tag{1.4}$$

where $\hat{\Lambda}$ is a consistent estimator of $E[h(z)'\rho\rho' h(z)]$ and plays the role of the optimal weighting matrix in expression (3.4.1).

According to the empirical framework described before, once supermarket-brand dummy variables are included, the error term of the model is $\Delta\xi_{jt} + \Delta\xi_{st}$ which can be computed as a function of the mean utilities $\delta_{.t}$, the data and the parameters. Following Berry (1994), this computation requires solving first for $\delta_{.t}$ from the system of equations resulting from the match of predicted and observed market shares: $s_{.t}(x, M, p_{.t}, \delta_{.t}; \theta_2) = S_{.t}$, where $s_{.t}(\cdot)$ is the predicted market share function defined in (3.3.23). From the inversion of this system we can express $\delta_{.t}$ as an explicit function of the observed market shares which allows us to write the error term in (3.4.1) as

$$\rho_{jst} = \delta_{mjst}(x, M, p_{.t}, S_{.t}; \theta_2) - (x_j\beta + x_s\lambda - \alpha p_{jst} + \varphi M_{ms} + \eta PL_{js} \times M_{ms} + \xi_j + \xi_s)$$

To solve for $\delta_{.t}$, BLP (1995) proposed a contraction mapping which uses starting values for δ and θ_2 and iterates up until it converges to some value of δ determined by a stopping rule provided by the econometrician. For a detailed exposition of the the contraction mapping and estimation procedure, see BLP (1995), Nevo (2000a, 2001) and Knittel and Metaxoglou (2012) .

1.3.5 Optimal Instruments

As previously stated, the inclusion of brand-supermarket fixed-effects captures the unobserved brand-supermarket characteristics and then the error term of the model becomes $\rho_{jst}(\theta) = \Delta\xi_{jt} + \Delta\xi_{st}$, which are the group-market deviations from the mean valuation of product and supermarket unobserved characteristics, respectively. Under the assumption that both firms and customers observe those characteristics and, consequently, their decisions account for these local deviations, we have then two sources of correlation with explanatory variables.

Prices are correlated with the local deviation of the mean valuation of product unobserved characteristics, $\Delta\xi_{jt}$. According to Nevo (2000a), differentiated products pricing models assume that firms know the unobserved (by the econometrician) characteristics of the good and use them to set prices, which are a function of marginal cost and a markup depending on demographics. On the other hand, $M_{.s}$ appears to be correlated with the local deviation from the mean valuation of supermarket unobserved characteristics, $\Delta\xi_{st}$. Given that they are firm-level programs[20] and under the assumption that the choice of a LP is driven, among other things, by supermarket choice, i.e. it depends on consumer's valuation of supermarket characteristics, there is some information about households subscription decisions contained in the error term.

Properties for efficient estimation of models based on conditional moment restrictions using optimal instruments dates back to late 70's with main contributions by Amemiya (1977), Chamberlain (1987) and Newey (1990).[21] However, the lack of an algorithm for the computation of those instruments in empirical studies, led empirical IO literature to use inefficient rather than optimal sets of instruments to treat the endogeneity of prices. BLP (1999) were the first to use an approximated version of optimal instruments in the evaluation of export restraints on automobiles. Reynaert and Verboven (2014) propose an altenative method to empirically compute Chamberlain's instruments for random coefficient models.

Following Newey's (1990) notation, consider an econometric model with the following moment restriction[22]

$$E[\rho(x_i, \theta_o)|z_i] = 0 \tag{1.5}$$

where $\rho(x, \theta)$ is a $Q \times 1$ residual vector, z are instrumental variables, θ is a 2×1 vector of parameters, and θ_o stands for the true value of this set of parameters, and $x_1, ..., x_n$ are i.i.d. observations on the data vector x_i, z_i making part of its components. Assuming homoscedasticity, the variance-covariance matrix conditional on the instrumental variables is $E[\rho(x, \theta_o)\rho(x, \theta_o)'|z_i] = \Omega$, with Ω constant. Using appropriate functions of the data (and the parameters), denoted $h(z_i)$, we can express the system in (1.5) as unconditional moment restrictions

$$E[h(z_i)\rho(x_i, \theta_o)] = 0$$

[20] A customer joining a program benefits from rewards no matter what store of the chain he purchases from.

[21] Instrumental variables estimation for problems with conditional moment restrictions and i.i.d observations were introduced, among others, by Kalejian (1971), Amemiya (1974, 1977), and Jorgenson and Laffont (1974). Amemiya (1977) proposed the computation of optimal instruments. This developments assumed parametric forms for the error terms. It was Chamberlain (1987) who studied the asymptotic properties of the IV estimator for nonparametric models, where all that we know is that the distribution function of the data satisfies the equality of the expected value of the residual to zero when multiplied by appropriate functions of the exogenous variables, and that efficiency bounds are attained when these functions are replaced by the optimal instruments. Finally, Newey (1990) proposed nonparametric estimation methods of optimal instruments for nonlinear simultaneous equations models.

[22] For the sake of exposition, in the general formulation of optimal instruments I replace panel subscripts by a single subscript i indicating a particular observation of the data. I will go back to the usual notation when I derive the particular optimal IVs for the model in this paper.

The optimal instruments are a particular function $h(\cdot)$ that allows us to consistently estimate the parameters of the model and attain the efficiency bound of the asymptotic variance-covariance matrix. This is[23]

$$h(z_i) = D(z)'\Omega^{-1} \qquad \text{where} \qquad D(z) = E\left[\frac{\partial \rho_{.t}(x,\theta_o)}{\partial \theta}\bigg| z_i\right] \tag{1.6}$$

According to Newey (1990), for the particular case of a single equation system where the residual is a scalar (as is the case of the demand model described in subsection 3.4.2) the optimal instruments become $h(z) = D(z)'$.[24] The number of instruments is equal to the number of parameters to be estimated in the model $\theta = (\theta_1, \theta_2)$.

The optimal instruments for the linear parameters are easy to compute as they are functions of observed data

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \beta'}\bigg| z_t\right] = E[x_j|z_t] = x_j \tag{1.7}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \alpha}\bigg| z_t\right] = E[p_{jst}|z_t] = x_j\gamma_1 + x_s\gamma_2 + w_{jst}\gamma_3 \tag{1.8}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \varphi}\bigg| z_t\right] = E[M_{ms}|z_t] = x_s\tau_1 + l_s\tau_2 \tag{1.9}$$

$$\begin{aligned} E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \eta}\bigg| z_t\right] &= E[M_{ms} \times PL_{js}|z_t] \\ &= E[M_{ms}|z_t] \times PL_{js} = (x_s\tau_1 + l_s\tau_2) \times PL_{js} \end{aligned} \tag{1.10}$$

The instruments resulting for the identification of β are product observed characteristics which are assumed to be exogenously set by producers and supermarkets, i.e. those attributes do not vary in response to department-specific demand shocks. As for the parameters of the endogenous variables α, φ and η, the instruments are the corresponding predicted values from first-stage OLS regressions for $p_{.t}$ and $M_{.s}$.

To compute, first, the predicted price in (1.8), I follow Reynaert and Verboven (2014) and assume for simplicity that marginal costs are linear and depend on

[23] For a complete discussion about IV estimation methods of nonlinear models, optimal instruments and efficiency bounds for nonlinear models and other models of interest, see Newey (1990).

[24] The error term denoted $\rho(x, \theta_o)$ corresponds to a general formulation of a model with simultaneous equations. As the model to be estimated here consists of a single equation, the $\rho(\cdot)$ function equals the error term of the demand model $\Delta\xi_{jt} + \Delta\xi_{st}$. For BLP models, for instance, where demand and supply equations are estimated simultaneously, the residual is a vector containing both the demand- and the supply-side error terms, $\rho = (\xi, \omega)'$.

product and store characteristics, and cost shifters $w_{.t}$, and that markets are competitive so that firms set prices at marginal cost.[25] However, as I do not observe any cost shifters in my data set, I use average regional prices of the same product in all the 21 French administrative regions (excluding the department to be instrumented for from the average price of the region it is located in) as proxies for marginal costs information. Following Nevo (2001), I claim that after controlling for brand-specific means, regional-specific valuations are independent of product valuations of people from other regions. This implies that in case a demand shock happens in one region, only the local price will be affected. This guarantees the exogeneity condition of prices. Now, prices of two departments in a country are linked by common marginal costs as long as they are produced (supplied) by the same manufacturer (retailer) or under a standardized process.[26] Table 1.B.1 in the Appendix displays first-stage results of price regressed on average regional prices.

As for the predicted value of the subscription decision, it is assumed to be a linear function of supermarket and loyalty program characteristics.[27] Therefore, the solution is more challenging than that of prices for the following reasons. First, in the data I use I do not observe neither household consumption patterns before subscription nor any other additional information on the intensity of purchases motivated by loyalty rewards, the effective amount of rewards obtained by households or the rate of coupon redemption. Second, I do not address the causality relationship between supermarket choice (and PL consumption) and subscription decisions to LPs.[28].

To deal with this issues, I collected information on the characteristics of each loyalty program and use them to have a very rough approximation to the predicted subscription decision indicator $\hat{M}_{.s}$. As in previous literature that uses product characteristics as valid instruments for price (see Berry, 1994 and BLP, 1995), I assume that as long as supermarket LPs are the same over the whole country, most LP characteristics are not set in response to market specific deviations of the mean valuation of the program. Hence, they can be used as exogenous instruments for

[25] Reynaert and Verboven (2014) examine both perfect and imperfect competition cases and obtain similar results.

[26] Although the independence assumption seems reasonable, there may be cases where it might not hold as, for example, a national demand shock as pointed out by Nevo (2001). However, in France prices are set at the local market level and due to some regulations that introduce rigidity (such as Galland Act), local prices might not be very responsive to national shocks.

[27] Demographics are definitely key factors determining the decision to join a given LP. However, in order to have valid instruments I exclude any variable correlated with the error term of the model.

[28] The causality issue implies answering the question: do consumers source more frequently a particular supermarket after they have joined its LP or do they subscribe to this LP because they are already loyal to that supermarket? Answering this question implies a structural model of store choice in the presence of loyalty programs, which will require to disentangle its determinants and a detailed data set on LPs characteristics, use and rewards. This is out of the scope of this paper and I leave it for future research.

a first-stage estimation of the model. Table 1.B.2 in Appendix displays estimation results for the regression of M_{is} on collected LP characteristics and demographics. Although the information collected is not enough to fully account for the variation in M_{is}, the results are quite appealing showing that the exogenous structure of the programs have some explanatory power.

Finally, the instruments for the nonlinear parameters in (1.11) are more difficult to compute as they are nonlinear functions of product characteristics and the expectation in (1.11) is a function of the true parameters of the demand function, namely, both linear θ_1 and nonlinear θ_2, this is

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \theta_2'}\bigg| z_t\right] = E\left[\frac{\partial \delta_{mjst}(s_{.t},\theta_2)}{\partial \theta_2'}\bigg| z_t\right] \tag{1.11}$$

This means that the instruments for the nonlinear parameters of the model cannot be directly computed from the data and require a first-stage estimation of the model. Following BLP (1999) I approximate the population moment in (1.11) using the Jacobian for the delta function, which yields[29]

$$\left.\frac{\partial \delta_{mjst}(\hat{s}_{.t},\theta_2)}{\partial \theta_2'}\right|_{\theta_2=\hat{\theta}_2}$$

1.4 Results

1.4.1 Logit results

Table 1.7 displays the results for the estimation of a Logit demand model for yogurt by regressing the log of the observed market share of each product minus the log of the share of the outside option $ln(S_{mjst}) - ln(S_{0t})$ on the variables described in subsection 4.3 as main covariates. Additionally, depending on the estimation method I control for product characteristics (columns (1) and (4)), brand fixed effects (all columns but (1) and (4)), and household characteristics (columns (3), (6) and (7)). Columns (4)-(7) in Table 1.7 display the results of 2SLS regressions using instrumental variables to treat the endogeneity of prices. As the purpose of this subsection is purely descriptive, I deal just with the endogeneity of prices using sets of inefficient instruments: in column (4) I use brand dummy variables, which play a

[29] Reynaert and Verboven (2014) provide an alternative version of BLP (1999) approximation to optimal instruments and compare the performance of the two methods using Monte Carlo simulations. They conclude that the gains in efficiency are small when using their method whereas the computational burden increases. For this reason, I stick to the BLP (1999) approximate version of the optimal instruments because its computation is easier.

similar role as BLP (1995) instruments, and in columns (5) through (7) I use the set of average regional prices as in Hausman (1996) and Nevo (2000a, 2001). Column (4) includes brand characteristics, whereas columns (5) through (7) include brand fixed-effects, and that is why it is no longer possible to use brand dummies as instruments. I use characteristics of each supermarket LP as controls in all regressions.

For every IV regression I conducted a Hausman test of over identifying restrictions, which always rejects the null hypothesis of the exogeneity of instruments, even in column (7) where demographics were replaced by department dummies which are supposed to capture local demand shocks in a better way. An explanation could be that given that the test is distributed as a chi-square, the large number of observations will cause any model to be rejected (Nevo, 2001). However, the IVs are individually and jointly significant at 1% level and the high first-stage R-squareds and F-statistics suggest that they have some power (Nevo, 2001).

A result of special interest is the estimate for the interaction between the membership and the private label indicators ($LPmember \times PLdummy$). In all regressions, the coefficient is positive and significant, and the magnitudes do not vary importantly when adding demographic controls, further interaction variables or when prices are instrumented. The coefficient means that the marginal valuation of PLs increases with the subscription to the program of the supermarket owning the brand. This finding provides evidence for the hypothesis that supermarket chains use LPs as a way to increase demand for PLs.

Additionally, regressions (3) and (6) include the interaction of price with the number of separate cards ($Price \times \#Subscriptions$), and the number of cards with a private label dummy ($\#Subscriptions \times PLdummy$). The former has a negative and significant coefficient, which will be positive in the full model with optimal instruments, suggesting there is a downward bias in the estimate maybe coming from the use of inefficient instruments for price. On the other hand, the latter interaction has a negative coefficient indicating that the marginal valuation of PL products decreases with the number of LPs subscriptions. This means that multisubscription weakens the effects of LPs on the demand for PLs.

1.4.2 Results from the Mixed Logit model

Table 1.8 displays the results of the full model with "optimal" instruments. A first-stage version of the random coefficients model with the same specification was estimated with inefficient instruments for price (average regional prices) and loyalty membership (LP characteristics) in order to be able to compute the set of 20 "optimal" instruments (see Table 1.B.3 in Appendix).

The first column contains the estimated means of the distributions of the individual marginal utilities. They are all significant and most preserve the same

Table 1.7: Results from a Logit for the demand for palin yogurt [a]

	OLS			IV			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Price(€/125g)	-4.345	-6.216	-4.502	-11.18	-7.335	-5.965	-6.386
	(0.094)	(0.114)	(0.159)	(1.181)	(0.280)	(0.463)	(0.196)
LP membership	-0.298	-0.282	-0.252	-0.438	-0.300	-0.281	-0.268
	(0.018)	(0.016)	(0.015)	(0.032)	(0.017)	(0.018)	(0.011)
PL dummy	-0.405	-0.140	-0.504	-1.149	-0.460	-0.637	-2.049
	(0.025)	(0.052)	(0.062)	(0.131)	(0.088)	(0.074)	(0.066)
LP membership×PL dummy	0.255	0.236	0.210	0.366	0.249	0.235	0.151
	(0.025)	(0.023)	(0.022)	(0.034)	(0.023)	(0.023)	(0.015)
Plastic	-0.451			-2.024			
	(0.034)			(0.273)			
Sugar	-0.125			-0.402			
	(0.020)			(0.051)			
Wholemilk	0.133			0.155			
	(0.011)			(0.012)			
#Subscriptions×PL dummy			-0.0561			-0.0974	
			(0.015)			(0.019)	
Price×#Subscriptions			-0.696			-0.396	
			(0.056)			(0.102)	
Average hh size			0.181			0.167	
			(0.009)			(0.010)	
Log income			-1.055			-1.050	
			(0.023)			(0.023)	
Car			1.306			1.287	
			(0.064)			(0.064)	
# stores visited the			0.110			0.105	
same week			(0.038)			(0.038)	
# of trips to the same			0.0357			0.0412	
store within a month			(0.010)			(0.010)	
Constant	-12.28	-35.37	-17.45	-5.888	-34.78	-16.91	10.54
	(0.284)	(0.560)	(0.680)	(1.128)	(0.576)	(0.698)	(0.801)
Fit/Test of over	0.142	0.235	0.297	5,337	138.2	131.4	79.9
Identification[b]				(1.145)	(10.851)	(10.851)	(10.851)
1st Stage R^2				0.749	0.841	0.909	0.843
1st Stage F-test				1,676	1,896	3,351	1,019
Instruments				Brand dummies	Prices	Prices	Prices

[a]Dependent variable $ln(S_{mjst}) - ln(S_{0t})$. Based on 37,662 observations. All parameters are significant at 5% level. All regressions include week dummy variables and with the exception of columns (1) and (4) all regressions include brand dummy variables. The regression in (7) includes department dummy variables. Asymptotically robust s.e. are reported in parentheses. All regressions include characteristics of each supermarkets loyalty program as controls.

[b] Adjusted R^2 for the OLS regressions, and a Hausman test of over identification for the IV regressions with the 0.95 critical values in parentheses.

sign as in the descriptive Logit model. The interaction between LP membership and the PL dummy is positive, supporting the previous result of a positive impact of loyalty programs on private label demand. Provided that loyalty rewards are mainly addressed to store brands, customers are more willing to consume them as members of a LP as compared to non members. By contrast, the coefficients for the variables in levels, namely PL dummy and LP membership dummy, are both negative, which reflects the fact that people in general value less store branded products with respect to NBs and that being member of a loyalty program implies some costs, respectively.

The second column displays the estimated standard deviations of the mean coefficients referred above. These estimates are the coefficients of the interactions of the right-hand side variables of the model with the unobserved demographics. All but two are significant (those for *Plastic* and *Sugar*), meaning that the unobserved demographics v have some explanatory power for the heterogeneity in consumer tastes. As for *Plastic* and *Sugar*, the non significance of the estimates mean that included household characteristics are enough to explain the variation in consumer tastes.

Most included household characteristics have significant estimates. One interesting result is the negative coefficient of the interaction between PL dummy and income, confirming the intuition that higher income households value less PL products relative to lower income ones. The interactions with the total number of loyalty cards held by households (*#Subscriptions*) provide interesting evidence as well. They are all negative, meaning in the first case (interaction with the constant) that the more memberships a household holds, the higher the costs it faces, this is consistent with the estimate of the indicator for LP membership. In the second case, the interaction with price shows that multi-subscribers are more price sensitive as a marginal increase in price would have a larger impact for those holding more cards. The intuition behind this result is that the retention effect of a LP is weaken by the fact that a consumer holding several cards have lower costs of switching supermarkets. Last, the coefficient of the interaction of the number of subscriptions and the PL dummy is negative, which indicates that the valuation of private labels decreases with the number of memberships to different LPs, mitigating the positive impact that LPs have on the demand for PLs.

1.5 Demand response to changes in LP membership

In this Section I show the results of two counterfactual experiments based on the previous estimates. Under the assumptions that both prices and unobserved product characteristics do not respond to membership decisions by consumers (at least in the short run), and that the utility from the outside good remains the same, I consider

Table 1.8: Results from the Mixed Logit model[a]

Variable	Means (β's)	Std. Deviations (σ's)	Interactions with hh characteristics: hh size	Income	# Subscriptions
Constant	-10.067[b]	0.478		0.131	-0.968
	(0.043)	(0.382)		(0.236)	(0.346)
Price	-12.963	1.057	0.505	0.900	-0.535
	(0.453)	(0.485)	(0.132)	(0.085)	(0.255)
LP member (M_{ms})	-0.825	0.685		0.182	
	(0.207)	(0.286)		(0.536)	
PL dummy	-11.063[b]	0.916	-0.862	-0.446	-1.243
	(0.025)	(0.192)	(0.357)	(0.229)	(0.255)
LP member×PL dummy	0.584	0.765			
	(0.211)	(0.245)			
Plastic	10.877[b]	0.144			
	(0.060)	(0.406)			
Sugar	6.840[b]	0.171			
	(0.053)	(0.409)			
Wholemilk	-1.513[b]	0.466			
	(0.013)	(0.156)			
GMM Objective	1.69E-06				
MD χ^2	9707686251				

[a] Based on 37,662 observations. Except where noted, parameters were estimated using GMM. All regressions include brand and week dummies. Asymptotically robust s.e. are given in parentheses.
[b] Estimated using a minimum-distance procedure.

two scenarios.[30] First, all the households in the sample are members to at least one loyalty program, i.e. I set M_{ms} in equation (3.3.22) to one for all $m = 0, 1$, and replace the variable $\#Subscriptions$, defined in subsection 1.3.3, by one for those non LP members in the baseline case. Intuitively, this situation may arise when either the reward system of a LP is good enough so that those customers not having strong preferences for such programs are better off by joining or when real and perceived costs of subscription are zero (or even negative).

In a second scenario, I assume no household is member of a loyalty program whatsoever, i.e. I set M_{ms} in equation (3.3.22) to zero for all $m = 0, 1$. Such a case may arise when consumers perceive either subscription as prohibitively

[30] Due to the data limitations referred previously, this is the best I can do to exploit my results. The model might as well be exploited to recover retailers' price-cost margins according to some assumptions on the conduct of the industry, as in previous literature. However, the nature of the grocery retailing sector in which supermarkets are not just distributors but also rivals of upstream firms, and the implied both vertical differentiation between PLs and NBs and horizontal differentiation between PLs, adds new dimensions to a supply model where firms' decisions on loyalty rebates should be accounted for. A structural model addressing all these features is needed. This is out of the scope of this paper and is left for future research.

costly (because supermarkets set too high a membership fee[31]) or rewards are not worthwhile or just unattainable.[32]

Table 1.9 displays the predicted annual aggregate demand in the baseline and the two counterfactual scenarios. It also shows the percentage change of the demand with respect to the baseline when modifying the membership variable. Interestingly, the model predicts that when everybody is member of at least one loyalty program ($M_{ms} = 1 \, \forall \, m = \{0, 1\}$) aggregate demand of former non-loyals increases in 29.7% while it slightly drops for formerly loyal subscribers.[33] Overall, aggregate demand increases by 1.5%, a low effect which is dominated by the larger subpopulation of 'loyals'. The intuition behind this negative reaction is that after having covered the market for loyalty program subscriptions, supermarkets' efforts to make attractive programs may be reduced which affects all those customers sensitive to discounts and loyalty rewards.[34]

In the second scenario ($M_{ms} = 0 \, \forall \, m = \{0, 1\}$) overall aggregate demand decreases with respect to that in the baseline. As expected, demand by baseline loyals decreases in 17.3%. Surprisingly, baseline non-loyals also reduce their total demand. One explanation for this might be that although they are not so concerned about discounts and loyalty rewards, there is a psychological effect stemming from the fact that when supermarkets are active in promotional activities, all customers may directly or indirectly benefit, even if they are not part of the loyalty program. Therefore, when loyalty programs are not there they may find it less interesting to keep their demand level.

Table 1.10 presents the total annual demand by product category in million euros. Interestingly, the model predicts that when everybody is member of at least one loyalty program (first scenario) aggregate demand for private labels increases in 75.8% for people who were non members before the change. The demand for national brands increases also by 1.4%, which may be interpreted as an indirect effect of loyalty programs. This result supports the initial hypothesis that LPs may be used

[31] In some retailing sectors in France such as clothing or department stores, loyalty programs giving permanent and special rebates, exclusive offers, and other benefits are offered to customers for a subscription fee either once or yearly of up to 30 €.

[32] Some loyalty programs require customers to pay a fraction in cash of the full price of the reward, such as FPP. Airlines ask FPP members to do a large number of "qualifying" (generally international) flights or to accumulate a given number of miles (which is often high) in a short period of time so as to reach a higher status and enjoy extra benefits. Some set short deadlines to expend the accumulated miles.

[33] To obtain the annual aggregate demand I compute the per brand-market (week-Department) aggregate demand, q_{jst}, as the predicted market share of each brand in a market s_{jst}, times the size of that local market (total number of consumers in a Department), $\mathcal{M}_t$. Then, I sum up local per-brand demands by brand across markets to get the yearly aggregate demand per brand, q_{js}. Finally, I aggregate across brands to obtain the total annual demand for yogurt.

[34] Recall that I am assuming that prices and supermarket-brand unobserved characteristics remain unchanged in the counterfactual scenarios.

Table 1.9: Aggregate demand under three schenarios (in million euros and % change)

Population subgroup	In million euros			% change	
	Baseline	$M_{ms}=1$	$M_{ms}=0$	$M_{ms}=1$	$M_{ms}=0$
Non-loyals	8.76	11.36	7.78	29.70	-11.17
Loyals	76.30	74.97	63.10	-1.74	-17.30
Total	85.06	86.34	70.88	1.50	-16.67

Notes: Column headers indicate counterfactual scenarios: everybody is member of LP, $M_{ms} = 1 \forall m = \{0,1\}$, and nobody is, $M_{ms} = 0 \forall m = \{0,1\}$. Row labels stand for the two original population subsamples according to their membership status.

as a marketing strategy to boost the demand for PLs, apart from other objectives. As for 'Loyals', i.e. those who were already LP members, demand seems to not be affected as it decreases by less than 1% with respect to the predicted demand in the baseline scenario. This is in line with what is expected as 'loyal' members may not have additional benefits but rather worse conditions as supermarket may reduce the benefits of LPs once the market is fully covered.

In the second scenario demand for both PLs and NBs decreases for the two groups. However, the impact is lower than that of the first scenario. Loyals' aggregate demand for PLs decreases by 27.2% whereas non-loyals' demand decreases by 18.1%. This negative reaction for both subgroups might be driven by psychological perceptions of indirect benefits from loyalty rewards as previously stated. Similarly, the demand for NBs decreases for both groups indicating that there might be an effect of loyalty programs on the whole product range: the idea that some yogurt brands purchases may give you rebates and other rewards, leads to a general increase in demand for all brands in the product category.

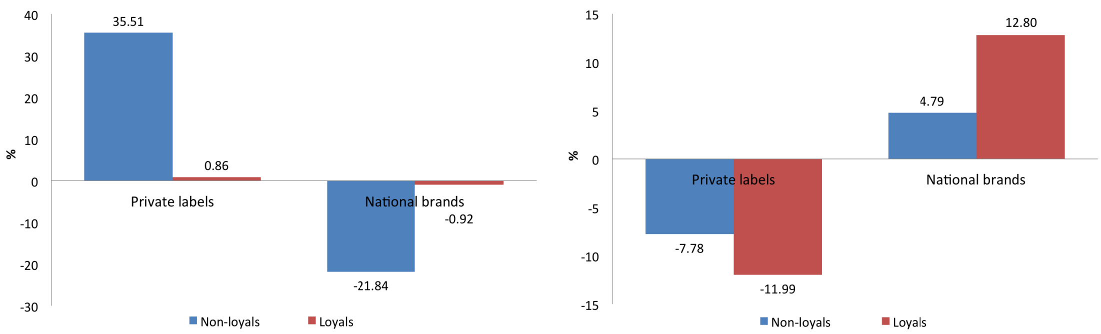
Figure 1.3 helps understanding what is behind people's reactions to changes in the membership status. It shows the percentage change of the shares of demand for PLs and NBs on the total demand with respect to the baseline values. Under the two scenarios, consumers seem to be substituting demand across product categories. On the left bar chart, a fully covered market for LP subscription makes PLs share to go up for both loyals and non-loyals. On the other hand, shares of demand for NBs decrease. Under the second scenario (bar chart on the right) we can see that impossibility to joint a LP leads to a decrease on the share of PLs on total demand, whereas NBs gain in importance for consumers. Notice that in the first case the impact is larger on the subgroup of non-loyals , as they are new to enjoy LPs benefits, and in the second case it is so on the subgroup of loyals, as they are more sensitive to promotional activities.

Table 1.10: Aggregate demand under three scenarios by brand type (in million euros)

Population subgroup	Private label			National brand		
	Baseline	$M_{ms}=1$	$M_{ms}=0$	Baseline	$M_{ms}=1$	$M_{ms}=0$
Non-loyals	3.34	5.86	2.73	5.43	5.50	5.05
Loyals	39.40	39.04	28.67	36.90	35.93	34.42
Total	42.73	44.91	31.41	42.33	41.43	39.47

Notes: Column headers indicate counterfactual scenarios: everybody is member of LP, $M_{ms}=1\,\forall\, m=\{0,1\}$, and nobody is, $M_{ms}=0\,\forall\, m=\{0,1\}$. Row labels stand for the two original population subsamples according to their membership status.

Figure 1.3: Percentage change of the share of demand by product type on total demand under scenario 1 (left) and scenario 2 (right)



I also compute the change in total consumer surplus (CS) for each scenario. Following Train (2009), the expected change in consumer surplus for individual i, provided that the price coefficient, $\tilde{\alpha}_i$, do not depend on income, i.e. the coefficient does not change when either income or price changes,[35] can be easily calculated as:

$$\Delta\mathbb{E}(CS_i) = \frac{1}{\alpha_i}\left[\ln\left(\sum_{s=1}^{S}\sum_{j=1}^{J}\exp(\delta^1_{mjst})\right) - \ln\left(\sum_{s=1}^{S}\sum_{j=1}^{J}\exp(\delta^0_{mjst})\right)\right], \qquad (1.12)$$

where δ_{msj} is defined by (1.2) and the superscripts 0 and 1 make reference to before and after the change in loyalty membership, respectively. The total mean change in consumer surplus is obtained by taking the average (1.12) over the whole sample and multiplying by the size of the national market which is the total population in France in 2006 (according to the official statistics by the INSEE, 61.1 million people) times 52 weeks (Nevo, 2000b).

[35] For a complete discussion on this, see Train (2009). For the computation of consumer surplus when the marginal utility of income varies with income, see McFadden (1999).

Table 1.11 displays the results of the change in total consumer surplus for each of the two simulated cases. In the first counterfactual scenario (first column of Table 1.11), where LPs market is fully covered, former non-loyals' CS increases by about 1.5% (36 million euros a year), whereas that of former loyals slightly decreases (0.6%). A similar trend is shown by the results of the second counterfactual scenario. The former non-members are not affected by the fact that now no one is willing to subscribe to a LP. However, former loyals are worse off with a reduction in total CS by 3.1% (49.4 million euros a year). Overall, customers are better off when everybody subscribes to at least one loyalty program, with a predicted increase of 13 million euros in CS, and worse off if nobody subscribes, with a predicted decrease of 26 million euros in total CS.

Table 1.11: Change in Consumer Surplus as a result of changes LPs market coverage (in millions of euros and percentage per year)

	$M_{ms} = 1$		$M_{ms} = 0$	
Group	millions of euros	%	millions of euros	%
Non-loyals	36.06	1.54	-3.48	-0.15
Loyals	-9.43	-0.59	-49.44	-3.11
Total	13.31	0.68	-26.46	-1.34

Notes: Column headers indicate counterfactual scenarios: everybody is member of LP,$M_{ms} = 1 \forall m = \{0, 1\}$, and nobody is, $M_{ms} = 0 \forall m = \{0, 1\}$. Row labels stand for the two original population subsamples according to their membership status.

1.6 Conclusions and further research

This paper studies the effects of supermarket loyalty programs on private label demand. Prior research has concluded, among other things, that loyalty programs are generally used to retain customers as it induces repeat buying, and as a discriminatory device. Evidence also suggests that retailers may be using LPs as a way to boost the demand for PLs. An empirical fact is that PLs are on average 20% of lower price relative to quality equivalent national brands. Yet, a common feature of supermarket loyalty programs is to reward private label purchases. This motivates the question I hoped to address in this paper: Why do profit-maximizing retailers give additional rebates on their lower-price own-brands?

This article adds to the literature of both private label and loyalty programs topics by using structural methods of demand estimation, and provides empirical support to a question that, to the best of my knowledge, has not yet been addressed.

Using a random coefficients Logit model to estimate a brand-level demand system for yogurt in France, I find that loyalty program members have a higher valuation for PLs as compared with non members, despite that my results are consistent with the general wisdom that store brands are less preferred products relative to similar national brands. When multi subscription is present, the effects of LPs are weaker: marginal valuation of private labels decreases with the number of subscriptions to different supermarket LPs and customers are more price sensitive.

By conducting some counterfactuals on the demand side, assuming exogenous changes in LPs subscription, I find evidence that unambiguously supports the hypothesis of this paper that LPs have a direct impact on the demand for PLs. In fact, aggregate demand for private labels of those people who formerly were not members of any LP considerably increases when they become members. On the other hand, in the absence of LPs (second scenario) private labels become less attractive products for everyone. Welfare analysis shows that consumers are in general better off when they all join at least one LP and worse off when no one is member to LPs, as compared with the baseline scenario. Results suggest that making subscription to LPs prohibitively costly may harm all consumers, even those non members in the baseline scenario.

Due to very limited data on loyalty programs, this paper cannot go further on the evaluation of LP effects on customer purchasing behavior. However, the results can be used as a motivation for pursuing further research on these topics in case richer data sets are available. In particular, with the help of descriptive results this paper sheds some light on questions that are worth asking. For instance, with the appropriate data modeling customer subscription decisions to supermarket LPs would help understanding why, if customers can join free of any charges and economic intuition indicates that each consumer able to subscribe should do so, not everybody subscribes or, among those who subscribe, just a very few join more than two programs. In fact, in France in 2006 15% of customers were not members of any supermarket LP and, among the members, 20% were affiliated to only one. Another interesting question to be explored is what are the effects of LPs on supermarket choice. This paper provides some descriptive results suggesting that members of a given supermarket loyalty program spend on average a higher portion of their income in that supermarket compared to non member customers. If supermarket and brand choice can be seen as an ordered choice problem in which store choice comes first, then LPs may play a key role in attracting consumers and making them one-stop shoppers.

Further research imply going deeper in the meaning of making customers loyal to a store or a particular brand and understanding the role played by loyalty programs in this context. By adopting a measure of loyalty, such as the share of wallet or the shopping intensity, it can be explored whether LPs are making customers loyal. A question that remains to be answered is thus: are non members more loyal than members?

Finally, from a policy perspective this paper may be seen as a first step for assessing more complex and interesting questions: Are loyalty programs being used by retailers as a strategy to increase buyer power? Moreover, taking into account the theory result that loyalty programs allow retailers to increase PLs prices, it is worth asking: Can loyalty programs be harmful for consumers at some point? Should policy authorities be cautious about the implementation of such marketing strategies? In order to answer such questions a structural model of the supply side is needed, which should include the horizontal (competition among retailers with their own brands) as well as the vertical dimension (competition among retailers and manufacturers, i.e, PLs against NBs). A counterfactual experiment on the supply side would allow to assess whether retailers are trying to exert some buyer power or not. I hope to think further about this in the future.

Appendix

1.A A theoretical background

This is a two-period model of oligopolistic competition with horizontal differentiation.[36] Consider two symmetric retailers ($j = 1, 2$) differentiated à la Hotelling and located at the two extremes of a line of length one. They supply the good at a constant marginal cost $c > 0$ and compete for a mass of customers indexed by $i \in [0, 1]$ uniformly distributed along the Hotelling line. I assume unit demands, i.e., each consumer buys at most one unit of the good in each period. Moreover, I assume the good is not storable. Consumer location x_t is independently and identically distributed over time. In particular, consumers cannot anticipate their second period location in period 1.[37]

[36] Caminal and Matutes (1990), analyse the effects of alternative pricing policies on duopoly competition with endogenous switching costs, and consider the commitment to a future rebate through a coupon as one of the pricing policies. The results coincide with those I obtain in this Section with the difference that they derive the main results through calibration whereas I solve for the equilibrium tariffs analytically. In any case, the model is intended to provide intuition that motivates the empirical analysis.

[37] Although this assumption is admittedly made for simplicity in order to introduce uncertainty in location and to avoid consumers anticipating second-period results from the first period, there are intuitive interpretations of it. We can think, for example, of a person who remembers running out of toothpaste in the morning when going out of her workplace, and go to buy to the closest supermarket or on the way home; another example is that of a person having a craving for pizza after work and stops by to buy one on the way home. This can be though of as uncertainty of tastes depending on the day or the weather: maybe in a cold-rainy day I can feel like eating lasagna but in a sunny day I would prefer eating salad or gazpacho instead.

At the beginning of the first period, retailers offer a loyalty program to consumers consisting of a coupon B of a given amount of money determined by period 1 purchases and redeemable in period 2 in the same retailer's stores. This rebate is given to a consumer conditional on his subscription to the retailer's LP in the first period. I assume, for simplicity, that consumers can only subscribe to one loyalty program and that the market is fully covered, i.e., the utility of joining a LP is large enough so that it is better to subscribe than the outside option.[38] Finally, I assume non-strategic forward-looking consumers with rational expectations.

The structure of the model defines a two-stage game as follows:

- *First stage:* Retailers determine simultaneously first-period prices and loyalty rebates to maximize their expected two-period profit function.
- *Second stage:* Prices, loyalty rebates and first-period market shares are realized. Each firm then decides its second-period price to maximize second-period profit taking into account its subscribers and the eventual "non-loyal" customers (switchers) it could have in every period.

I am looking for subgame-perfect equilibria.

Definition 1.1. The set of choices $\{p_{11}^*, B_1^*, p_{21}^*\}$ for firm 1 and $\{p_{12}^*, B_2^*, p_{22}^*\}$ for firm 2 is said to be a subgame-perfect equilibrium of the game if

- Second stage: For any given first-period prices and loyalty rebates $\{p_{11}, B_1; p_{12}, B_2\}$, and for any given first-period market shares $\{\alpha_1(p_1), \alpha_2(p_1)\}$ with $p_1 = (p_{11}, p_{12})$, the prices p_{21}^* and p_{22}^* constitute a Nash equilibrium.
- First stage: Given the second-period optimal prices $\{p_{21}^*, p_{22}^*\}$ and market shares $\{D_1, D_2\}$, the quadruple $\{p_{11}^*, B_1^*; p_{12}^*, B_2^*\}$ is a Nash equilibrium.

We are interested in the set of interior equilibria, therefore we impose an upper bound on B_j, which must be small enough so that the consumers' demands remain interior:

$$0 \leq B_j \leq \overline{B}$$

1.A.1 Second period

Assume $\{p_{1j}, B_j\}_{j=1,2}$ as given. Moreover, I assume that a mass $\alpha_j \in (0,1)$ of consumers subscribed to firm j's loyalty program in the first period. The remaining

[38] The subscription to a given LP does not imply that consumer has committed herself to repeat purchases from the same retailer in the second period. As a consequence, consumers have always the choice to go to the retailer with the lower price given their second period location and the size of B, which becomes a switching cost for those who decide to go to a different retailer.

$(1-\alpha_j) \equiv \alpha_k$ subscribed to rival's LP. α_j and α_k denote, hence, firm j, k's first-period market share, respectively.

The utility of a consumer i located at x_2 who has joined firm j's LP is:

$$\begin{array}{ll} u - (p_{2j} - B_j) - \tau x_2 & \text{if he continues buying from retailer } j \\ u - p_{2k} - \tau(1 - x_2) & \text{if he switches to retailer } k, \ \forall k = 1, 2 \end{array}$$

where u is consumer i's valuation of the good and τ is a transportation cost parameter which is constant over time and symmetric across consumers.

Assuming that the consumer who is indifferent between the two retailers is located at $x_2 = D_j$, firm j's market share is defined by:

$$p_{2j} - B_j + \tau D_j = p_{2k} + \tau(1 - D_j)$$

Solving for D_j yields firm j's second-period market share:

$$D_j(p_{2j}, p_{2k}, B_j) = \frac{1}{2} + \sigma \left[p_{2k} - (p_{2j} - B_j)\right], \ \forall j, k = 1, 2 \tag{1.A.1}$$

where

$$\sigma \equiv \frac{1}{2\tau} \tag{1.A.2}$$

is a parameter indicating the degree of substitutability between the two retailers.

Firm j's profit maximization problem writes as:

$$\max_{p_{2j}} \quad \pi_{2j} = (p_{2j} - c)(\alpha_j D_j + (1 - \alpha_j) D_k) - \alpha_j D_j B_j$$

FOC:

$$\alpha_j D_j + (1 - \alpha_j) D_k + (p_{2j} - c) \left[\alpha_j \frac{\partial D_j}{\partial p_{2j}} + (1 - \alpha_j) \frac{\partial D_k}{\partial p_{2j}}\right] - \alpha_j \frac{\partial D_j}{\partial p_{2j}} B_j = 0 \tag{1.A.3}$$

Plugging $\frac{\partial D_j}{\partial p_{2j}} = \frac{\partial D_k}{\partial p_{2j}} = -\sigma$ into (1.A.3) and solving for p_{2j} yields:

$$p_{2j} = \frac{c}{2} + \frac{\tau}{2} + \alpha_j B_j + \frac{p_{2k} - (1 - \alpha_j) B_k}{2} \tag{1.A.4}$$

By the symmetry of the model, p_{2k} has a similar expression to the previous equation. Plugging it to (1.A.4) and solving for p_{2j} we have:

$$p^*_{2j}(\alpha_j) = c + \tau + \alpha_j B_j, \ \forall j = 1, 2 \tag{1.A.5}$$

or

$$p^*_{21}(\alpha_1) = c + \tau + \alpha_1 B_1 \quad \text{and} \quad p^*_{22}(\alpha_2) = c + \tau + \alpha_2 B_2 \tag{1.A.6}$$

Note that if we set $B_j = 0$, we get $p = c + \tau$ which is the standard Hotelling price with unit demands.

1.A.2 First period

A consumer located at x_1 will have the following instantaneous utility:

$$\begin{array}{rl} u - p_{1j} - \tau x_1 & \text{if he goes to retailer } j \\ u - p_{1k} - \tau(1 - x_1) & \text{if he goes to retailer } k \end{array}$$

Conditional on buying from retailer j in the first period, he will buy again in the second period to this retailer if x_2 satisfies:

$$x_2 \leq D_j \equiv \frac{\tau + p_{2k} - p_{2j} + B_j}{2\tau}$$

The consumer i's expected second-period surplus is then:

$$\delta \left[\int_{x=0}^{D_j} (u - p_{2j} + B_j - \tau x)dx + \int_{D_j}^{x=1} (u - p_{2k}(\alpha_k) - \tau(1-x))dx \right] \tag{1.A.7}$$

performing the integrals and rearranging, yields:

$$\delta \left\{ (p_{2k} - p_{2j} + B_j + \tau)D_j - \tau D_j^2 + u - p_{2k} - \frac{\tau}{2} \right\}$$

Hence, consumer i's lifetime utility is:

$$u - p_{1j} - \tau x_1 + \delta \left\{ (p_{2k} - p_{2j} + B_j + \tau)D_j - \tau D_j^2 + u - p_{2k} - \frac{\tau}{2} \right\} \tag{1.A.8}$$

Similarly, conditional on buying from retailer k in the first period, the consumer continues buying from her if:

$$x_2 \geq D_k \equiv \frac{\tau + p_{2k} - p_{2j} - B_k}{2\tau}$$

Then, consumer i's expected total surplus conditional on buying from k is:

$$u - p_{1k} - \tau x_1 + \delta \left\{ (p_{2k} - p_{2j} - B_k + \tau)D_k - \tau D_k^2 + u - p_{2k} + B_k - \frac{\tau}{2} \right\} \tag{1.A.9}$$

A consumer located at $x_1 = \alpha_j$ who is indifferent between retailers makes $(1.A.9) = (1.A.8)$, which results in firm j's market share:

$$\alpha_j = \frac{1}{2} + \tilde{\sigma} \left[(2p_{1k} - \delta B_k) - (2p_{1j} - \delta B_j) \right] \tag{1.A.10}$$

where

$$\tilde{\sigma} \equiv \frac{\tau}{4\tau^2 + \delta(B_j + B_k)^2}$$

is a 'modified' substitutability parameter that takes into account the dynamic effects of LPs. Note that if we set $B_j = B_k = 0$, we obtain the standard market share in the static Hotelling model.

Equation (1.A.10) tells us that even though the consumer does not receive an immediate rebate, he perceives current prices lower as if the loyalty discount were immediate. Also, that a delayed rebate does not have the same impact on the demand as an immediate one as long as the discount factor reduces the impact of this loyalty rebate, stimulating only a fraction of demand increase per unit of price reduction.

Assuming that firms discount the future at the same rate δ as consumers, firm j's overall profit function is given by:

$$\Pi_j = \pi_{1j}(p_{1j}, p_{1k}) + \delta\pi_{2j}\left(\alpha_j(p_{1j}, p_{1k}, B_j, B_k)\right) \tag{1.A.11}$$

where:

$$\begin{aligned}\pi_{1j} &= (p_{1j} - c)\alpha_j(p_{1j}, p_{1k}, B_j, B_k)\\ &= (p_{1j} - c)\left(\frac{1}{2} + \tilde{\sigma}\left[(2p_{1k} - \delta B_k) - (2p_{1j} - \delta B_j)\right]\right)\end{aligned} \tag{1.A.12}$$

and

$$\begin{aligned}\pi_{2j} &= (p_{2j} - c)(\alpha_j D_j + (1 - \alpha_j)D_k) - \alpha_j D_j B_j\\ &= \frac{\tau^2 - \alpha_j(1 - \alpha_j)(B_j + B_k)B_j}{2\tau}\end{aligned} \tag{1.A.13}$$

Firm j's problem is then:

$$\max_{\{p_{1j}, B_j\}} \Pi_j$$

FOCs:

$$\frac{\partial \Pi_j}{\partial p_{1j}} = \frac{1}{4\tau^2 + \delta(B_j + B_k)^2}\left[2\tau\left(p_{1k} - 2p_{1j} + \tau + \delta\frac{(B_j - B_k)}{2} + c\right) + \delta\frac{(B_j + B_k)^2}{2}\right] + \frac{\delta}{2\tau}\left[(B_j + B_k)B_j\frac{\partial \alpha_j}{\partial p_{1j}}\left(-(1 - \alpha_j) + \alpha_j\right)\right] = 0$$

In a symmetric equilibrium, $p_{1j} = p_{1k} = p_1^*, \alpha_j = 1/2, B_j = B_k = B$:

$$p_1^* = c + \tau + \frac{\delta B^2}{\tau}, \tag{1.A.14}$$

$$p_2^* = c + \tau + \frac{B}{2} \tag{1.A.15}$$

And the respective profits are:

$$\pi_1^* = \frac{1}{4\sigma} + \delta\sigma B^2 \qquad\qquad \pi_2^* = \frac{1}{4\sigma} - \frac{\sigma B^2}{2} \tag{1.A.16}$$

$$\Pi^* = (1+\delta)\frac{1}{4\sigma} + \delta\frac{\sigma B^2}{2} \tag{1.A.17}$$

Where σ is the standard substitutability parameter as defined in (1.A.2). Note that if we set $B = 0$ we get $\pi_1 = \pi_2 = \frac{1}{4\sigma}$, which is the Hotelling profit in a model with unit demands.

As shown in the previous equations, retailers' instantaneous profits can be some periods higher and some others lower than those they would obtain in absence of LPs. However, overall profits are higher thanks to the rise in prices induced by LPs. According to equation (1.A.17), overall profits increase with the loyalty rebate. Given the transportation costs, the higher the B the better for the retailer.

1.B First-stage results

In this Section I present the regressions conducted as a first step to obtain predicted price, membership and the deltas necessary to compute the optimal instruments.

1.B.1 Modeling prices

As pointed out previously, I assume that the price of a particular brand in a department is a linear function of brand and supermarkets characteristics and the average prices of the same brand in all the regions of the country (excluding the department from the average price of the region the market is located in). These prices are used as proxies of marginal costs under the assumption that prices contain information on common marginal costs in all its stores across the country, and then retail prices contain information on that common costs (see Section 4 for more details).

Table 1.B.1 displays the results of the linear regression of the price of the brand variety j in supermarket s at t on average prices of the same brand in the 21 administrative regions of France, and brand dummy variables and week dummy variables as controls. The fact that all the estimates are significant at 1% level and the high R-squared suggest that prices have important explanatory power.

Table 1.B.1: Results from a linear regression for price

Variable	Estimates	Standard errors
pR1	-0.632	0.132
pR2	-2.902	0.194
pR3	-1.218	0.227
pR4	-1.546	0.148
pR5	-1.591	0.100
pR6	-2.717	0.194
pR7	-2.820	0.288
pR8	-1.915	0.216
pR9	-1.940	0.170
pR10	-0.895	0.140
pR11	-6.149	0.276
pR12	-2.236	0.191
pR13	-0.863	0.164
pR14	-2.346	0.167
pR15	-3.253	0.212
pR16	-1.008	0.135
pR17	-1.660	0.117
pR18	-1.233	0.139
pR19	-1.310	0.174
pR20	-2.881	0.283
pR21	-6.385	0.338
Adj. R^2	0.840	

Notes: Based on 37,662 observations. Robust s.e. in parentheses. Regressions include brand and time dummy variables.

1.B.2 Modeling the membership decision

The information I have about loyalty programs is very limited: a dummy variable indicating whether a given household was member of the loyalty program of the supermarket it purchased from in 2006 is all I have. Information such as the date of subscription coupons (and amounts) given to households and redemption rates is not available. Due to these data limitations, I cannot fully explain the determinants of the individual decision to subscribe to a loyalty program. Nevertheless, information on loyalty programs characteristics is available on the web and this, along with the household scanner data, can be used to get a clue on how loyalty program characteristics and household characteristics are correlated with the membership status. Further, this exercise helps me evaluate the performance of potential instruments to treat the endogeneity of the loyalty membership indicator.

I collected some data on the characteristics of each loyalty program that do not vary by market, namely, a dummy variable equal to one if the program includes both NBs and PLs in the reward system and zero otherwise, a dummy variable equal to one if the customer can either subscribe to the program by Internet or at least download the subscription form and zero otherwise, the average rebate in euros per 100 € spent in the supermarket, and dummies for whether the program has an upper bound for accumulating rebates, thresholds of points or money needed to get a coupon and a the number of months points can be hold before expiration.[39] Along with this information, I use household characteristics and other variables such as the number of different stores visited by the household members within a week, the number of shopping trips to a same retailer in a month and the number of members to the LP of a given supermarket in the rest of the country (excluding the department where the household was located).

Table 1.B.2 displays the results of OLS and Logit regressions of the dummy variable indicating household membership to a supermarket loyalty program on the variables previously described. After a first specification with all the covariates, some LP characteristics were either non significant at 5% level or had a different sign than expected or both. Thus, I excluded them form the final specification displayed in the table.

Although the set of covariates used do not explain the whole variation of the dependent variable, this is confirmed by the low adjusted R^2, the results are quite appealing. On the one hand, in the Logit regression (column 2) all parameters are significant, and on the other hand, the signs of the variables are according to the intuition. In the first place, the probability of joining supermarket j loyalty program increases with the number of members in the rest of the country, which means that the more popular the program is, the more attractive it is to join, most probably because this is a signal of the program quality. Moreover, this probability also increases with household characteristics such as household size, age, car ownership and living in urban areas which facilitates the access to the retailer, and the number of trips the household members do to the retailer j, being an indicator of shopping intensity, the more often a customer goes to the same store, the more informed about prices and promotions she is and, consequently, the more likely to benefit from rebates.

Second, individual characteristics such as the income and the multi-store shopping behavior make consumers to be less interested in joining LPs because either they are less sensitive to small price cuts or they are less likely to get rewards. In fact,

[39] To obtain this information, I reviewed carefully each supermarket program's "Terms and Conditions of use" document and identify some common categories describing the general structure of the programs. These documents were obtained from each supermarket's website in 2011 and 2012. Although the information is used to instrument a decision that could have been taken even before 2006 (the year of the database used in this article), all the programs where launched before 2006 and I believe the structure has not changed a lot since then.

Table 1.B.2: Results from OLS and Logit regressions for LP membership

Variable	OLS	Logit
LPs characteristics		
Log total Sup. LP members	0.141^{**}	0.892^{**}
excluding location depart.	(0.028)	(0.185)
NBs included	0.0499^{**}	0.322^{**}
	(0.020)	(0.137)
Online access to inscription	0.0410^{**}	0.261^{**}
form	(0.012)	(0.074)
Log of average reward per	0.0110^{**}	0.0692^{**}
100 **€** spent	(0.005)	(0.033)
HH shopping behavior		
# of different stores	-0.272^{**}	-1.643^{**}
visited a week	(0.044)	(0.235)
# trips to the same	0.156^{**}	1.127^{**}
retailer per month	(0.011)	(0.093)
Demographics		
Household size	0.0288^{**}	0.190^{**}
	(0.003)	(0.023)
Log income	-0.0132	-0.0928^{*}
	(0.009)	(0.057)
Log age	0.0365^{**}	0.221^{**}
	(0.015)	(0.088)
Car	0.0683^{**}	0.369^{**}
	(0.022)	(0.114)
Lives in city	0.00869	0.0615
	(0.009)	(0.058)
Constant	-0.632^{**}	-7.515^{**}
	(0.254)	(1.660)
N	9,976	9,976
R^2	0.035	
adj. R^2	0.034	

Robust standard errors in parentheses.
** Significant at 5% and * at 10% level.

one can expect that as income increases supermarket loyalty programs become less attractive because wealthier people may be less sensitive to little price cuts and also because loyalty rebates are generally addressed to middle quality goods rather than to high quality branded products. Regarding the multi-store shopping behavior, it is more difficult to accumulate points or money in a loyalty account when purchases are not concentrated in one retailer, which makes individuals with this shopping profile either less interested to join any loyalty program or willing to join all of them, but not likely to use them as rewards will never be attainable.

Finally, concerning LP characteristics included in the regression, Logit results show that the probability of joining a LP increases with the fact that both NBs and PLs purchases give the customers points to get a future rebate compared to those programs including only PLs, also with the possibility of joining the program online (or at least having access to the inscription form) and the average reward in euros, i.e, the better rewarding programs are the more likely to attract customers to their loyalty program.

1.B.3 First-stage estimation of the full model

The results of the first-stage estimation of the full model using the set of inefficient instruments for the endogenous variables are basically of the expected sign. However, most estimates are statistically non significant due to the very large robust standard errors obtained. In addition, some convergence problems were experienced as the size of the estimates and in some cases the sign were quite sensitive to the provided set of starting values. All this suggests that the instruments used to identify the parameters of the full model are not performing well, even though they are similar to the instruments used in some standard papers. This may be due to the fact that they are just helping to identify the parameters in the linear part of the model (θ_1) and not those of the whole model (θ_1, θ_2). This fact empirically supports the need to use optimal instruments as the best way to overcome endogeneity problems in this kind of models.

1.B.4 Own- and cross-price elasticities

Table 1.B.4 presents the medians of the distribution of estimated own- and cross-price elasticities for the leading 12 brand varieties based on their national market shares. The columns in the table indicate the price that changes, and the rows market shares responses, so that for instance, the entry located in row 2 column one gives the elasticity of brand 2 with respect to a change in the price of brand 1.

One of the advantages of a random-coefficients model, as discussed previously, is that it gives more realistic substitution patterns allowing for different brands to have

Table 1.B.3: First-stage results from the Mixed Logit model with inefficient instruments[a]

Variable	Means (β's)	Std. Deviations (σ's)	Interactions with hh characteristics: HH size	Income	# Subscriptions
Constant	-3.631[b]	1.353		-0.967	0.922
	(1.388)	(1.865)		(13.454)	(4.787)
Price	-6.399	2.301	-1.748	0.764	1.331
	(5.253)	(6.090)	(12.456)	(25.647)	(14.523)
LP member (M_{ms})	-8.338	6.223		-0.967	
	(6.896)	(2.246)		(9.011)	
PL dummy	-2.634[b]	2.755	-1.331	0.764	-7.803
	(2.270)	(3.569)	(4.305)	(4.120)	(5.237)
LP member×PL dummy	7.935	4.347			
	(3.381)	(3.344)			
Plastic	-0.035[b]	2.114			
	(0.483)	(1.266)			
Sugar	-0.223[b]	0.562			
	(0.240)	(4.803)			
Wholemilk	-2.998[b]	0.670			
	(1.035)	(3.779)			
GMM Objective	1,045				
MD χ^2	4139681				

Notes: [a] Based on 37,662 observations. Except where noted, parameters were estimated using GMM. All regressions include brand and week dummies. Asymptotically robust s.e. are given in parentheses.
[b] Estimated using a minimum-distance procedure.

different values for the elasticity due to the change in the price of a given price, as opposed to the simple Logit. As a consequence, estimated elasticities vary by brand and by market.

Table 1.B.4: Median own- and cross-price elasticities

Brand	b1	b3	b7	b8	b9	b11	b12	b15	b16	b18	b21	b23	b24	b28	b29
PLs															
1	-2.612	0.017	0.006	0.004	0.005	0.006	0.010	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000
2	0.033	0.262	0.159	0.119	0.121	0.162	0.180	0.004	0.031	0.002	0.002	0.005	0.027	0.013	0.003
3	0.009	-15.995	0.045	0.029	0.022	0.032	0.035	0.001	0.007	0.000	0.000	0.002	0.008	0.003	0.001
4	0.030	0.269	0.125	0.081	0.081	0.119	0.163	0.002	0.027	0.001	0.001	0.006	0.025	0.009	0.002
5	0.041	0.387	0.271	0.144	0.148	0.239	0.244	0.006	0.043	0.003	0.003	0.009	0.050	0.016	0.005
6	0.009	0.157	0.053	0.035	0.036	0.061	0.049	0.001	0.011	0.001	0.000	0.002	0.009	0.003	0.001
7	0.005	0.052	-8.893	0.018	0.017	0.022	0.019	0.001	0.004	0.000	0.000	0.001	0.004	0.002	0.000
8	0.004	0.040	0.019	-7.310	0.012	0.017	0.021	0.000	0.003	0.000	0.000	0.001	0.004	0.002	0.000
9	0.003	0.023	0.016	0.007	-4.777	0.012	0.011	0.000	0.002	0.000	0.000	0.001	0.003	0.001	0.000
10	0.015	0.091	0.053	0.037	0.037	0.040	0.044	0.002	0.010	0.001	0.000	0.002	0.010	0.004	0.001
11	0.005	0.040	0.022	0.015	0.014	-10.365	0.027	0.001	0.004	0.000	0.000	0.001	0.004	0.002	0.000
12	0.009	0.049	0.031	0.017	0.020	0.031	-12.624	0.001	0.008	0.000	0.000	0.002	0.009	0.003	0.001
NBs															
13	0.008	0.100	0.038	0.027	0.027	0.039	0.047	0.003	0.028	0.002	0.001	0.005	0.030	0.011	0.003
14	0.005	0.047	0.037	0.015	0.016	0.024	0.047	0.003	0.024	0.001	0.001	0.005	0.026	0.008	0.002
15	0.000	0.002	0.001	0.001	0.001	0.001	0.002	-1.834	0.001	0.000	0.000	0.000	0.001	0.000	0.000
16	0.001	0.012	0.007	0.004	0.004	0.005	0.009	0.001	-11.919	0.000	0.000	0.001	0.004	0.002	0.000
17	0.002	0.020	0.009	0.007	0.008	0.012	0.016	0.001	0.009	0.001	0.000	0.001	0.008	0.005	0.001
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.837	0.000	0.000	0.000	0.000	0.000
19	0.010	0.072	0.048	0.025	0.031	0.038	0.089	0.004	0.028	0.003	0.001	0.007	0.025	0.016	0.003
20	0.009	0.142	0.057	0.028	0.020	0.050	0.044	0.005	0.034	0.004	0.002	0.009	0.033	0.011	0.003
21	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.491	0.000	0.000	0.000	0.000
22	0.001	0.016	0.005	0.004	0.005	0.005	0.002	0.000	0.003	0.000	0.000	0.001	0.003	0.001	0.000
23	0.000	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.000	-2.694	0.001	0.000	0.000
24	0.001	0.007	0.003	0.003	0.003	0.003	0.006	0.000	0.002	0.000	0.000	0.001	-10.906	0.001	0.000
25	0.006	0.056	0.024	0.021	0.017	0.030	0.037	0.003	0.018	0.001	0.001	0.005	0.017	0.009	0.002
26	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
27	0.022	0.169	0.091	0.061	0.070	0.080	0.150	0.009	0.070	0.003	0.003	0.015	0.048	0.031	0.006
28	0.000	0.002	0.001	0.001	0.001	0.001	0.002	0.000	0.001	0.000	0.000	0.000	0.001	-4.371	0.000
29	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.852
30	0.004	0.047	0.026	0.015	0.014	0.022	0.025	0.002	0.019	0.001	0.001	0.003	0.018	0.007	0.002
31	0.003	0.025	0.012	0.010	0.008	0.014	0.016	0.002	0.009	0.001	0.000	0.002	0.009	0.004	0.001

Notes: Entry (i, j), where i denotes row and j column, gives the percent change of brand i market share to a one percent change in price of brand j. Each entry was obtained as the median of the elasticities from the 7,704 markets. In the horizontal axis the brands with the largest market shares are displayed. The full matrix is available on demand.

Chapter 2

Multiproduct retailing and consumer shopping patterns: The role of shopping costs

Abstract: We structurally identify consumer shopping costs —real or perceived costs of dealing with a store— using scanner data on grocery purchases of French households. We present a model of demand for multiple stores and products consisting of an optimal stopping problem in terms of individual shopping costs. This rule determines whether to visit one or multiple stores at a shopping period. We then estimate the parameters of the model and recover the distribution of shopping costs. We quantify the total shopping cost in 18.7 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 1.53 € and mean total transport cost of 17.1 € per trip. We show that consumers able to source three or more grocery stores have zero shopping costs, which rationalizes the low proportion of three-stop shoppers observed in our data. Theory predicts that when shopping costs are taken into account in economic analysis, some seemingly pro-competitive practices can be welfare reducing and motivate policy intervention. Such striking findings remain empirically untested. This paper is a first step towards filling this gap.

JEL classification: D03, D12, L13, L22, L81.

Keywords: Grocery retailing, supermarket chains, shopping costs, one- and multistop shopping, Method of Simulated Moments.

2.1 Introduction

Consumers have heterogeneous shopping patterns (see Figure 2.1 below). This heterogeneity might be explained by several factors such as preferences, demographics, geographic location, information frictions, differentiated retailers, and time availability for shopping activities. Previous literature has introduced a concept that accounts for some (or most) determinants: *shopping costs* (Klemperer, 1992, Klemperer and Padilla, 1997, Armstrong and Vickers, 2010, and Chen and Rey, 2012, 2013). In line with this literature, we will call *shopping costs* all real or perceived costs a consumer incurs when sourcing a grocery store. Economic theory shows that in a context of multiproduct retailing and consumer shopping costs, several practices that would otherwise be considered competitive and good from a social welfare perspective can be less competitive. However, there is not much empirical support for such findings, in part because the introduction of shopping costs in a structural model of demand is a challenging task. This motivates the following questions. First, is it possible to quantify shopping costs from observed consumer shopping behavior? Second, will accounting for shopping costs in a multiproduct demand model lead to a better understanding of consumer heterogeneity in shopping patterns? Finally, to what extent the inclusion of shopping costs would be crucial for policy analysis? In this paper, we develop and estimate a structural model of multiproduct demand for groceries in which shopping costs play a key role in consumer decision making. This framework enables us to identify the distribution of consumer shopping costs from data on grocery purchases.

We say that two consumers have heterogeneous shopping patterns when they visit a different number of stores within the same shopping period. Therefore, a consumer sourcing a single store within, say, a week will be a *one-stop shopper* and a consumer visiting several separate suppliers within the same week will be a *multistop shopper*. Consumer shopping costs, which may depend on stores' characteristics (e.g. transport costs depend on store location; the opportunity cost of time from shopping depends on store size) and may as well be informative about consumers' tastes for shopping, account for such differences.[1]

The inclusion of shopping costs in the analysis of multiproduct demand and supply may change policy conclusions dramatically. Consider, for instance, the case of multiproduct retailers competing head-to-head by selling homogeneous products. In the presence of shopping costs, customers will stick with a single retailer because the benefit from visiting an additional supplier need not compensate the shopping

[1] Klemperer (1992) distinguishes among consumer costs in the following way: "...consumer's total costs include purchase cost and utility losses from substituting products with less-preferred characteristics for the preferred product(s) not actually purchased [transport costs of the standard models à la Hotelling] (...) Consumers also face shopping costs that are increasing in the number of suppliers used." p.742.

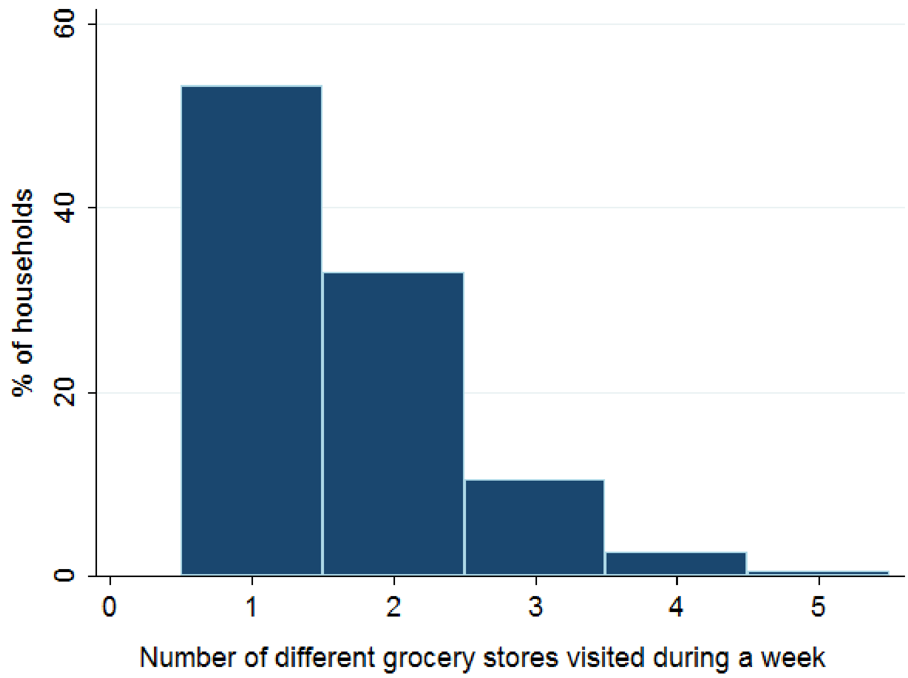
cost. As a consequence, competition is reduced and prices are higher. In contrast, if product lines are differentiated, retailers may be tempted to undercut prices to make one-stop shoppers become multistop by patronizing several separate suppliers (Klemperer, 1992). Further, the presence of shopping costs may lead to the introduction of too many varieties of products with respect to the social optimum. When a retailer introduces a new product, the mass of one-stop shoppers increases because more consumers prefer to concentrate purchases with the retailer supplying a wider product range and save on shopping costs. As a consequence, rivals' profits decrease (Klemperer and Padilla, 1997).

Moreover, shopping costs in a context of multiproduct retailing may change the way we understand below-cost pricing, commonly considered as predatory. Large retailers can adopt loss-leading strategies when competing with smaller rivals to price discriminate between one- and multi-stop shoppers. From this perspective, it is more profitable to keep rivals in the market and motivate customers with low shopping costs to source multiple stores, than pushing them out. Hence, pricing below cost turns out to be an exploitative device rather than a predatory practice (Chen and Rey, 2012). Finally, in a setting of competition between large retailers, in which each has a comparative advantage on some products, cross subsidization strategies may be competitive. Below-cost pricing is again not predatory and it can be good for consumer welfare. Banning this practice may hurt consumers and reduce social welfare (Chen and Rey, 2013).

From an empirical point of view, we can readily find support for the idea that shopping patterns are heterogeneous and that this heterogeneity is explained by differences in shopping costs. Figure 2.1 displays the distribution of the population by the average number of different retailers visited within a week. Moreover, we performed reduced form regressions of the number of different supermarkets visited in a week (which constitutes an indicator of multistop shopping behavior) on demographic variables that are proxies for shopping costs (such as income, age, household size, number of children under 16, etc.) and control for household storage capacity, among others. We found strong empirical evidence showing that multistop shopping depends on how busy the household members could be, i.e. how costly it might be to spend a lot of time in shopping activities.

This paper provides a framework to assess the role of shopping costs in explaining heterogeneous shopping patterns. To do so we develop a structural model in the spirit of the main theoretical contributions on the topic. Consumer optimal shopping behavior is given by a threshold strategy where the choice between one- or multistop shopping depends on the size of individual shopping costs. We are able to take the model to data through parametric specifications of consumer utility and shopping cost along with some distributional assumptions on the unobserved shocks. We use scanner data on household grocery purchases in France in 2005, which is representative of French households and contains information on a wide

Figure 2.1: Distribution of household by average number of stores visited in a week, 2005



Notes: The observed distribution has a longer tail than displayed by the graph as we observe households visiting up to 8 separate retailers per week. However, 99.8% of the observations are concentrated up to 5 stops.
Source: TNS Worldpanel data base.

product range and household demographics. As for grocery stores, an additional data set allows us to observe store characteristics and location.

By solving the implied optimal stopping problem of a consumer who needs to decide how many stores to source, we are able to recover the distribution of shopping costs. We quantify the total shopping cost in 18.7 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 1.53 € and the total transport cost of 17.1 €per trip to a given store. Moreover, we are able to compute the transport and total costs of shopping by store format. Transport and total costs of shopping are decreasing in the size of the stores, on average, as smaller formats are closer to downtowns. The largest total shopping cost, 24.7 €, are incurred by consumers who source big-box stores, because they are farther away from downtown. Sourcing a supermarket or a hard-discounter implies total costs of shopping of 14.3 € and 13.4 € per trip, respectively. Finally, the costs of sourcing a convenience store, 4.8 € per trip, are the lowest provided that they are located in downtown. We find that individuals who source three or more stores in a week have zero shopping costs. This might be an indicator that those households actually visiting more than two separate stores a week should have a strong preference for shopping. Finally, the predicted proportions of shoppers by number of stops are 90.1% of one-stop shoppers, 9.7% of two-stop shoppers and only 0.26% do three-stop shopping.

Related literature

The literature including or measuring explicitly consumer-related costs from an empirical point of view, can be summarized in three categories: *i*) search cost literature,[2] *ii*) switching costs literature,[3] and *iii*) shopping costs literature. In recent years there has been a considerable number of contributions developing models and empirical strategies that allow to identify search costs —these include Hong and Shum (2006), Moraga-Gonzalez et al. (2011), Hortaçsu and Syverson (2004), Dubois and Perrone (2010) and Wildenbeest (2011), and switching costs —these include Dubé et al. (2010), Handel (2010) and Honka (2012).

Less attention has been put on shopping costs. To the best of our knowledge, few empirical papers include explicitly shopping costs when it comes to explain time use or supermarket choice. Brief (1967) models consumer shopping patterns in a Hotelling framework, and estimates transportation as part of consumers' shopping costs.[4] His identification strategy consists of using 'the shopping costs elasticity of demand', as he claims these costs are not directly identifiable. Aguiar and Hurst (2007) evaluate how households substitute time for money by optimally combining shopping activities with home production. They argue that multistop shoppers exist because they want to reduce the price paid for a good, which requires more time. As opposed to them, one-stop shoppers may find it optimal to become frequent customers of the same store and benefit from sales and discounts. All this implies a cost in terms of the time needed to carry out the shopping activity, which is accounted for in their modeling framework.

[2] Both shopping and search costs are often referred to as the opportunity cost of time when people go search (for search costs)/shopping (for shopping costs). The difference stems from the purpose of the time spent, whether the consumer ends up buying a product she was looking for or not, and the available information on prices or product characteristics in different locations (sellers). Search costs appear whenever consumers face search frictions caused by information asymmetries. Shopping costs account for the real and opportunity costs related to the shopping activity which may include a previous search if needed.

[3] As stated by Kemplerer and Padilla (1997), shopping costs differ from switching costs in that the latter derives from the economies of scale from repeated purchases of a product while the former is associated with economies of scope from buying related products.

[4] Brief (1967) claims that the final price paid by a consumer has two components, namely, the "pure" price of the good and the marginal cost of shopping for it. These shopping costs include both explicit, such as transportation costs, and implicit, such as the opportunity costs of shopping, which are related to the "purchaser's valuation of time and inconvenience associated with the shopping trip.".

In the analysis of store choice in the presence of shopping costs, our paper closely relates to Shciraldi, Seiler and Smith (2011). They evaluate the effects of big-box retailing on competition, allowing for the fact that customers may do one- or two-stop shopping. This observed heterogeneity allows them to identify individual shopping costs. However, our approach differs from theirs in at least one important way. In line with previous theory literature, we adopt the view that heterogeneous shopping patterns stem from differences in shopping costs as a modeling feature. In other words, in our model the number of stops is endogenously determined by a stopping rule involving the extra utility and extra costs of sourcing an additional store. This fact enables us to empirically identify the distribution of shopping costs. In this sense, our approach is more closely related to the empirical literature on search costs previously mentioned. In particular, our setup relates to Hortaçsu and Syverson (2004), and Dubois and Perrone (2010).

The rest of the paper is organized as follows. Section 3.2 presents the data and a preliminary analysis of consumers' shopping behavior based on descriptive statistics and reduced-form regressions. Section 3.3 outlines the structural model of multiproduct demand and consumer shopping behavior in the presence of shopping costs. Section 2.4 describes our empirical strategy, discusses identification and presents the estimation procedure. Section 2.5 describes the results. We examine the robustness of our results in Section 2.6. Finally, Section 2.7 concludes and discusses directions for further research.

2.2 Grocery retailing, shopping patterns and opportunity cost of time

This Section aims at giving an overview of the data we use, and a first look at customers' shopping behavior.

2.2.1 The data

This paper uses two complementary data sets. Data on household purchases is obtained from the TNS Worldpanel data base by the TNS-Sofres Institute. It is homescan data on grocery purchases made by a representative sample of 7,490 households in France during 2005. These data are collected by household members themselves with the help of scanning devices. Most households integrating the panel were randomly sampled since 1998 (the TNS Worldpanel is a continuous panel database starting from 1998). Every year, a bunch of new randomly selected households is added to the panel either to replace other households rarely reporting data or to increase sample size.

The data set contains information on 352 different grocery products from around 90 grocery stores including hyper- and supermarkets, convenience stores, hard-discounters and specialized stores. The data is reported at the purchase level, so we observe product characteristics such as total quantity, total expenditure, the store where it was purchased from, brand, etc. In addition, the data include a range of household demographics such as household size, number of children, location, income, number of cars, internet access, storage capacity, etc.

On the other hand, data on stores' characteristics is obtained from the Atlas LSA 2005. It includes information by store category (Hyper-, supermarket, convenience and hard-discount stores) on store location, surface, number of checkouts, parking spots, etc. In particular, store location is key to our analysis as it will enable us to identify transportation costs. This will become apparent in Section 2.4.1.

2.2.2 Customer profile

Table 2.1 gives summary statistics for demographic characteristics of french households observed in the data. The average household in France consists of three members, the household's head age[5] being 51 years old, with approximately 2,350 € of income per month and at least one car. Only half of the households in the sample reported having internet access at home which may give a clue on why internet purchases are not so important in our data set. As for storage capacity and home production, 79% of the households have storage rooms at home and 69% an independent freezer, which may explain low frequency of shopping for some households or one-stop shopping behavior. In particular, it is remarkable that about 39% reported vegetable production at home, which along with the fact that less than 30% of the households are located at rural areas, may be an indicator lower frequency of shopping for these households.

Table 2.2 displays details on consumer shopping patterns. On average, households tend to favor multistop shopping. The average french household visits two separate grocery stores in a week and tend to do a single trip per week to the same store. The average number of days between shopping occasions is 5 days. Notice there is some heterogeneity here, which is indicated by a standard deviation of 4.7 days: some households go every day to a grocery store whereas for some others it takes up to ten days to go back to a store.

Larger store formats are preferred by consumers: on average, the two most frequently visited store formats are Supermarkets and Hypermarkets with 48.4% and 40.5% share on total visits per week. Convenience stores, the small downtown stores supplying a reduced product range generally at higher prices, get the lowest share of visits, with 1.9%. Although convenience stores have the advantage of being

[5] By household head we mean the person mainly in charge of the household's grocery shopping.

Table 2.1: Summary statistics for household characteristics

Variable	Mean	Median	Sd	Min	Max
Demographics					
Household size	2.96	3	1.38	1	9
Income (€/month)	2,352	2,100	1,106	150	7,000
Children under 15 (prop. of hh)	0.35	0	0.48	0	1
Household head's age	50.6	49	14.32	22	76
Lives in city	0.73	1	0.44	0	1
Car	1.55	2	0.80	0	8
Home internet access	0.49	0	0.50	0	1
Storage capacity					
Independent freezer	0.69	1	0.46	0	1
Freezer capacity $> 150L$	0.58	1	0.49	0	1
Storage room at home	0.79	1	0.41	0	1
Vegetables production at home	0.39	0	0.49	0	1

Source: TNS Worldpanel data base.

within walking distance to households location, as opposed to hypermarket that are located outside city centers, the preference for larger stores may be explained by several factors such as bulk shopping, lower prices, sales and promotions (that may be more intense in larger stores) and a larger product range.

Interestingly, households tend to concentrate purchases of particular product categories in the same store format. Table 2.3 gives transition probabilities of visiting a particular store format this week for dairy products conditional on the store format sourced the previous week. The probability of keeping the same store format in most cases is larger than the probability of switching store formats. In particular, the lowest probabilities of switching are for those households sourcing hyper- and supermarkets in the past, which is in line with the preference for larger store formats reported in Table 2.2. Moreover, those households patronizing specialized and other smaller stores ('others') are more likely to switch to a hyper- or a supermarket next period.

Age can be seen as a good indicator of the opportunity cost of time. Aguiar and Hurst (2007) find that older people often pay lower prices because their frequency of both shop trips and retailers visited increases, presumably due to a lower cost of time. In our data we found a similar relationship between shopping frequency indicators and age. Figure 2.2 shows that both the number of trips per store and the number of different stores visited a month increase with age. Older people go shopping more frequently performing more visits to the same retailer as well as more visits to separate retailers than their younger counterparts. This can be thought of

Table 2.2: Summary statistics for household shopping patterns

Variable	**Mean**	**Median**	**Sd**	**Min**	**Max**
No. Trips to same grocery store/week	1.37	1	0.72	1	7
No. separate grocery stores visited/week	1.65	1	0.83	1	8
Days between visits	5.09	4	4.73	1	232
Visits by format (% of total/week)					
Hypermarket	40.48	32.2	34.4	0	100
Supermarket	48.38	47.6	32.6	0	100
Convenience	1.92	0.0	8.7	0	100
Hard discount	9.22	3.7	11.6	0	50

Source: TNS Worldpanel data base.

Table 2.3: Transition matrix for purchases of dairy products by store format

		Purchase at t				
		Hyper	**Super**	**Convenience**	**Hard discount**	**Other**
	Hyper	0.68	0.16	0.16	0.27	0.28
At	**Super**	0.17	0.67	0.25	0.31	0.37
t+1	**Convenience**	0.01	0.01	0.46	0.01	0.02
	Hard discount	0.12	0.12	0.09	0.38	0.10
	Other	0.03	0.04	0.04	0.02	0.23

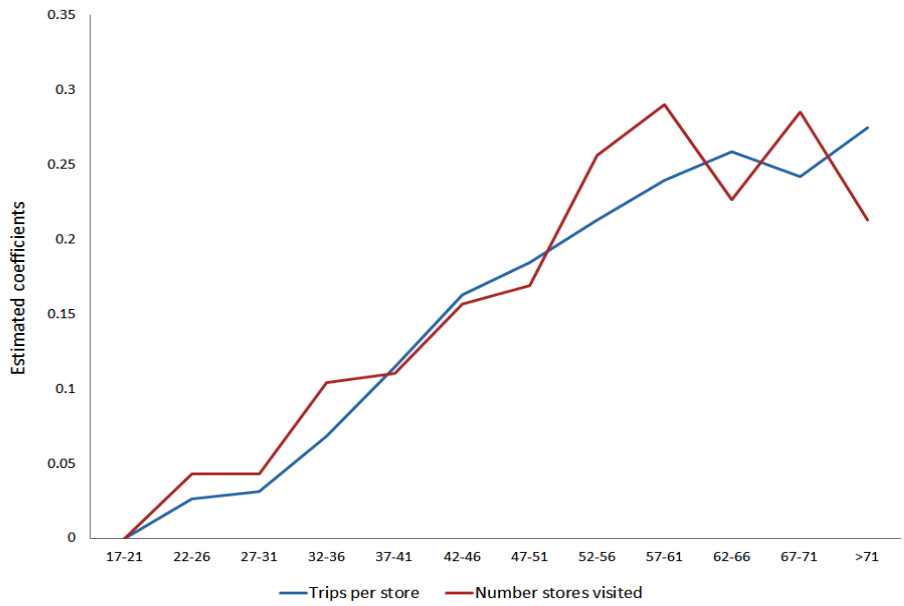
as older people with higher taste for shopping and quality doing more multistop shopping in order to get the best products. It might as well be interpreted as a way to search for the best deals, from an information friction viewpoint. However, the low shopping costs reasoning seems to be more appealing to us because frequent shopping allows people to be better informed about prices and promotional activities without the need to do a search each time they want to go shopping.

2.2.3 Reduced-form results

Recall that shopping costs are the costs of dealing with a store. This implies that multistop shopping, i.e. visiting several separate suppliers in a given shopping period, should be negatively correlated with the consumers' physical as well as time costs. Such a correlation will constitute key empirical evidence of the role of shopping costs on consumer shopping behavior.

In line with theory, we measure multistop shopping as the number of different suppliers visited within a week by the consumer. We regress this variable on the distance from household location to stores and a set of household demographic

Figure 2.2: Frequency of shopping by age ranges, 2005



Notes: Both lines show the results of independent regressions of each variable (*Trips per store* and *Number of stores visited*) on age categories and other demographic controls (income, hh size, car dummy, storage capacity, etc.). Results are based on 5 million observations. All estimates are significant at 1% confidence level.

characteristics which proxy opportunity cost of time, to study the correlation between shopping costs and multistop shopping behavior. Dummy variables to control for region fixed-effects are added in all regressions. Supermarket and time dummies are included gradually in order to assess their effect on the estimates. Further, we add some controls on household storage capacity that can determine the frequency of shopping during the week, namely, type of living place (apartment, farm), storage room, independent freezer, and the size of the largest freezer at home. Table 2.4 gives the results. Most coefficients are of the expected sign and statistically significant at 1% confidence level.

Results provide us with strong empirical evidence on how households' ability to patronize multiple stores depends on how costly it will be in terms of time and distance. Interestingly, we find that larger households living in urban areas tend to favor multistop shopping. On the other hand, higher income people as well as households with babies do less stops on average due, presumably, to a larger opportunity cost of time. Similarly, internet access reduces the number of stops as people can shop online and use home delivery services, which might involve savings on transport costs and time. Growing vegetables at home also reduces the number of stops people want to make probably due to lesser needs for staples. People living in an apartment tend to source more stores as compared to those who live in a house. In contrast, those who live in a larger place, such as a farm, do less stops as compared to families living in a house. This can be explained by the fact that in general, people living in apartments are more likely to be located at or closer to downtown than families living in houses (that tend to be located farther away from city centers) and farms. It might as well indicate that apartments have lower storage capacity than houses and farms.

As expected, distance to stores is negatively correlated with the number of stores visited in a week (see column (1) of Table 2.4). The more distant is the store from consumer location, the larger the transport costs. Notice that distances were excluded from specifications displayed in columns (2) and (3) due to the inclusion of store fixed-effects that are capturing location as a store characteristic that does not vary over time. Finally, in specification given in column (1) we find a negative correlation with car ownership, which can be explained by the fact that people with a car can do bulk shopping at a big-box store, generally located outside downtown areas. However, this relationship becomes positive and non significant in specifications (2) and (3) as we introduce store and time dummies.

2.3 Consumer shopping behavior with shopping costs

Our general strategy is to identify all parameters of the model and retrieve shopping costs cutoffs by setting out a model of demand for multiple grocery products. This way, we can avoid any difficulties related to unobserved data on costs and structure of the supply side.

Our structural model allows for consumer heterogeneity in two dimensions, namely, in the valuation for a particular product and in shopping costs. To keep exposition simple and without loss of generality, we present a model of three grocery stores which will capture the basic intuition of one- and multistop shopping behavior and the role of shopping costs.

2.3.1 General set-up

Demand for grocery products is characterized by different consumers indexed by $i = \{1, \ldots, I\}$ with idiosyncratic valuations for grocery products $k = 1, \ldots, K$.[6] Although valuations and demands may vary with time, we drop the time subscript

[6] Assuming all consumers have access to the same product range might appear strong. However, this help us reducing dimensionality issues in the estimation part. An extension of the model would relax this assumption and allow for heterogeneous choice sets.

Table 2.4: Results for number of different stores visited per week

Variable	(1)	(2)	(3)
HH head's age	0.0025***	0.0032***	0.0032***
	(0.0000)	(0.0000)	(0.0000)
Log Income	-0.0541***	-0.0106***	-0.0104***
	(0.0009)	(0.0009)	(0.0009)
HH size	0.0781***	0.0691***	0.0692***
	(0.0004)	(0.0004)	(0.0004)
Car	-0.0177***	0.0030	0.0031
	(0.0019)	(0.0019)	(0.0019)
Lives in city	0.0416***	0.0517***	0.0516***
	(0.0009)	(0.0009)	(0.0009)
Lives in an appartment	0.0699***	0.0620***	0.0622***
	(0.0011)	(0.0011)	(0.0011)
Lives in a farm	-0.1791***	-0.1605***	-0.1601***
	(0.0028)	(0.0027)	(0.0027)
Baby	-0.1155***	-0.0901***	-0.0898***
	(0.0012)	(0.0011)	(0.0011)
Home internet access	-0.0147***	-0.0086***	-0.0086***
	(0.0008)	(0.0008)	(0.0008)
Grow vegetables home	-0.0122***	-0.0095***	-0.0101***
	(0.0009)	(0.0009)	(0.0009)
Distance to store (km)	-0.0002***		
	(0.0000)		
Constant	1.8866***	1.9142***	1.9153***
	(0.0073)	(0.0075)	(0.0080)
HH storage capacity controls	Yes	Yes	Yes
Dummies per region	Yes	Yes	Yes
Store FE		Yes	Yes
Week FE			Yes
R^2	0.0249	0.074	0.0764

Notes: Regressions are based on 4.72 million observations. Asymptotically robust s.e. are reported in parentheses.
*** Significant at 0.1%.

t for the sake of exposition unless it is strictly necessary. A customer i purchasing product k from store $r \in \{0, \ldots, R\}$ derives a net utility $\overline{v}_{ikr}$.[7]

Consumers want to purchase bundles of these products. Let $\mathcal{B} = \{1, \ldots, R^K\}$ be the set of all possible bundles consisting of combinations of products-stores available in the market, i.e. our bundles account not only for which product was purchased, but what supplier it was purchased from as well. A consumer can either concentrate all her purchases with a single store (*one-stop shopping*) or buy subsets of products from several separate suppliers (*multistop shopping*). At the end of the day, each individual's shopping behavior will be determined by her idiosyncratic cost of shopping.

In the formulation of the model, we focus on the fixed component of the total shopping costs that may account for consumers' taste for shopping. From now on, we will refer to this fixed cost as "shopping costs" and denote it s_i. Physical transport costs, which are an important component of the total cost of shopping, will be accounted for in the empirical implementation of the model by including distances to stores in the utility specification (see Section 2.4).[8] Accordingly, shopping costs are assumed to be independent of store characteristics (size, facilities, location, etc.) and time invariant. Furthermore, we assume s_i is a random draw from a continuous distribution function $G(\cdot)$ and positive density $g(\cdot)$ everywhere.

Finally, we suppose consumers are well informed about prices and product characteristics. Therefore, we assume away information frictions and so consumers' need for searching activities to gather information about prices, qualities and the like.[9]

A consumer i is supposed to have an optimal shopping behavior. This implies she should optimally make a decision that involves choosing between being a one-stop or a multistop shopper and where to go and buy each of the K products of his desired bundle b.

Suppose there are three grocery stores in the market indexed by $r \in \{A, B, C\}$. A consumer will favor multistop shopping if her shopping costs are small enough, otherwise she will optimally concentrate all her purchases with a single store. Roughly speaking, the choice set of consumer i will be restricted by the number of separate stores she can source given her shopping costs, so that her choice will

[7] For now, we do not specify a functional form for the utility as it is not necessary for setting out the model. We will assume a parametric specification at the empirical implementation stage in Section 2.4.

[8] Due to some data limitations, we can only compute distances from the zip code of a given household to the zip code of a given store. Consequently, transport costs will be the same for all individuals living in the same zip code area. See Section 2.4.1 for further details.

[9] This might seem a strong assumption, even though we believe frequent grocery shopping make better informed households and reduce the need to engage in costly search. A more general set up would allow for positive search costs. However, this is out of the scope of this paper and we leave it for future research.

consist of picking the mix of products-stores that maximize the overall value of the desired bundle. In this sense, a three-stop shopper who can visit all three stores will pick the best product from the three alternatives in the market within each category. A two-stop shopper will pick the mix of two stores maximizing the utility of the desired bundle from all the combinations of products-stores possible. Her final bundle will consist of two sub-bundles each containing the best product out of two alternatives in each product category. Finally, a one-stop shopper will pick the store offering the largest overall value of the whole bundle of products.

Formally, let D_{ir}, for all $r \in \{A, B, C\}$ denote the distance traveled by a consumer i from his household location to store $r's$ location, and γ a parameter that captures consumer's valuation of the physical and perceived costs of traveling that distance. Define the utility net of transport costs, of a shopper that can only source one of the three stores in the market as

$$v_i^1 = \max \left\{ \sum_{k=1}^{K} \overline{v}_{ikA} - \gamma D_{iA}, \sum_{k=1}^{K} \overline{v}_{ikB} - \gamma D_{iB}, \sum_{k=1}^{K} \overline{v}_{ikC} - \gamma D_{iC} \right\}. \tag{2.3.1}$$

Similarly, a two-stop shopper has net utility given by

$$\begin{aligned} v_i^2 = \max \Bigg\{ & \sum_{k=1}^{K} \max\{\overline{v}_{ikA}, \overline{v}_{ikB}\} - \gamma(D_{iA} + D_{iB}) \,, \\ & \sum_{k=1}^{K} \max\{\overline{v}_{ikA}, \overline{v}_{ikC}\} - \gamma(D_{iA} + D_{iC}), \\ & \sum_{k=1}^{K} \max\{\overline{v}_{ikB}, \overline{v}_{ikC}\} - \gamma(D_{iB} + D_{iC}) \Bigg\}. \end{aligned} \tag{2.3.2}$$

Finally, a consumer able to source the three stores has net utility given by

$$v_i^3 = \sum_{k=1}^{K} \max \{\overline{v}_{ikA}, \overline{v}_{ikB}, \overline{v}_{ikC}\} - \sum_{r \in \{A,B,C\}} \gamma D_{ir}. \tag{2.3.3}$$

Notice that expressions in (2.3.1), (2.3.2), and (2.3.3) are particular cases of a more general utility function in which, conditional on shopping costs, a n-stop shopper is picking the subset of stores that maximize the overall utility of the desired bundle. For a one-stop shopper, these subsets are singletons, for a two-stop shopper they contain two elements and for a three-stop shopper each subset of stores contains exactly the number of stores in the market, which is why she does not need to maximize over mixes of suppliers.[10]

[10]The general expression of the utility and choice of a n-stop shopper are described in Appendix 2.A.

Suppose $v_i^1 - s_i > 0$ so that all consumers will go shopping at least once. To determine the number of stops to be made, consumer i will compare the extra utility of doing n-stop shopping with the extra costs, taking into account that the total cost of shopping increases with the number of different stores visited. A consumer will optimally decide to do three-stop shopping only if the net utility of visiting three separate stores is larger than what she could obtain by doing either one- or two-stop shopping instead. Formally,

$$v_i^3 - 3s_i \geqslant \max\{v_i^2 - 2s_i, v_i^1 - s_i\}$$

Let $\delta_i^3 \equiv v_i^3 - v_i^2$ be the incremental utility of visiting three stores rather than two, and $\Delta_i^3 \equiv v_i^3 - v_i^1$ be the extra utility of deciding to source either one or three stores. The optimal shopping rule for a three-stop shopper is

$$s_i \leqslant \min\left\{\delta_i^3, \frac{\Delta_i^3}{2}\right\} \tag{2.3.4}$$

A consumer will optimally decide to do two-stop shopping if and only if

$$v_i^2 - 2s_i \geqslant \max\{v_i^1 - s_i, v_i^3 - 3s_i\}$$

Similarly, let $\delta_i^2 \equiv v_i^2 - v_i^1$ be the incremental utility of sourcing two stores rather than one. Hence, a consumer i will do two-stop shopping as long as

$$\delta_i^3 < s_i \leqslant \delta_i^2 \tag{2.3.5}$$

Finally, a consumer will optimally decide to do one-stop shopping if and only if

$$v_i^1 - s_i \geqslant \max\{v_i^2 - 2s_i, v_i^3 - 3s_i\}$$

from which we can derive the optimal shopping rule of a one-stop shopper as

$$s_i > \max\left\{\delta_i^2, \frac{\Delta_i^3}{2}\right\} \tag{2.3.6}$$

In general, the optimal shopping rule for consumer i indicates that she will choose the mix of suppliers to maximize her utility, conditional on the extra shopping cost being at most the extra utility obtained from sourcing additional stores. Equations (2.3.4), (2.3.5) and (2.3.6) suggest we can derive critical cutoff points of the distribution of shopping costs. It is necessary though to determine how are δ_i^2, δ_i^3 and $\Delta_i^3/2$ ordered. From six possible orderings only one survives,[11] namely,

[11]We explain why this is so in Appendix 2.B.

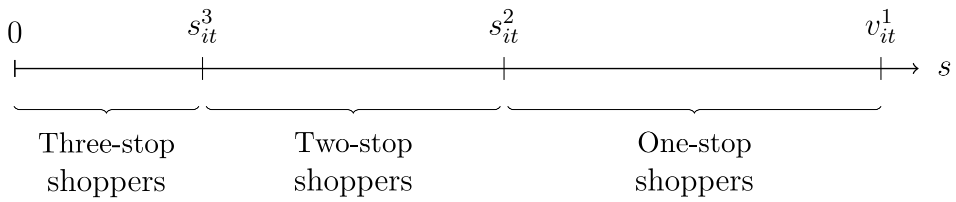
$$\delta_i^3 < \frac{\Delta_i^3}{2} < \delta_i^2, \tag{2.3.7}$$

Under this ordering, the highest possible shopping costs of any consumer able to do multistop shopping at either two or three stores in equilibrium are given respectively by the following critical cutoff points:

$$\begin{aligned} s_{it}^2 &= \delta_{it}^2, && \text{for two-stop shopping, and} \\ s_{it}^3 &= \delta_{it}^3, && \text{for three-stop shopping.} \end{aligned} \tag{2.3.8}$$

Notice that these cutoff points depend on the period of purchase —the subscript t was added— because it depends on utilities that may vary across periods. This contrast with individual shopping costs which are assumed to be time invariant. Cutoffs in (2.3.8) say that for given shopping costs, consumers only care about marginal extra utility of visiting an additional store to make their final decision on how many stores they should optimally source. Moreover, one-, two- and three-stop shopping patterns arise and will be defined over all the support of $G(\cdot)$ –see Figure 2.3.[12]

Figure 2.3: One-, two- and three-stop shopping



2.3.2 Aggregate demand

Let $\mathcal{B}_{2i}, \mathcal{B}_{3i} \in \mathcal{B}_i$ be subsets of bundles involving two- and three-stop shopping, respectively. Recall our previous assumption $v_{it}^1 - s_i > 0$ for all $i = 1, \ldots, I$, which means that all consumers will do at least one shopping trip per week. This implies that the outside option is chosen with probability zero, i.e. $G(v_{it}^1) = 1$. The intuition behind this is as follows: a likely outside alternative to grocery shopping is home production, which consists of households transforming time and market goods into consumption products according to a given home production function (see Aguiar and Hurst, 2007). Yet, even if the household chooses to produce at home most of its

[12]Notice that the kind of behavior according to which a shopper evaluates extreme choices such as visiting all retailers against only one does not appear to be relevant here.

preferred products, there is still a bunch of them that will be too costly to produce compared to the retail price (e.g. toothpaste, toothbrush, cleaning products, bulbs, medicines, etc.). Then, we can think of household members going from time to time to a store to get the set of products they are not able to produce at home (or even the inputs to produce at home the final products they wish to consume).[13]

Aggregate demand for product $k = 1, \ldots, K$ supplied by retailer r is given by

$$\begin{aligned} q_{krt}\left(\mathbf{p_t}\right) = \left[1 - G\left(s_{it}^2\left(\mathbf{p_t}\right)\right)\right] P_{it}^1(X_{\mathcal{B}_i};\theta) \\ + \left[G\left(s_{it}^2\left(\mathbf{p_t}\right)\right) - G\left(s_{it}^3\left(\mathbf{p_t}\right)\right)\right] \prod_{\{b\,\in\,\mathcal{B}_{2i}\,|\,kr\,\in\,b\}} P_{it}^2(X_{\mathcal{B}_i};\theta) \\ + G\left(s_{it}^3\left(\mathbf{p_t}\right)\right) \prod_{\{b\,\in\,\mathcal{B}_{3i}\,|\,kr\,\in\,b\}} P_{it}^3(X_{\mathcal{B}_i};\theta), \end{aligned} \tag{2.3.9}$$

where P_{it}^3 is the probability that a one-stop shopper decides to stop at r, P_{it}^2 is the probability that a two-stop shopper chooses to source retailer r as one of the two retailers she will optimally stop at, and P_{it}^3 is the probability that a three-stop shopper decides to pick a bundle b including product kr. All these probabilities are known by consumers.

The own- and cross-price elasticities of demand are given by the standard formula $\eta_{krht} = \frac{\partial q_{krt}}{\partial p_{kht}} \frac{p_{kht}}{q_{krt}}$, for all $h \in \{A, B, C\}$. It is important to note that a price change may affect not only the market shares per type of shopper but also the shopping costs cutoff values provided they depend on utilities. As a consequence, the distribution of shoppers between one-, two- and multistop shopping changes. In fact, an increase in product k's price at retailer r reduces the indirect utility of consumer i making a stop at r. She may therefore consider to make less stops and purchase a substitute for this product from rival retailer, say h, as the extra gain in utility from sourcing an additional store may not compensate the extra shopping cost.

2.4 Empirical implementation

As described in Section 3.3, consumer choice set consists of bundles of products that can be purchased from one or several stores. Accordingly, if we consider R stores and K products, we would have to deal with a choice set of R^K alternative bundles for each individual, which grows exponentially as R or K increases, resulting in a dimensionality problem which will make estimation challenging and burdensome, whereas it might not change the results in an important way. We circumvent this problem by restricting attention to a reduced set of products and grocery stores. We

[13] The outside option might as well be thought of as not shopping on a weekly basis (for instance, going once a month or every other month). However, in our data the proportion of households not purchasing on a weekly basis corresponds to 8%.

select pre-packaged bread, ready-to-eat breakfast cereal and yogurt as the products to be included in our analysis, provided that they meet the following conditions. First, they are staples and so they are frequently purchased and heavily consumed by french households (see Table 2.1). Second, they belong to different categories of products, which ensures we can observe enough variation in shopping patterns as people may tend to concentrate purchases of the same category in a particular store but might want to diversify across categories. Finally, these products are likely to be of unit demand, i.e., consumers tend to consume one serving of the product at a time and to not mix varieties (see Table 2.1 for details on how we define servings).

Table 2.1: Chacteristics of the selected producs

Product	**Serving** (in grams)	**Consumers (% of pop.)**[a] Kids	Adults	**Position among 352 products**[b]
Yogurt	125	90.7	83	2
Bread	28	95.2	98.5	20
Breakfast cereals	34	60.4	16.8	30

[a] *Source:* Étude Inca (Afssa) 2006-2007 by *Agence Française de Securité Sanitaire des Aliments.* Yogurt and pre-packaged bread appear in the Inca study as part of broader categoires including similar products, namely, "bread and dried bread" and "Ultra-fresh dairy", respectively. Percentages of consumption correspond to consumption of all products in the categories.
[b] These are the positions of the considered products in a ranking of the 352 observed products in our data set, TNS Worldpanel 2005 by frequency of purchase.

Concerning grocery stores, we restrict attention to the two leading supermarket chains in France selected according to national market shares in 2005. The remaining grocery stores observed in our data are treated as part of a composite store which sells the three products we referred to above and constitute an outside option to the two leading chains. In other words, consumers have three alternative stores in their choice set: two insiders and an outside option. This will be enough to describe one- and multistop shopping behavior and to estimate shopping costs cutoffs.

Notice that a bundle can consist partially or fully of products purchased from the outside retailer. Consider, for example, the case of three stores {A, B, O} supplying three products $k = 1, 2, 3$. Let two bundles be $b = \{1_A, 2_B, 3_O\}$ and $b' = \{1_O, 2_O, 3_O\}$. The former will be the choice of a three-stop shopper purchasing product 1 from store A, product 2 from B and product 3 from the outside store O, whereas the latter corresponds to the choice of a one-stop shopper purchasing all products from the outside store. We call the latter bundle the outside good or the zero bundle, $b = 0$.

We empirically specify the utility of consumer i from purchasing good k from store r at time t as

$$\overline{v}_{ikrt} = \begin{cases} -\alpha p_{krt} + X_{kr}\beta + \xi_t + \epsilon_{ikrt}, & \text{if} \quad r = \{A, B\} \\ \epsilon_{ikOt}, & \text{if} \quad r = O \end{cases} \tag{2.4.1}$$

where, p_{krt} is the price of good k at store r, X_{kr} are product-store observed characteristics, ξ_t are time fixed effects, ϵ_{ikrt} is an idiosyncratic shock to utility, which rationalizes all remaining week-to-week individual variation in choices, and α and β are parameters common to all individuals. For simplicity, we normalize the mean utility of the product varieties supplied by outside store to zero.

Notice that equations (2.3.1) through (2.3.3) along with equation (2.4.1) fully specify the utilities of one and multi-stop shoppers as a function of price of the product, product characteristics, and distance to the stores, among others. Put it that way, our utility accounts for both vertical and horizontal dimensions of consumers' valuations for products. The former is captured by included product-store characteristics. The horizontal differentiation aspect is captured by distances which vary across postal codes.[14]

Further, we assume that individual shopping costs are a parametric function of a common shopping cost across all consumers ς, which can be thought of as the minimum cost every consumer bears due to the need of going shopping, and an individual deviation from this mean η_i, which rationalizes the individual heterogeneity in shopping costs, this yields

$$s_i = \varsigma + \eta_i \tag{2.4.2}$$

we assume $\eta_i \sim \mathcal{N}(0, 1)$.

Remark that even though the choice set for all consumers is the same (i.e. all products from all retailers are available for purchase), consumers with large shopping costs visiting an inferior number of retailers than there is in the market are not able to choose the first best option from each product category. Therefore, shopping costs limit the set of alternatives available for one- and two-stop shoppers. Under our setup, this can be thought of as the result of a constrained maximization processes rather than suboptimal choices or mistakes.

2.4.1 Identification

Equation (2.3.8) show that we can identify critical cutoff points of the distribution of shopping costs if we are able to both observe the optimal shopping patterns

[14]If several goods are purchased at the same retailer, the distance to it will only be counted once; the distance will be divided evenly across goods purchased from the same retailer.

of one-stop and multistop shoppers and identify the parameters of the per product utilities involved in the computation of the nth cutoff point. In other words, for each individual we need to identify the utility of her actual choice, say two-stop shopping, and the utility she would have derived had she chosen one-stop shopping instead. To do this, we exploit the panel structure of our data. For most households we observe enough cross-section variation in choices of products and stores, which allows us to identify the utility parameters. In particular, the price coefficient is separately identified from the mean utility from choice data alone due to the observed variation in prices per product. The predicted probabilities will vary due to this variation in prices, which generates enough moments for identification.

On the other hand, (fixed) shopping costs and shopping costs cutoffs are identified from the observed week-to-week variation in shopping patterns, i.e. a household making one-stop shopping this week might be doing multistop shopping next week, meaning that a given household can be more or less time constrained in different weeks. A key point in the identification of fixed shopping costs is the inclusion of other sources of shopping costs that may vary across retailers and periods. An important component in this class of costs is transport costs. Following Dubois and Jódar-Rosell (2010), we empirically identify transport costs by including distances from households' locations designated by postal codes. All households located at a same postal code will have the same distance to retailers nearby.[15] The inclusion of distances to stores will be useful for two purposes: they will capture the horizontal dimension of consumers' preferences for product characteristics and, on the other hand, will allow as to identify the disutility of transport. By adding this information to the model along with the unit demand assumption, the remaining variation in shopping costs across consumers can be interpreted as a pure idiosyncratic shopping cost that is constant across stores, consistent with our set up.

Finally, the identification of aggregate demand requires the computation of the mass of one-, two- and three-stop shoppers, which in equation (2.3.9) are defined as the differences of the distribution of shopping costs $G(\cdot)$ evaluated at two different cutoff values. Given our setup, we are able to compute those values from the empirical distribution of customers between one-, two- and three-stop shopping that we observe in our data.

There may be some endogeneity problems, in particular that of the correlation between prices and the utility shock. In addition, the method of estimation we apply and describe below relies on moment conditions, which requires a set of exogenous instruments. To account for this, we follow Nevo (2001) and use average regional prices of the product to be instrumented for as IVs. These IVs are standard in the IO literature and are proved to work well. We provide further details on the validity of instruments in subsection 2.4.3.

[15] Due to data limitations, we do not observe the exact locations of neither households nor retailers but postal codes only. As a consequence, we are not able to compute exact distances.

2.4.2 Estimation

In this section we present details on how we estimate the utility parameters, and the mean and cutoff values of the distribution of shopping costs. We estimate the parameters of the model presented in the previous section using the data described in Section 3.2. Consistent with this reduced product set and the assumptions of the model described in Section 3.3, the final sample we use consists of local areas where we observe one-, two- and three-stop shopping behavior and households purchasing at least one unit of each product considered here (see Appendix 2.C for further details on how we define units and how we deal with these three goods in a discrete choice context).

The key point of our estimation strategy is to exploit population moment conditions and estimate the parameters of the model by the method of moments for reasons that will become clear below. Therefore, we need to express our discrete choice problem as moments and match population moments with empirical moments in the data. Recall the choice problem we are analyzing. A consumer who wish to buy a set of products K, faces a set $\mathcal{B}$ of mutually exclusive and exhaustive alternatives consisting of combinations of products and retailers available in the market. She will purchase the K products from $n \in \{1, 2, 3\}$ stores, call it bundle $b \in \mathcal{B} = \{1, \dots, 27\}$, such that she can obtain the highest utility net of shopping costs. This maximizing behavior defines the set of unobservables leading to the choice of bundle b as

$$A_{ibt}(X_{\mathcal{B}}; \theta) = \{(\epsilon_{it}, \eta_i) | v^n_{ibt} - ns_i > v^m_{ib't} - ms_i \ \forall m \in \{1, 2, 3\}, b' \in \mathcal{B}\}$$

where $X_{\mathcal{B}}$ is a matrix of characteristics of all alternatives including prices. The response probability of alternative b as a function of characteristics of products and retailers, given the parameters, is given by

$$P_{\mathcal{B}}(b|X_{\mathcal{B}}; \theta) = \int_{A_{ibt}} dF(\epsilon) dF(\eta) \tag{2.4.3}$$

A natural way to estimate the parameters of the model seem to be the maximization of the log-likelihood function

$$L(X_{\mathcal{B}}, d, \theta) = \sum_{i,b,t} \mathbb{1}_{ibt} \log P_{\mathcal{B}}(b|X_{\mathcal{B}}; \theta) \tag{2.4.4}$$

However, given the functional form of the utilities specified in equations (2.3.1) through (2.3.3), maximum likelihood estimation turns out to be extremely difficult to implement as the likelihood of the problem is very nonlinear in the utility shocks. We overcome this problem by using the Method of Simulated Moments (MSM) introduced by McFadden (1989) and Pakes and Pollard (1989).

Let $d_{ibt} = \mathbb{1}\{v^n_{ibt} - ns_i > v^m_{ib't} - ms_i\}$ be an indicator function taking on 1 when bundle $b \in \mathcal{B}$ implying n number of stops is chosen by consumer i and zero otherwise. This information is observed in the data for each consumer i every week. The expected value of d_{ibt} conditional on a set of measured characteristics $X_{\mathcal{B}}$ writes as

$$\mathbb{E}[d_{ibt}|X_{\mathcal{B}}, \theta] = P_{\mathcal{B}}(d_{ibt} = 1|X_{\mathcal{B}}; \theta) \tag{2.4.5}$$

To simulate $P(\cdot)$, we proceed as follows:

1. We build the whole choice set consumers face independently of their shopping costs. This is, we construct bundles as all possible combinations of three retailers and three goods. As a whole, we obtain a choice set of 27 bundles that account for all possible shopping patterns.

2. We assume the shock to utility ϵ_{ikrt} is distributed i.i.d. type one extreme value and take S random draws $\epsilon^s_{ikrt} \forall s = 1, \ldots, S$ per individual, product, retailer and week. Similarly, we assume the shopping costs shock η_i is distributed i.i.d. standard normal and take S random draws $\eta^s_i \ \forall s = 1, \ldots, S$ per individual. Consistent with our assumption of constant shopping costs, we replicate this draws for all retailers and periods whenever we observe purchases by consumer i.

3. Using a vector of initial parameter values, $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \varsigma_0)$ randomly drawn from a normal distribution, along with drawn shocks $(\epsilon^s_{ikrt}, \eta^s_i)$ we are able to compute utilities for all product-retailer choices and consumers, as well as shopping costs to simulate the consumer choice problem described in our modeling framework for each $s = 1, \ldots, S$.

4. From these simulations, we observe what bundle (stores-products combination) maximizes the utility net of shopping costs of each individual in a given week and form an indicator variable for the implied choices, which we denote $d^s_{ibt} \forall \ b \in \mathcal{B}, s = 1, \ldots, S$.

5. Finally, we approximate the choice probability as

$$\check{P}_{\mathcal{B}}(d_{ibt} = 1|X_{\mathcal{B}}, \theta) = \frac{1}{S} \sum_{s=1}^{S} d^s_{ibt} \tag{2.4.6}$$

Plugging the simulated statistics into (2.4.5), rearranging and introducing instruments that may be functions of $X_{\mathcal{B}}$ (we defer to the next subsection the discussion of the instruments we use), we have the following moment conditions

$$\mathbb{E}\begin{bmatrix} w_{1i}\left(d_{i1t}-\check{P}_{\mathcal{B}}(d_{i1t}=1|X_{\mathcal{B}},\theta)\right) \\ \vdots \\ w_{Ni}\left(d_{i27t}-\check{P}_{\mathcal{B}}(d_{i27t}=1|X_{\mathcal{B}},\theta)\right) \end{bmatrix}=0$$

We estimate the parameters of the model by making the sum of the squares of the residuals inside the expectation above across individuals as close as possible to zero. Formally,

$$\min_{\theta}\left[\sum_{i=1}^{I} Q(w_i, X_{\mathcal{B}}, d_{it}, \theta)\right]'\left[\sum_{i=1}^{I} Q(w_i, X_{\mathcal{B}}, d_{it}, \theta)\right],$$

where $Q(\cdot)=\left[w_{1i}\left(d_{i1t}-\check{P}_s(d_{i1t}=1|X_{\mathcal{B}},\theta)\right),\ldots,w_{Ni}\left(d_{i27t}-\check{P}_s(d_{i27t}=1|X_{\mathcal{B}},\theta)\right)\right]'$.

The Method of Simulated Moments (MSM) estimator is then given by

$$\hat{\theta}_{MSM}=\arg\min_{\theta}[d-P(\theta)]'W'W[d-P(\theta)],$$

where $W=[w_1,...,w_I]$ is a $N\times I$ matrix of instruments.

Given the way simulated probabilities are computed in (2.4.6), they are not continuous in θ. It implies that the objective function previously described, which is a sum of simulated probabilities, is not continuous either. As a consequence, analytical methods cannot be used in the optimization process nor standard optimal instruments (which are derivatives of the simulated probabilities evaluated at a consistent estimator of the true parameters) nor the computation of standard errors (which require the use, among other things, of the first derivative of the GMM objective function). These discontinuities do not jeopardize the consistency of simulation estimators. Pakes and Pollard (1989) derive asymptotic properties for a broad class of simulation estimators (including McFadden's MSM) that cover cases where the objective function is discontinuous in the parameters. In practice, to circumvent the discontinuity problem we use a numerical search ('Pattern search') method in the optimization process. As for the computation of standard errors, we apply parametric bootstrap methods.

2.4.3 Instruments

In order to obtain consistent estimates of the parameters of the model, we require to deal with the potential correlation of prices with the error term of the model, ϵ_{ikrt}. In our framework, this error term, known by the consumer but unobserved to the econometrician, is interpreted as a shock to utility that affects demand. If we assume that firms, that may observe these shocks through the observed demand curves, will react to changes in ϵ_{ikrt} by adjusting prices, it will bias the estimate of price sensitivity, α.

To treat this endogeneity problem we assume for simplicity that marginal costs are linear and depend on product and store characteristics and cost shifters, and that markets are competitive so that firms set prices at marginal cost.[16] However, as we do not observe any cost shifters in our data set, we use average regional prices of the same product in all the 21 French administrative regions (excluding the department to be instrumented for from the average price of the region it is located in) as proxies for marginal costs information. Following Nevo (2001), after controlling for product-retailer-specific means, individual shocks might still be correlated within a city but are uncorrelated with product valuations of people from other regions. This implies that in case a demand shock happens in one region, only the local price will be affected. This guarantees the exogeneity condition of prices. Now, what makes average regional prices good instruments is the fact that prices from two different locations (cities, departments, etc.) in a country are linked by common marginal costs as long as they are produced (supplied) by the same manufacturer (retailer) or under a standardized process.[17]

2.5 Results

Table 2.1 displays MSM estimates of the utility parameters, according to two specifications. The first column corresponds to the simplest model including the main covariates and controlling for product and time fixed-effects. The second column shows the results of a specification including IVs. Most coefficients are significant, and results are as expected: demands are downward sloping and the estimate for the distance shows that the value of a product decreases as the retailer is farther away from customer's dwelling. The estimate for mean shopping costs is positive (as expected) and significant in both regressions.

After introducing IVs in the model, we obtain a larger price estimate which may indicate a downward bias in the estimate without instruments. On the other hand, the coefficients for distance seemed to be biased upwards, as we obtained a lower estimate. Finally, the men shopping cost estimate does not differ when we add IVs.

Table 2.2 displays the estimates for the mean shopping cost and the distance in euros. It also shows the values in euros of the average cutoffs of the distribution

[16] In a discrete choice framework, Reynaert and Verboven (2014) examine both perfect and imperfect competition cases and obtain similar results.

[17] Although the independence assumption seems reasonable, there may be cases were it cannot hold as, for example, a national demand shock as pointed out by Nevo (2001).

Table 2.1: Estimates for the utility parameters and shopping costs[a]

Variable	(1)	(2)
Price (€/basket[b])	-1.43^{**}	-1.91^{**}
	(0.72)	(0.83)
Distance (km)	-9.03^{***}	-8.14^{***}
	(3.18)	(1.80)
Mean Utility Bread (28gr)	1.98^{**}	-0.12
	(0.91)	(1.29)
Mean Utility Cereal (35 gr)	-0.22	-0.01
	(0.59)	(0.53)
Mean Utility Yogurt (125 gr)	0.20	2.21^{*}
	(0.83)	(1.25)
Mean Shopping Costs	2.92^{***}	2.93^{***}
	(0.15)	(0.31)
Time fixed-effects	Yes	Yes
Instruments[c]		Yes

Notes: [a] Based on 6,192 observations consisting of purchases of the three considered products made by 2,929 in 2005. Bootstrap standard errors are in parenthesis.
[b] A basket contains a serving of each of the considered products: a slice of bread (28g), a bowl of cereal (35g) and one yogurt (125g).
[c] Instruments include prices of the same good from other geographic locations, as well as bundle dummy variables.
*,**,*** are significant at 10, 5 and 1% confidence levels.

of shopping costs in euros, calculated following equation (2.3.8) and using predicted utilities. In order to translate these values into euros, we divided each of them by the absolute value of the estimated price coefficient. The estimate for the distance, obtained in principle as the disutility of transport, is reinterpreted here as a cost. To do this, we took the absolute value of the original estimate and divide it by the absolute value of the price coefficient.

In line with this, the average fixed cost of shopping is 1.5 € per trip. In addition, visiting a grocery store implies a cost of 4.26 € per km, for the average consumer. The distance between the median consumer's dweling to a store is 4 km, which multiplied by the transport cost per km gives a total transport cost of 17.1 €. Summing up with the mean shopping cost per trip, gives an average total cost of shopping of 18.7 € per retailer sourced (see Table 2.3).

As for shopping costs cutoffs, our results indicate that consumers should have almost-zero shopping costs to be able to source more than two retailers in a week. This rationalizes the small proportion of three-stop shoppers observed in our data.

Notice that the threshold of three-stop shopping, s^3, in column (2) of Table 2.2 is negative. As stated previously, shopping costs may account for consumer's taste for shopping. In line with this, a shopper having a negative shopping cost means that she has a stronger taste for shopping, so that using multiple suppliers makes her total cost of the shopping experience lower than it would be had she decided to concentrate purchases with a single supplier.

One-stop shoppers are all those having shopping costs beyond 2.12 € per trip. A former one-stop shopper will find it optimal to source an additional retailer if her shopping costs were slightly lower than 2.12 €, yet sourcing a third retailer may require a large decrease in shopping costs, such as having more time available or enjoying a lot multi-stop shopping in a given week. The estimates allow us to retrieve the predicted proportion of shoppers by number of stops: 90.1% are one-stop shoppers, 9.7% are two-stop shoppers and only 0.26% do three-stop shopping.

Table 2.2: Mean shopping costs, mean distance and average shopping costs cutoff (across periods and consumers) in euros[a]

	(1)	(2)
Total shopping costs		
Mean shopping cost	2.04	1.53
Mean transport cost	6.31	4.26
Average shopping costs cutoffs		
One-two stops ($\hat{s}^2$)	2.85	2.12
Two-three stops ($\hat{s}^3$)	0.02	-0.02
Predicted distribution of shoppers (% of total)		
One-stop shoppers		90.07
Two-stop shoppers		9.68
Three-stop shoppers		0.26

Notes: [a] To transform estimates into euros, we divide each coefficient by the absolute value of the price coefficient.
[b] To interpret the coefficient for distance as a transport cost, we take the absolute value of the original estimate presented in Table 2.1. It is negative in principle because it enters an utility function, expressing therefore a disutility of transportation.

Table 2.3 gives total transport costs and total cost of shopping (transport plus fixed shopping costs) by store format. The median distance to a big-box store (or hypermarket) is 5.4 km, which multiplied by the transport costs per km gives a total transport cost of 23.2 €, and by adding the mean shopping cost of 1.53 €per trip to a store, gives a total cost of shopping of 24.7 € the average consumer bears each time he visits a large store. Transport and total costs are decreasing in the size of the stores, on average, as smaller formats are closer to downtowns. Sourcing a supermarket or a hard-discounter implies transport costs of 12.8 € and 11.9 € per trip, and total costs of shopping of 14.3 € and 13.4 € per trip, respectively. Finally, the costs of sourcing a convenience store are the lowest provided that they are located in downtowns: the median distance to a convenience is 0.8 km, the transport costs are 3.2 € and the total costs of shopping are 4.8 € per trip.

Table 2.3: Transport costs and total shopping costs (fixed plus transport), by store format (averages across periods and consumers) in euros[a]

Store format	Median Distance (km)[a]	Transport costs (€)[b]	Total costs of shopping (€)[c]
Hypermarket	5.4	23.2	24.7
Supermarket	3.0	12.8	14.3
Hard discounter	2.8	11.9	13.4
Convenience	0.8	3.2	4.8
Overall average	4.0	17.1	18.7

Notes: [a] We use the median of the distance and not the mean, to avoid the effects of outliers.
[b] Computed as the mean transport cost, 4.26 €/km given in column (2) of Table 2.2, times the median distance.
[c] Computed as the sum of Transport costs plus the mean shopping cost of 1.53€per trip, in column (2) of Table 2.2.

In Table 2.4, we present own- and cross-price elasticities. Due to the discontinuity of the predicted choice probabilities described in the estimation section, we cannot compute the derivatives of the demand functions with respect to price analytically. To overcome this problem, we simulated a price increase of 20% for one product at a time, recomputed the utilities for each product and each individual, and retrieved predicted choice probabilities again, to finally get new demands. We take the difference between the demand after the price increase and the baseline demand, and divide it by the price change. Following the standard formula, we then obtain price elasticities of demand as the product of the numerical derivative and the original price, divided by the baseline quantity.

As expected, we obtain negative own-price elasticities and positive cross-price elasticities for the same product category across retailers. This indicates that, on average, consumers may switch retailers when the price of the desired product increases in their patronized retailers. Interestingly, within retailer cross-price elasticities are negative. This means that a price increase in a particular product causes a drop in demand for all other products the consumer intends to purchase. This complementarity effect might be driven by the larger mass of one-stop shoppers. For given prices of the products, a one-stop shopper should pick the retailer in which she derives the maximum value of the desired bundle. If the price of a product category raises in the chosen retailer, the shopper would need to source a competing retailer due to the impossibility of sourcing two or more.

Table 2.4: Mean elasticities (across periods and consumers)

Changing price	Retailer 1			Retailer 2			Outisde retailer		
	Bread	Cereal	Yogurt	Bread	Cereal	Yogurt	Bread	Cereal	Yogurt
Retailer 1									
Bread	-0.0044	-0.0042	-0.0039	0.0040	0.0040	0.0037	0.0061	0.0053	0.0075
Cereal	-0.0070	-0.0080	-0.0065	0.0054	0.0059	0.0053	0.0074	0.0084	0.0095
Yogurt	-0.0087	-0.0088	-0.0098	0.0069	0.0069	0.0076	0.0072	0.0069	0.0125
Retailer 2									
Bread	0.0041	0.0040	0.0038	-0.0046	-0.0043	-0.0040	0.0064	0.0056	0.0079
Cereal	0.0059	0.0062	0.0057	-0.0069	-0.0080	-0.0064	0.0078	0.0090	0.0104
Yogurt	0.0074	0.0074	0.0080	-0.0089	-0.0090	-0.0100	0.0079	0.0076	0.0136

Notes: Elasticities were computed according to the standard formula: $\eta_{ikrht} = \frac{\partial q_{ikrt}}{\partial p_{kht}} \frac{p_{kht}}{q_{ikrt}}$, using numerical methods to approximate partial derivatives due to the discontinuity of predicted choice probabilities. Row titles indicate the product which price is changing. Column headers indicate the sensitivity of the demand for a particular product to a 20% price change.

2.6 Robustness checks

A first concern when using simulated methods is whether the results are sensitive to changes in starting values. To be sure that our estimates were robust to changes in the vector of initial parameters, θ_0, we performed the whole estimation process described in Subsection 3.4.2 using ten different sets of pseudorandom draws from a normal, as starting values. We obtained similar estimates at each iteration which may as well be interpreted as an indicator of convergence. The final results, which are shown in Table 2.1 are those corresponding to the minimum value of the objective function out of ten available.

We also conducted a sample selection check. The final sample used for the estimates presented previously was selected by restricting attention to those households purchasing the three products considered here in a given week, consistent with our assumption of inelastic demand for a unit of each product. We therefore dropped households not fulfilling this condition. To find out if our results are robust, we use an alternative sample with tighter restrictions on the selection of the households, namely, if we observed a zip code with at least one household not purchasing the three products in a given week, we drop the entire local market. We are left with 1,027 observations corresponding to purchases made by 541 households. We use the same estimation method and instruments as for our final results. However, due to the few observations in this sample, we do not include product fixed-effects. Results are similar in the direction and statistical significance of estimates, except for the distance that become non significant with the use of IVs (see Table

2.1). In this sense, the results do not seem to be driven by sample selection. However,concerning the magnitude of estimates we have a remarkable difference. This might be driven by the fact that there is much less variation in the new sample and the omission of product dummies (that capture consumers' valuation for product characteristics). In particular, the average shopping costs cutoffs, expressed in euros, are smaller as compared to those in Table 2.1. This is due, in part, to a larger estimate of the price coefficient. Nevertheless, their relative position remains similar and lead to the same conclusions as those derived before.

Table 2.1: Results based on an alternative sample[a]

Variable	(1)	(2)
Price (€/basket[b])	-3.75***	-3.78***
	(0.61)	(0.25)
Distance (km)	-13.48**	-5.71
	(5.34)	(3.88)
Mean Shopping Costs	0.80***	0.41***
	(0.23)	(0.08)
Time dummies	Yes	Yes
Instruments		Yes
Av. shopping costs cutoffs (in €)		
One-Two stops ($\hat{s}^2$)		0.26
One-Three stops ($\hat{\Delta}^{31}/2$)		0.13
Two-Three stops ($\hat{s}^3$)		0.01

Notes: [a]Based on 1,027 observations of purchases made by 541 households. Bootstrap standard errors are in parenthesis.
[b] A basket contains a serving of each of the considered products: a slice of bread (28g), a bowl of cereal (35g) and one yogurt (125g).
*,**,*** are significant at 10, 5 and 1% confidence levels.

2.7 Concluding remarks

Theory has shown that in the presence of shopping costs, i.e. real or perceived costs of dealing with a supplier, policy conclusions might change dramatically. In particular, some pro-competitive practices, such as head-to-head competition with homogeneous product lines (Klemperer, 1992) or the introduction of a new product variety (Klemperer and Padilla, 1997), can hurt consumers and motivate policy intervention. On the other hand, some seemingly anti-competitive practices, such as below-cost pricing, can be welfare enhancing and should not be banned (Chen and Rey, 2013).

From an empirical point of view, this motivates many important questions that remain unanswered. First, is it possible to quantify shopping costs from consumers' observed shopping behavior? Second, will accounting for shopping costs in an empirical model of multiproduct demand lead to a better understanding of consumer heterogeneity in shopping patterns? Finally, to what extent the inclusion of shopping costs would be crucial for policy analysis? This paper presents and then estimates a model of multiproduct demand for groceries in which customers, that differ in shopping costs, can choose between sourcing one or multiple retailers in the same shopping period. This framework allows us to retrieve the distribution of shopping costs.

We quantify the total shopping cost in 18.7 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 1.53 € and the total transport cost of 17.1 € per trip to a given store. Moreover, we are able to compute the transport and total costs of shopping by store format. Transport and total costs of shopping are increasing in the size of the stores, on average, as smaller formats are closer to downtowns. The largest total shopping costs, 24.7 €, are incurred by consumers who source big-box stores, because they are farther away from downtown. Sourcing a supermarket or a hard-discounter implies total costs of shopping of 14.3 € and 13.4 € per trip, respectively. Finally, the costs of sourcing a convenience store, 4.8 € per trip, are the lowest provided they are located in downtown. We find that individuals who source more than two suppliers in a week have zero (even negative) shopping costs. This rationalizes the low proportion of individuals making three and more stops in the same week observed in the data. This might be an indicator that those households actually visiting more than two separate stores a week should have a strong preference for shopping. In fact, The predicted proportions of shoppers by number of stops are 90.1% of one-stop shoppers, 9.7% of two-stop shoppers and only 0.26% do three-stop shopping.

There are several avenues for further research that can be empirically addressed using our framework. A first avenue is related to below-cost pricing. According to the OECD (2005), laws preventing resale below-cost (RBC) and claiming to protect high-price, low-volume stores from large competitors who can afford lower prices might be introducing unnecessary constraints. Evidence from countries without RBC laws shows that smaller competitors need not be pushed out of the market if they are not protected. Chen and Rey (2012, 2013) show that in the presence of shopping costs, loss-leading strategies and cross subsidies are not predatory, and the latter might even be welfare enhancing. Empirical evidence showing what would happen if RBC laws are eliminated would help in this debate.

A second avenue concerns the implications of product delisting. In recent years, a considerably concentrated retail sector has brought the attention on the possible consequences of retailer buyer power on upstream firms. A retailer can, for example, stop carrying a product to punish a particular supplier for not agreeing on her requests. It might as well use delisting as a threat, so that she can get better terms of trade. How will demand react to the delisting of a product? Will consumers substitute brands in the same store or will decide to source an alternative store? What is the role of shopping costs in this decision? These are questions to be addressed.

To carry out such policy analyses, a more comprehensive and flexible framework allowing for multiple brands per category in each supermarket as well as the possibility of elastic choices by consumers (bundles containing zero, one or multiple products as opposed to a fixed number) is needed. Our structural model can be readily extended to cover such changes. However, the empirical implementation of such a flexible framework is challenging and computationally burdensome, in particular because each product added to the problem increases the dimension of the choice set exponentially. This and other related issues are part of our current research efforts that we hope will allow us to come up with a solution in the near future.

Finally, theoretical and empirical analyses should be done on retailers' motivations to raise consumers shopping costs and the consequences of such strategies for competition and consumer welfare. One-stop shopping make more powerful retailers. Klemprer (1992) predicts that if consumers are not interested to source multiple retailers, prices will tend to be higher. It might be the case that consumers face such high shopping costs that they are not able to do multistop shopping even if they would like to. Retailers might use their market power to raise customers shopping costs by making the shopping experience more tiring or complicated, so that their share of one-stop shoppers increases.

Appendix

2.A The utility function of a n-stop shopper

We can give a general expression for the optimal decision rule of a n-stop shopper, $n \in N = \{1, \cdots, R_i\}$, $R_i \leqslant R$, being R the total number of grocery stores in the market, as follows. Assume a n-stop shopper compares bundles of the desired products from all the possible combinations of n stores. Denote each of these combinations by $j \in \{1, \cdots, J_i^n\}$, where according to combinatorics theory, the total number of combinations of R elements taken n at a time is given by $J_i^n = R_i!/n!(R_i - n)!$ Consumer i will choose the mix j of n stores such that

$$\sum_{k=1}^{K_i} \max\{v_{ikrt}\}_{r\in j} \geqslant \sum_{k=1}^{K_i} \max\{v_{ikr't}\}_{r'\in l} \; \forall \; l = 1, \cdots, J_i$$

For instance, in a context with $R = 3$ stores, a one-stop shopper $n = 1$ will pick the best combination of one store out of $J_i^1 = 3$ possible {A},{B},{C}, and pick the best mix such that it yields the largest overall value of the desired bundle. Similarly, a two-stop shopper, $n = 2$, will compare all $J_i^2 = 3$ possible combinations of two stores ({A,B},{B,C},{A,C}) and pick the best according to the rule above. For a three-stop shopper, $n = 3$, the number of combinations of three stores taken three at a time is $J_i^3 = 1$, i.e. {A,B,C} which explains why he is not maximizing over several subsets of stores in equation (2.3.3).

2.B Cases for extra utilities ordering

As stated in Section 3.3, we can derive critical cutoff points on the shopping costs distribution from equations (2.3.4), (2.3.5) and (2.3.6) as functions of δ_{it}^2, δ_{it}^3 and $\Delta_{it}^3/2$. As these numbers represent utilities for different, say, products, their ordering can vary from a consumer to another. Therefore, we need to establish what the cutoffs would be in a case by case analysis.

From three objects, we can have six possible orderings:

$$\begin{aligned} &(C1) \quad \delta_{it}^2 > \frac{\Delta_{it}^3}{2} > \delta_{it}^3, && (C2) \quad \delta_{it}^3 > \frac{\Delta_{it}^3}{2} > \delta_{it}^2, \\ &(C3) \quad \frac{\Delta_{it}^3}{2} > \delta_{it}^3 > \delta_{it}^2, && (C4) \quad \frac{\Delta_{it}^3}{2} > \delta_{it}^2 > \delta_{it}^3, \\ &(C5) \quad \delta_{it}^3 > \delta_{it}^2 > \frac{\Delta_{it}^3}{2}, && (C6) \quad \delta_{it}^2 > \delta_{it}^3 > \frac{\Delta_{it}^3}{2}, \end{aligned}$$

From the six cases above, only $(C1)$ survives, the remaining are contradictory. To see why, notice that the incremental utility of sourcing two additional stores, $\Delta_{it}^3 := v_{it}^3 - v_{it}^1$, can be written as the sum of the two marginal utilities of going from one to two stores and from two to three. This is: $\Delta_{it}^3 = \delta_{it}^2 + \delta_{it}^3$. Therefore, if we assume, for instance, that $\frac{\Delta_{it}^3}{2} > \delta_{it}^3$ as in in $(C3)$, then

$$\frac{v_{it}^3 - v_{it}^2}{2} + \frac{v_{it}^2 - v_{it}^1}{2} > v_{it}^2 - v_{it}^1 \equiv \delta_{it}^3$$

which after some manipulations leads to $\delta_{it}^2 > \delta_{it}^3$, i.e. a contradiction. In a similar fashion, the proofs for the other cases follow.

2.C Data manipulation for structural estimation

Three products are taken into the analysis, ready-to-eat breakfast cereals, yogurt and pre-packaged bread, which are among the most purchased products by french households. It is often the case that people do not only buy one brand, or even one unit of the same brand at a time but several varieties to have different choices at home (different flavors, fruit contents, etc.). However, following Nevo (2001), we claim that an individual normally consumes one yogurt (125 grams per portion), one serving of cereal (35 grams per portion), and one serving of bread (28 grams per portion) at a time, so that the choice is discrete in this sense. Of course there could be cases in which some people consume more than one brand, or serving, at a time. Although we believe this is not the general case, the assumption can be seen as an approximation to the real demand problem.

In our scanner data we do not observe prices but total expenditure and total quantity purchased for each product and store sourced by each household. Consequently, a price variable was created in the following way: first, we compute the sum of expenditures over local markets(defined by zip codes), month, and stores and number of servings of each product purchased by each consumer. Second, we divided the total expenditure on a given product-store made by all consumers living in the same zip code in a month by the the total number of servings to obtain a common unit price. If the information to compute a unit price is missing, we replace it with the average across local markets within the same period. By constructing our price variable this way, we are assuming that consumers have rational expectations. Due to data limitations, we do not account for manufacturers' nor stores' promotional activities or sales of any kind.

Last, to compute distances between the store and the household location we follow Dubois and Jódar-Rosell (2010). Data on stores location was obtained from *LSA/Atlas de la Distribution 2005*, which contains information on most french stores involved in groceries distribution. The information was merged with the household data using the name of the store, the zip code of the consumer's residence and the surface of the outlet. For each store, we find the closest outlet to the consumer thanks to zip codes and geographical data. Only one outlet per store chain was included in this set.

Chapter 3

The role of nonlinear pricing and resale price maintenance on nominal price stability

Abstract: This paper empirically examines the role of nonlinear contracts between manufacturers and retail stores, and Resale Price Maintenance (RPM) on nominal price stability. It is widely accepted in the literature that the incomplete transmission of costs shocks into retail prices is explained by the existence of markup adjustment and price adjustment costs. The vertical conduct of the industry and the existence of vertical restraints such as RPM might introduce further price stickiness or reinforce it. I present a structural model of vertical relations between manufacturers and retailers allowing for nonlinear contracts and vertical restraints, and accounting explicitly for retail price rigidity by including fixed costs of price adjustment in retailer's profit function. Using micro data on sales of ready-to-eat breakfast cereals from a large supermarket chain in Chicago, I estimate demand, retrieve upstream and downstream markups, and compute bounds of retail price adjustment costs. Results show that the total costs the retailer bears for adjusting prices of its products in a year lie between 1.6% and 3% of its total revenue, on average.

JEL Codes: L11, L13, L42.

Keywords: Nonlinear contracts, vertical restraints, resale price maintenance, costs shocks, incomplete pass-through, nominal price stability, menu costs, adjustment costs, random-coefficients Logit.

3.1 Introduction

Nominal price rigidity and its effects on monetary policy are a central concern of Macroeconomics. A widely held view is that aggregate price inertia is determined by how responsive individual goods prices are to cost or exchange rate shocks (Midrigan, 2011; Kehoe and Midrigan, 2012). Actually, evidence shows that retail prices are not very responsive to changes in nominal costs and exchange rates. This is the so-called incomplete pass-through of costs shocks to nominal retail prices. Does the structure of the industry matters when it comes to explain individual goods price stickiness? Engel (2002) points out three sources of such incomplete transmission that are related to the industry structure and firms' strategic behavior: the existence of local costs, markup adjustment either by retailers or manufacturers or both, and nominal price rigidity. A growing empirical Industrial Organization (IO) literature has made important advances in understanding incomplete pass-through by looking at how vertical relations and vertical restraints such as Resale Price Maintenance (RPM), a practice through which manufacturers determine prices retailers charge to consumers, shape markup adjustment. Less attention has been put on price rigidity. We know from microeconomic theory that RPM makes prices less responsive to costs shocks (Jullien and Rey, 2007). The objective of this paper is to empirically examine the role of nonlinear contracts between manufacturers and retail stores, and RPM on nominal price stability.

There is a considerable number of theoretical and empirical contributions to the study of the incomplete transmission of costs shocks to nominal prices and aggregate price inertia. Empirical research motivated by macroeconomic theory has mainly focused on providing evidence on the importance of price rigidity (Midrigan, 2011; Eichenbaum et al., 2011; Levy, et al. 2010; Kehoe and Midrigan, 2012), the frequency of price changes and the duration of nominal prices, and the sources of such rigidities using reduced-form methods (Levy et al., 1997; Dutta et al., 1999; Peltzman, 2000; Chevalier et al., 2003; Goldberg and Campa, 2006; Leibtag et al., 2007; and Levy et al., 2010).

In the empirical IO literature we find contributions covering a variety of methods, perspectives of the incomplete transmission problem and applications to particular industries. These contributions include papers providing evidence on how the vertical structure of the industry affects the degree in which costs shocks pass-through to nominal prices (Hellerstein and Villas-Boas, 2010, Bonnet et al., 2013). Other articles analyze the sources of this incomplete pass through. Three main sources have been accounted for: local non-traded costs (Goldberg and Verboven, 2001; Hellerstein, 2008; Nakamura and Zerom, 2010); strategic markup adjustment (Bettendorf and Verboven, 2000, Goldberg and Verboven, 2001, Nakamura and Zerom, 2010, Hellerstein and Villas-Boas, 2010, Goldberg and Hellerstein, 2013, and Bonnet et al., 2013); and price rigidity in the form of costs of price adjustment (Slade, 1998, Nakamura and Zerom, 2010, and Goldberg and Hellerstein, 2013). Most papers use structural models of vertical relationships between manufacturers and retailers to show that the type of relationships between upstream and downstream firms and the presence of vertical restraints play an important role in the degree of price responses to costs and/or demand shocks.

I set out a modeling framework that closely relates to two of those contributions. Goldberg and Hellerstein (2013) address the question of the incomplete transmission of exchange rate shocks to local currency prices of imported beer in the U.S. They set out a structural model in which linear tariffs characterize the industry's vertical conduct. Unless previous research, they control explicitly for price rigidity by including fixed costs of price adjustment in the profit functions of both manufacturers and retailer in a static framework. These costs capture everything that prevents a firm from adjusting the price in a period (menu costs, opportunity costs of time and effort to find a new optimal price, advertisement costs, etc.), which helps rationalizing why we often observe a given retail price that remains constant for several weeks and/or that is set again after a temporary price reduction (the so-called *regular price* in Macroeconomics literature).

On the other hand, Bonnet et al. (2013) are the first to empirically investigate the role of nonlinear pricing and RPM on incomplete pass-through of costs into retail prices, focusing on how the vertical structure of the industry affects strategic markup adjustment. They analyze the market for coffee in Germany in which retail prices are positively correlated with raw coffee prices but vary significantly less. They find that when manufacturers can implement two-part tariffs contracts with RPM, the share of a cost shock that is passed-through to retail prices is larger than in the presence of other types of contracts, because RPM restricts retailers' ability to make strategic markup adjustment.

My paper may be regarded as a combination of Bonnet et al. (2013) and Goldberg and Hellerstein (2013). As in Bonnet et al., I specify the supply side according to several distinct models of vertical relationships, namely, linear pricing, and two-part tariffs contracts with and without RPM. As in Goldberg and Hellerstein, I explicitly account for price rigidity by including price adjustment costs to the optimization problem of the retailer and use the model and estimates to quantify bounds on retail price adjustment costs. My focus is, however, on a totally different sector: the ready-to-eat (RTE) breakfast cereal industry.

This is a particularly suitable industry to study vertical relations and vertical restraints. It is characterized by high concentration, high price-cost margins, very intensive non-price competition through aggressive advertising campaigns, product proliferation due to rapid introduction of new brands, and substantial coupon issuing by manufacturers. On top of that, price competition is considerably less intense. RTE breakfast cereals are highly consumed by U.S. households with a penetration rate near to 100% for cold cereals and 65% for hot cereals and sales of roughly $10 billion in each category in 2013. Furthermore, cereal prices appear to be quite rigid, even though input prices are not (see Figure 3.1).

This work adds to the literature in several ways. First, it fills a gap by setting out a model that allows nonlinear contracts and vertical restraints interact with the price adjustment problem that changes the way optimal prices are set as compared to the standard static profit maximization problem. Second, it shows how the vertical conduct of the industry help rationalizing the observed price rigidity. Last, it adds to the literature on the U.S. RTE cereal industry, by shedding new light on how manufacturers and retailers relate and how pricing decisions are made in the presence of adjustment costs.

My empirical strategy relies on the consistent estimation of demand, which I specify as a random-coefficients Logit model. I set out three structural models of supply (linear pricing, simple two-part tariffs, and two-part tariffs with RPM) in a context of static Nash-Bertrand oligopolistic competition with several competing manufacturers and a single retailer that carries all products. The retailer faces fixed costs of repricing whenever it decides to adjust the price of a product. At each period, the retailer weighs the costs and the benefits of changing the price of each product and makes one of two decisions: if benefits exceed costs, it sets a new price that maximizes current period profit; otherwise, it keeps the same price from previous period which implies a deviation from first order conditions of the static profit maximization problem.

Using data from *Dominik's Finer Foods*, a large supermarket chain in Chicago, that contains among other things information on weekly prices and quantities sold at the universal product code (UPC) level for 224 weeks, I consistently estimate the demand parameters which are identified without the need of the supply side thanks to the panel structure of the data. Next, I recover retail as well as wholesale margins and marginal costs according to each industry conduct specified. With all these elements in hand, I use the structural model to compute upper and lower bounds of adjustment costs.

I find that the linear pricing specification of the supply side gives biased results for price adjustment costs relative to two-part tariffs with RPM. In fact, under linear pricing I obtain that the retail chain is willing to change the price of a product if it obtains, on average, an extra profit of at least US \$109, whereas under two-part tariffs with RPM this amount is US \$98.84. On the other hand, I obtain that adjustment costs are bounded above by US \$447 on average under either supply conduct. Simple two-part tariffs (i.e. without RPM) give similar results to those of linear pricing. To have an idea of the relative importance of these magnitudes, I compute the share of the sum of each lower and upper bounds of adjustment costs on retailer's total revenue for the entire period considered here (224 weeks). I find that the share of adjustment costs on total revenue is 7.27% for lower bounds and 10.12% for upper bounds in the linear tariffs case, and to 6.5% and 10.14% respectively, in the two-part tariffs case.

There is, however, a big caveat related to the the fact that I model the market with a single retailer as if it was a local monopolist. When there is no downstream competition, simple two-part tariffs are sufficient to solve the double marginalization problem and attain monopoly profits. In such a context, if manufacturers can use RPM they should obtain no better result than that they get with simple two-part tariffs. RPM does make a big difference when there is competition among retailers as simple two-part tariffs no longer suffice to maintain monopoly profits (Rey and Vergé, 2010). With the appropriate data set containing such detailed information on multiple retailers, I would be able to offer an empirical analysis of multiple common agency. I expect to do this in a future version of this paper. In the meantime, the results I present here are intuitive, consistent with theory and give an idea of the importance of the vertical conduct of the industry in the study of an economic problem such as price stickiness.

The rest of the paper is organized as follows. Section 3.2 gives an overview of the data used in the paper and presents a preliminary analysis of price rigidity in the RTE cereal industry. Section 3.3 outlines the structural demand model as well as the supply models of vertical relationships between manufacturers and retailers. Section 3.4 discusses details of the empirical implementation of the model. Section 3.5 presents the results. Finally, Section 3.6 concludes and discusses directions for future research.

3.2 The market, data and reduced-form analysis

This Section aims at giving an overview of the data, the cereal industry in the United States and some preliminary results based on descriptive statistics and reduced-form regressions.

3.2.1 Overview

The data I use in this paper comes from *Dominick's Finer Foods* (DFF), the second largest supermarket chain in the Chicago metropolitan area. *Dominick's* database is provided by the Kitls Center for marketing at the University of Chicago Booth School of Business and is publicly available.[1] It is scanner data reported by each store in the sample at the Universal Product Code (UPC) level for 29 categories of packaged products in 93 stores of the chain for 400 weeks between September 1989 and May 1997.

[1] Go to:
http://research.chicagobooth.edu/kilts/marketing-databases/dominicks/dataset.

The database contains weekly information on retail price, quantity sold, promotional activity and the percent gross margin the store makes on each sale of a UPC to consumers. The latter variable can be used to compute the average acquisition cost (AAC) of each brand, which gives information on wholesale prices.[2] During the data collection period, DFF set prices according to four 'price zones' (high-, medium- and low-price zones, and 'Club-fighter' zone), which are defined by geographic location and/or nearby competitors. However, as DFF pricing policy is chainwide, prices across stores are highly correlated and follow a similar pattern.[3]

A particularly appealing industry to study vertical relations and vertical restraints is that of the ready-to-eat (RTE) breakfast cereal. RTE breakfast cereals are highly consumed by U.S. households with a penetration rate near to 100% for cold cereals and 65% for hot cereals and sales of roughly $10 billion in each category in 2013. The particular characteristics of this industry have been widely documented: high concentration, high price-cost margins, very intensive non-price competition through aggressive advertising campaigns, product proliferation, due to rapid introduction of new brands, and coupon issuing by manufacturers. On top of that, price competition is considerably less intense.[4]

Interestingly, prices in this industry appear to be quite rigid, even though prices of commodities such as wheat, corn and sugar, three main ingredients of RTE cereals, vary quite often, which makes it also suitable to investigate the sources of incomplete pass-through (see Figure 3.1). In fact, a first look at price series in the database shows that one can easily identify what macroeconomics literature calls *regular* price, i.e. a price that remains unchanged during several weeks or that after a temporary change (such as a sales price), returns to the same level as before. Figure 3.1 displays retail price and AAC series for two different brands of RTE cereals from May 1990 to September 1994 as well as nominal prices of three commodities commonly used as inputs of RTE cereals.

The common pattern in the bottom panel is retail prices that remain at the same *regular* level for several weeks, even though we observe some temporary reductions in between, and jumps to another level that becomes the new regular price. By contrast, the AACs do not show the same pattern, although we can observe some stability, they vary more frequently. This may suggest that in this industry manufacturers and retailers may be using nonlinear contracts instead of linear tariffs (double marginalization). The presence of RPM seems plausible as well. In fact, evidence from other industries shows that when the vertical conduct of the industry is based

[2] It is a very rough proxy of wholesale prices though, as long as it does not contain information on replacement costs or the last transaction price. For a detailed description of the database and a discussion on the computation of average acquisition costs, see Peltzman (2000).

[3] For zones 1,2 and 3 the main rival is 'Jewel', the largest store chain in Chicago. As for the club-fighter price zone, the main rival is 'Club foods', a discount chain (see Peltzman, 2000).

[4] For a detailed description of this industry, see for example Schmalensee (1978), Nevo (2000b,2001).

on double marginalization, changes in the wholesale price of a product generally lead to changes in the retail price in the same direction (see Goldberg and Hellerstein, 2013). In the case of cereals, however, we observe that not only retail prices are not responding very frequently to changes in AACs, but also in some cases they lie below this cost. Although this is suggestive of something else might be driving vertical relationships, we cannot conclude anything by simply looking at the data. This is the motivation to use structural methods.

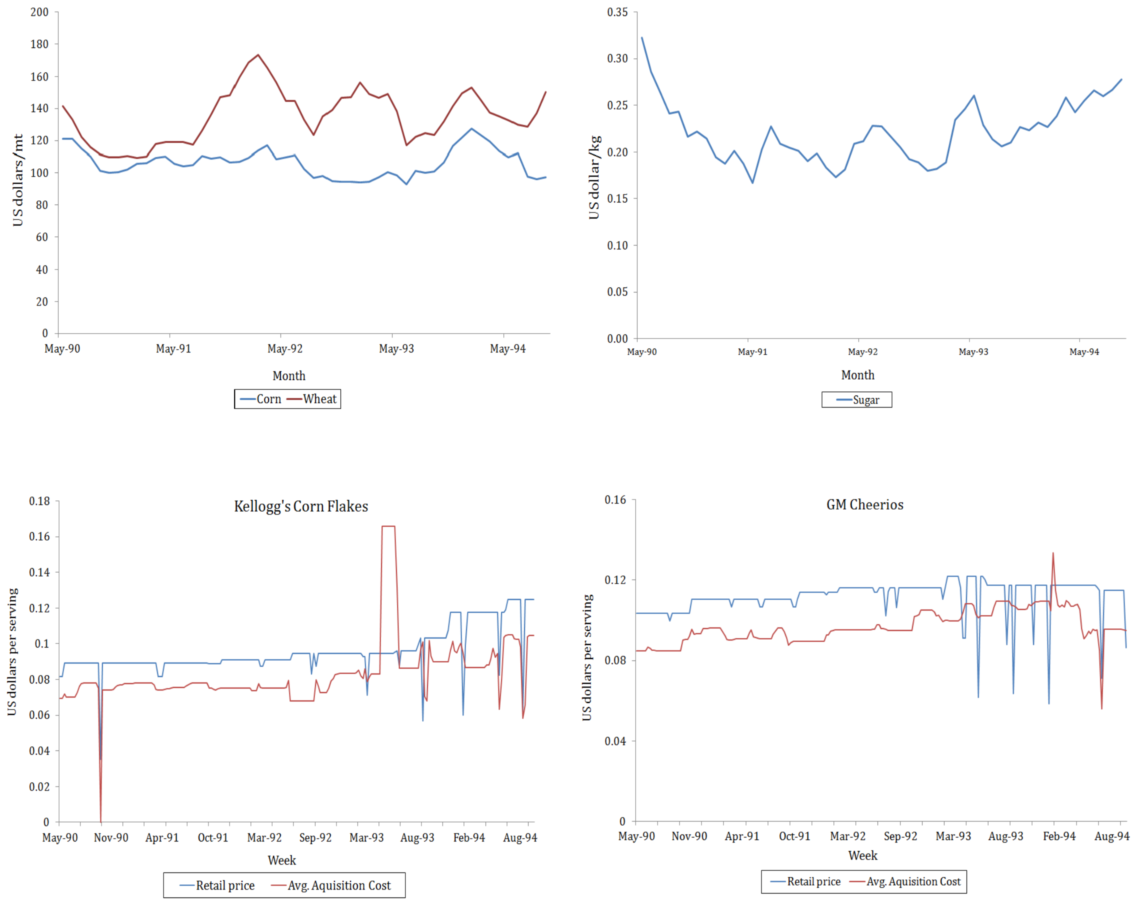


Figure 3.1: Monthly world prices of Corn and Wheat (Top-left), and Sugar (Top-right). *Source: World Bank*. Weekly retail price and average acquisition cost for Corn Flakes (bottom-left), and Cheerios (bottom-right). Prices are from an arbitrary store of the Medium price tier. *Source:* Dominick's database.

Table 3.1 reports summary statistics on prices, frequency of price changes and number of weeks a given price remains constant. These numbers confirm that retail prices are quite stable as compared to AACs, that appear to change more frequently in the same period. In fact, a dummy variable taking on 1 if the price observed in a particular week was different from the price of the same product in the previous week, suggests that the retail price of a product varies 21% of times in 224 weeks, while the AAC of the same product varies 57% of times in the same period. Moreover, the average duration of a given retail price is nearly double of that of a given AAC: 15.5 weeks against 8.7 weeks.

Table 3.1: Summary statistics for retail price and Avg. Acquisition cost (AAC)

Variable	Mean	Median	Std. Dev.	Min	Max
Retail price (cents/serving)	10.65	10.88	2.49	2.29	16.93
AAC (cents/serving)	9.08	9.25	2.15	0	21.79
Dummy for retail price change (=1 if Yes)	0.21	0	0.41	0	1
Dummy for AAC price change (=1 if Yes)	0.57	1	0.50	0	1
Duration of given retail price (No. Weeks)	15.48	13	10.50	1	58
Duration of a givenAAC price (No. Weeks)	8.69	6	7.38	1	38

Source: Dominick's database.

3.2.2 Reduced-form analysis

To have an idea of the magnitude of the impact of costs shocks on retail prices of RTE cereals, I perform three linear regressions with the log of retail price as the dependent variable on observable costs shifters. Table 3.2 displays the results. The first regression (column 1) includes an employment cost index for total compensation of workers in goods producing industries in the United States. The second regression (column 2) has in addition the logs of nominal prices of key inputs for cereal production, namely, wheat, corn and oil. Finally, a third regression (column 3) substitutes all previous covariates by the log of the average acquisition cost of each product reported by retail chain. All regressions include product, time and price zone dummy variables to account for observed and unobserved product characteristics, and time and zone fixed effects.

Interestingly, all estimated elasticities are very low indicating that only a small proportion of costs shocks are passed-through to consumer prices. The specification given in column 2, for instance, predicts that a 10% percent increase in employment costs leads to a 1.3% increase in retail prices. Similarly, a 10% increase in the price of corn leads to a 1.1% rise in cereal prices. Even lower elasticities are obtained for the price of wheat and oil, although they are not statistically significant. Column 3 shows that the degree of retail price responses to changes in AACs, which contains information on all input costs, is not very high either. An increase in the AAC of a product leads to an increase in the retail price of that product in 2.0%. This preliminary result is a clear evidence of the incomplete response of cereal prices to upstream costs and reinforces the question this paper tries to address. Not only retail prices are not fully responding to changes in input prices but also show to be rigid to changes in wholesale costs, even if these are a very important component of total costs of distribution. In a similar analysis, Goldberg and Hellerstein (2013) find that retail prices fully respond to changes in wholesale prices and that infrequent price adjustment was driven by rigid wholesale prices. My results then suggest that a complex vertical structure of the industry may be playing an important role in this incomplete transmission of costs and wholesale price movements to retail prices.

Table 3.2: Results from linear regressions (variables in logs)

Variable	(1)	(2)	(3)
Labor cost index	0.084**	0.132***	—
	(0.039)	(0.048)	
AAC	—	—	0.204***
			(0.044)
Wheat	—	0.009	—
		(0.033)	
Corn	—	0.114*	—
		(0.068)	
Oil	—	0.004	—
		(0.016)	
Constant	-2.630***	-3.374***	-2.352***
	(0.168)	(0.532)	(0.136)
R^2	0.9081	0.908	0.922

Notes: Based on 14,784 observations. All regressions include brand, week and price zone fixed effects.
*,** ,***Significant at 10%, 5% and 1% level, respectively.

3.3 The model

In this Section I set out structural models of demand and supply, with a special focus on characterizing vertical relationships between manufacturers and a single retailer in the presence of retail costs of price adjustment. Then I derive expressions for adjustment costs bounds under three alternative specifications for the supply side.

3.3.1 Supply Models

The approach presented here is similar to that of previous literature that investigates the sources of price stability from a vertical relations perspective. I set out three alternative models of vertical relations that can potentially fit the case of the industry under study, namely, linear tariffs, two-tariffs with RPM, and simple two-part tariffs (i.e. without RPM). This has at least two advantages: first, I can compare results across models and form a prior about the role of nonlinear contracts on adjustment costs, which should be evaluated with the help of counterfactual simulations; and second, I avoid imposing a particular structure arbitrarily to the data.[5] To account for price rigidities on retail prices, I follow closely Goldberg and Hellerstein (2013) and add a fixed cost of price adjustment to retailer's profit function. The derivation of lower and upper bounds for these adjustment costs follows a revealed-preferences approach, according to which the retailer optimally decides whether or not to adjust the price at each period by comparing profits she would obtain under each alternative scenario.

I consider a model of single common agency, i.e. there are several competing upstream firms, indexed by $f = \{1, \ldots, N\}$, distributing their products through a common downstream retailer, r. Accordingly, the retailer carries all products, which are indexed by $j = \{1, \ldots, J\}$. Markets are defined as a 'price zone'-week combination and denoted $t = 1, \ldots, T$.

As previously stated, the structural models presented in this paper account for retail price rigidity. The argument is rather positive and comes from what I observe in the data: whereas retail prices for all products considered here show a similar pattern according to which a regular price (a price remaining constant for several weeks) can be easily identified, the AACs seem to change more often and need not keep the same proportion with respect to retail prices. Recall that some descriptive statistics point out that while retail prices change roughly 21% of times, the AACs change about 57% of times in 224 weeks. In terms of modeling, this means that unlike the retailer, each manufacturer sets prices that satisfy first order conditions of the static profit maximization problem.[6]

The retail firm bears a fixed cost of repricing whenever she decides to adjust retail prices with respect to the previous period levels. These costs are denoted A^r_{jt} in case the retailer wishes to adjust the retail price of product j in the current period. As it will become apparent below, these costs capture all the remaining variation that is not accounted for by included covariates. In economic terms, this is interpreted as all factors that make a firm refrain from adjusting the price of a product with respect to its previous-period level and potentially deviate from the optimum implied by static profit maximization. Consequently, these costs may include, among other things, menu costs, management costs, time and effort costs, and price advertising costs (Goldberg and Hellerstein, 2013).

[5] A preferred specification can be selected using nonnested tests of model selection developed by Rivers and Vuong (2002) and applied to vertical relations by Bonnet and Dubois (2010,2015). In the next iteration of the paper I will use such tests.

[6] This does not mean that manufacturers do not bear costs of repricing, but these should be certainly lower than those faced by the retailer.

3.3.1.1 Linear tariffs

In a context of linear tariffs, manufacturers set prices first and the retailer follows by setting retail prices taking wholesale prices as given. This form of vertical interactions leads to the well known double marginalization result. From a horizontal perspective, manufacturers act as oligopolists competing against rivals à la Nash-Bertrand. In the following, I first present the problem of the retailer and then the problem of manufacturers, following a backward induction reasoning.

■ **Retailer problem**

Suppose retailer r carries all J products existing in the market. Her profit function at t writes as

$$\Pi_t^r = \sum_{j=1}^{J} \left[(p_{jt} - w_{jt} - c_{jt}) s_{jt}(p_t) M - \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A_{jt}^r \right] \tag{3.3.1}$$

where p_{jt} is the retail price of product j at t, w_{jt} the wholesale price of product j paid by retailer r at t, c_{jt} is the constant marginal cost of distribution of product j at t, $s_{jt}(p_t)$ is the market share of product j at t that depends on the vector of prices of all products in the market, and M is the size of the market. Notice that the term A_{jt}^r is preceded by the indicator function $\mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}}$ which takes on one if the current price of j is different from the previous period price, and zero otherwise. This means that r bears this cost only if she decides to change the price of product j at period t.

In the presence of price adjustment costs, setting a new price every period is costly and might eventually be unprofitable as compared to the profit the retailer would make by leaving the price unchanged. An optimal behavior by the retailer implies weighing benefits of adjusting the price of each product in the current period with the costs. If extra profit of setting a new price exceeds adjustment costs, it is optimal to do so. Otherwise, it is optimal to leave the price of that product constant. In this sense, optimal price setting with positive fixed costs of repricing need not coincide with the standard static profit maximization behavior according to which optimal prices must always satisfy the current period first order conditions (FOCs).

The optimal price adjustment problem leads to two possible cases: either the price of a product changes from previous period (Case 1) or the price remains constant from previous period (Case 2). I describe the two in detail in what follows.

Case 1: The price changes from the previous period ($p_{jt} \neq p_{jt-1}$). Retailer r will be willing to change the price of product j at time t if the total profit net of repricing costs exceeds the profit she would have made by leaving the price constant, i.e. if for all $k \neq j$

$$(p_{jt} - w_{jt} - c_{jt})s_{jt}(p_t)M + \sum_{k \neq j}(p_{kt} - w_{kt} - c_{kt})s_{kt}(p_t)M - A^r_{jt}$$
$$\geqslant (p_{jt-1} - w_{jt} - c_{jt})s^c_{jt}(p_{jt-1}, p_{-jt})M + \sum_{k \neq j}(p_{kt} - w_{kt} - c_{kt})s^c_{kt}(p_{jt-1}, p_{-jt})M, \tag{3.3.2}$$

where $s^c_{.t}(p_{jt-1}, p_{-jt})$ denotes market shares in the counterfactual scenario. The retailer determines the price of product j by maximizing (3.3.1). Assuming the existence of a pure-strategy Nash equilibrium in retail prices and that these prices are strictly positive, first-order conditions (FOCs) of the problem in (3.3.1) are as follows:

$$s_{jt}(p_t) + \sum_{k=1}^{J}(p_{kt} - w_{kt} - c_{kt})\frac{\partial s_{kt}}{\partial p_{jt}} = 0. \tag{3.3.3}$$

The FOCs yield a system of equations, one for each product j in r's product range. To write this system in matrix notation, define S_p as a $J \times J$ matrix containing market shares responses to changes in retail prices, with entry $S_p(j,i) = \frac{\partial s_j}{\partial p_i}$ for $j, i \in \{1, \ldots, J\}$. Further, let γ_t denote the vector of price-cost margins which from the FOCs is given by[7]

$$\gamma_t \equiv p_t - w_t - c_t = -S_p^{-1}s(p_t) \tag{3.3.4}$$

Using (3.3.2) and rearranging terms, an upper bound for the adjustment costs of product j is given by:

$$\begin{aligned} A^r_{jt} \leqslant \overline{A^r_{jt}} = \Big[&(p_{jt} - w_{jt} - c_{jt})s_{jt}(p_t) - (p_{jt-1} - w_{jt} - c_{jt})s^c_{jt}(p_{jt-1}, p_{-jt}) \\ &+ \sum_{k \neq j}(p_{kt} - w_{kt} - c_{kt})(s_{kt}(p_t) - s^c_{kt}(p_{jt-1}, p_{-jt}))\Big]M, \end{aligned} \tag{3.3.5}$$

Case 2: The price does not change from the previous period ($p_{jt} = p_{jt-1}$). Retailer r may find it optimal to leave the price of product j unchanged from previous period, if the adjustment costs are high enough so that it is more profitable to not change the price, even if this may imply that the price does not satisfy the

[7] In a context of multiple retailers carrying differentiated products, retail margins write $I_r\gamma_t = -(I_rS_pI_r)^{-1}I_rs(p_t)$ for all $r = 1, \ldots, R$, with I_r being retailer r's ownership matrix, i.e. a diagonal matrix of order J with jth entry equal to 1 if product j is in r's product range and zero otherwise. With a single retailer carrying all products in the market, the ownership matrix is the identity of order J.

current period FOCs, i.e. if for all $k \neq j$

$$
\begin{aligned}
&(p_{jt-1} - w_{jt} - c_{jt}) s_{jt}(p_{jt-1}, p_{-jt}) M + \sum_{k \neq j} (p_{kt} - w_{kt} - c_{kt}) s_{kt}(p_t) M \\
&\geqslant (p^c_{jt} - w_{jt} - c_{jt}) s^c_{jt}(p^c_{jt}, p_{-jt}) M + \sum_{k \neq j} (p_{kt} - w_{kt} - c_{kt}) s^c_{kt}(p^c_{jt}, p_{-jt}) M - A^r_{jt},
\end{aligned}
\tag{3.3.6}
$$

where p^c_{jt} and $s^c_{.t}(p^c_{jt}, p_{-jt})$ denote price and market shares in the counterfactual scenario of a price adjustment. Using inequality (3.3.6) and rearranging terms, a lower bound for the adjustment costs of product j is given by

$$
\begin{aligned}
A^r_{jt} \geqslant \underline{A^r_{jt}} = \big[&(p^c_{jt} - w_{jt} - c_{jt})\; s^c_{jt}(p^c_{jt}, p_{-jt}) - (p_{jt-1} - w_{jt} - c_{jt}) s_{jt}(p_{jt-1}, p_{-jt}) \\
&+ \sum_{k \neq j} (p_{kt} - w_{kt} - c_{kt})(s^c_{jt}\; (p^c_{jt}, p_{-jt}) - s_{kt}(p_t))\big]\; M,
\end{aligned}
\tag{3.3.7}
$$

■ Manufacturer problem

Each manufacturer f sets optimal wholesale prices by solving the following optimization program

$$
\max_{\{w_{jt}\}} \; \Pi^f_t = \sum_{j \in G_f} (w_{jt} - \mu_{jt}) s_{jt}(p_t(w_t)) M \tag{3.3.8}
$$

where G_f denotes manufacturer f's product range, and μ_{jt} its constant marginal cost of production of product j.

Assuming the existence of a pure-strategy Nash equilibrium in wholesale prices, first order conditions of the problem in (3.3.8) are as follows

$$
s_{jt}(p_t) + \sum_{k \in G_f} \sum_{l=1}^{J} (w_{kt} - \mu_{kt}) \frac{\partial s_{kt}}{\partial p_{lt}} \frac{\partial p_{lt}}{\partial w_{jt}} = 0, \quad \text{for all } j \in G_f
$$

The FOCs yield a system of equations, one for each product j in manufacturer f's product line. In order to write this system in matrix notation, define P_w as a $J \times J$ matrix containing retail prices responses to changes in wholesale prices, with entry $P_w(j,i) = \frac{\partial p_j}{\partial w_i}$ for $j, i \in \{1, \dots, J\}$. Moreover, let manufacturer f's ownership matrix I_f be a diagonal matrix of dimension J with jth entry equal to 1 if product j is in her product range and zero otherwise. Finally, denote $\Gamma_t \equiv w_t - \mu_t$ the vector of wholesale margins. FOCs imply that for all $f = 1, \dots, N$ are given by

$$
I_f \Gamma_t = -(I_f P_w S_p I_f)^{-1} I_f s(p_t) \tag{3.3.9}
$$

For wholesale margins to be identified, I need to be able to compute the matrix Pw. I follow Bonnet and Dubois (2015) and obtain an expression for this matrix by differentiating retailer's FOCs, under the assumption that retailers act as Stackelberg followers of manufacturers and set retail prices given wholesale prices. Formally, for all $j, k = 1, \dots, J$, the derivative of equation (3.3.3) with respect to wholesale prices is given by

$$\sum_{i=1}^{J} \frac{\partial s_{kt}}{\partial p_{it}} \frac{\partial p_{it}}{\partial w_{kt}} - \frac{\partial s_{kt}}{\partial p_{jt}} + \sum_{i=1}^{J} \frac{\partial s_{it}}{\partial p_{jt}} \frac{\partial p_{it}}{\partial w_{kt}} + \sum_{i=1}^{J} \left[(p_{it} - \mu_{it} - c_{it}) \sum_{l=1}^{J} \frac{\partial^2 s_{it}}{\partial p_{jt} \partial p_{lt}} \frac{\partial p_{lt}}{\partial w_{kt}} \right] = 0 \tag{3.3.10}$$

The expression of the system of equations defined by (3.3.10) requires the computation of matrices of second price derivatives of market shares with respect to retail prices of all products. I compute those matrices by taking the vector of first derivatives of the J market shares with respect to the price of product j and differentiate each entry of this vector with respect to each price $p_1, \dots, p_J$. This results in a matrix of second derivatives per product j. Let the jth matrix of second derivatives of market shares with respect to retail prices be given by

$$S_p^{p_j} \equiv \begin{bmatrix} \frac{\partial^2 s_{1t}}{\partial p_{jt} \partial p_{1t}} & \cdots & \frac{\partial^2 s_{Jt}}{\partial p_{jt} \partial p_{1t}} \\ \vdots & & \vdots \\ \frac{\partial^2 s_{1t}}{\partial p_{jt} \partial p_{Jt}} & \cdots & \frac{\partial^2 s_{Jt}}{\partial p_{jt} \partial p_{Jt}} \end{bmatrix}$$

Equation (3.3.10) can be written in matrix notation as

$$P_w S_p + P_w (S_p)' + P_w (S_p^{p_1} \gamma_t | \dots | S_p^{p_J} \gamma_t) - S_p = \mathbf{0},$$

where notation (a|b) means horizontal concatenation of vectors a and b. Rearranging terms and solving for P_w yields

$$P_w = S_p \left[S_p + S_p^{'} + (S_p^{p_1} \gamma_t | \dots | S_p^{p_J} \gamma_t) \right]^{-1} \tag{3.3.11}$$

3.3.1.2 Nonlinear contracts

Suppose now that manufacturers and retailers sign nonlinear contracts in the form of two-part tariffs. RPM may take place. I follow the literature (Bernheim and Whiston, 1985; Rey and Vergé, 2010; Bonnet and Dubois, 2010,2015; and Bonnet et al., 2013) and characterize subgame perfect equilibria of the following game. Manufacturers make take-it-or-leave-it offers of contracts to retailers consisting of a fixed franchise fee F_{jt} and a price per unit of product j, w_{jt}. The offer will consist also of a retail price p_{jt} whenever manufacturers can use RPM. Each manufacturer announces contracts to retailers. These offers are private information. Then the retailer announces which contracts she is willing to accept. These announcements are public information. If all offers are accepted, retailers (manufacturers if RPM) set retail prices and contracts are implemented. On the other hand, if one offer is rejected firms earn zero profits and the game ends.

Let the profit function of the retailer be given by

$$\Pi_t^r = \sum_{j=1}^{J} \left[(p_{jt} - w_{jt} - c_{jt}) s_{jt}(p_t) M - F_{jt} - \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A_{jt}^r \right] \tag{3.3.12}$$

Manufacturer f sets wholesale prices w_{kt} and franchise fees F_{kt} by maximizing the profit function given by

$$\Pi_t^f = \sum_{k \in G_f} \left[(w_{kt} - \mu_{kt}) s_{kt}(p_t) M + F_{kt} \right] \tag{3.3.13}$$

subject to retailer's participation constraint

$$\Pi_t^r \geqslant \overline{\Pi}_t^r$$

where $\overline{\Pi}_t^r$ is retailer's reservation value capturing what he would have got from the best outside alternative had he chosen to reject an offer. Participation constraints must be binding, otherwise there will still be room for increases in fixed fees, F_{jt} (Rey and Vergé, 2010). I normalize the reservation value to zero and use binding participation constraints to find an expression for the franchise fees. Rearranging for F_j yields

$$\sum_{j=1}^{J} F_{jt} = \sum_{j=1}^{J} \left[(p_{jt} - w_{jt} - c_{jt}) s_{jt}(p_t) M - \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A_{jt}^r \right]$$

Provided that the retailer carries brands of all manufacturers, we can decompose the total sum of franchise fees as $\sum_j F_{jt} = \sum_{j \in G_f} F_{jt} + \sum_{j \notin G_f} F_{jt}$. Plugging this in the previous expression and rearranging yields

$$\sum_{j \in G_f} F_{jt} = \sum_{j=1}^{J} \left[(p_{jt} - w_{jt} - c_{jt}) s_{jt}(p_t) M - \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A_{jt}^r \right] - \sum_{j \notin G_f} F_{jt}$$

Plugging this expression into f's profit function in (3.3.13), yields:[8]

$$\begin{aligned} \Pi_t^f = & \sum_{k \in G_f} (p_{kt} - \mu_{kt} - c_{kt}) s_{kt}(p_t) M + \sum_{k \notin G_f} (p_{kt} - w_{kt} - c_{kt}) s_{kt}(p_t) M - \sum_{j \notin G_f} F_{jt} \\ & - \sum_{j \in G_f} \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A_{jt}^r - \sum_{j \notin G_f} \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A_{jt}^r \end{aligned} \tag{3.3.14}$$

[8] See Bonnet and Dubois (2010) for details on how to find the final expression of manufacturer's profit.

This equation shows that manufacturer f bears retailer's adjustment costs on both her products and rivals' products.

■ Two-part tariffs with RPM

Suppose manufacturers are able to use RPM. Then, in addition to wholesale prices and fixed fees they will set retail prices as long as by doing so they can always replicate retail prices and profits that would result in a context of no RPM, independently of the strategies of rivals, i.e. whenever possible, the use of RPM is a dominant strategy for manufacturers (Rey and Vergé, 2010).

In this context, wholesale prices do not play a direct role on manufacturer f's own profit, but rather a strategic role at the horizontal dimension. In fact, in addition to controlling retail prices, manufacturers use franchise fees to extract profits from the retailer, which makes them indifferent to the level of wholesale prices of their own products. On the other hand, manufacturer f's wholesale prices can affect rivals in two ways: through market shares that are functions of the vector of prices, and through retail prices provided they are decreasing functions of the vector of wholesale prices w^*. There is thus more instruments than targets in this problem, and consequently a continuum of equilibria, one for each vector of wholesale prices. Empirically, this implies a problem of identification that needs to be accounted for. See subsection 3.4.3 below for details on how I circumvent this problem.

As previously described, the presence of adjustment costs shapes optimal price setting behavior implying that for some periods retail prices may not satisfy first order conditions, as it is more profitable to leave the price constant. In the context of RPM, is the manufacturer who weighs benefits and costs of changing the retail price of its own products. Again, two cases arise.

Case 1: The retail price changes from the previous period ($p_{jt} \neq p_{jt-1}$). Manufacturer f is willing to adjust the retail price of product j at period t if

$$\begin{aligned}(p_{jt} - \mu_{jt} - c_{jt})s_{jt}(p_t)M + &\sum_{k\in G_f}(p_{kt} - \mu_{kt} - c_{kt})s_{kt}(p_t)M \\ &+ \sum_{k\notin G_f}(p^*_{kt} - w^*_{kt} - c_{kt})s_{kt}(p_t)M - A^r_{jt} \\ \geqslant (p_{jt-1} - \mu_{jt} - c_{jt})s^c_{jt}(p_{jt-1}, p_{-jt})M &+ \sum_{k\in G_f}(p_{kt} - \mu_{kt} - c_{kt})s^c_{kt}(p_{jt-1}, p_{-jt})M \\ &+ \sum_{k\notin G_f}(p^*_{kt} - w^*_{kt} - c_{kt})s^c_{kt}(p_{jt-1}, p_{-jt})M, \quad k \neq j\end{aligned} \tag{3.3.15}$$

Optimal retail prices are then set by manufacturer f by solving the program given by[9]

$$\max_{\{p_{kt}\}_{k\in G_f}} \sum_{k\in G_f} (p_{kt} - \mu_{kt} - c_{kt}) s_{kt}(p_t) + \sum_{k\notin G_f} (p_{kt} - w_{kt} - c_{kt}) s_{kt}(p_t),$$

The FOCs of this program, for all $j \in G_f$ write as

$$s_{kt}(p_t) + \sum_{k\in G_f} (p_{kt} - \mu_{kt} - c_{kt}) \frac{\partial s_{kt}}{\partial p_{jt}} + \sum_{k\notin G_f} (p^*_{kt} - w^*_{kt} - c_{kt}) \frac{\partial s_{kt}}{\partial p_{jt}} = 0 \tag{3.3.16}$$

In matrix notation, the FOCs write as follows

$$I_f s(p_t) + I_f S_p I_f (\gamma_t + \Gamma_t) - I_f S_p (I - I_f) \Gamma_t = \mathbf{0}$$

Rearranging for total margins yields, for all $f = 1, \ldots, N$

$$I_f(\gamma_t + \Gamma_t) = -(I_f S_p I_f)^{-1} \left[I_f s(p_t) - I_f S_p (I - I_f) \Gamma_t \right] \tag{3.3.17}$$

Notice that the system of equations depend on both retail margins γ_t and wholesale margins Γ_t. This entails a problem of identification as long as there are more unknowns than equations. It emerges as a consequence of the use of RPM by manufacturers, as discussed previously. Further restrictions should be imposed for identification (see subsection 3.4.3 for details).

By rearranging terms in (3.3.15), we obtain an upper bound for the adjustment costs of a manufacturer that can exert RPM

$$\begin{aligned} A^r_{jt} \leqslant \overline{A^r_{jt}} = \Big[& (p_{jt} - \mu_{jt} - c_{jt}) s_{jt}(p_t) - (p_{jt-1} - \mu_{jt} - c_{jt}) s^c_{jt}(p_{jt-1}, p_{-jt}) \\ & + \sum_{k\in G_f} (p_{kt} - \mu_{kt} - c_{kt})(s_{kt}(p_t) - s^c_{kt}(p_{jt-1}, p_{-jt})) \\ & + \sum_{k\notin G_f} (p^*_{kt} - w^*_{kt} - c_{kt})(s_{kt}(p_t) - s^c_{kt}(p_{jt-1}, p_{-jt})) \Big] M, \quad k \neq j. \end{aligned} \tag{3.3.18}$$

Case 2: The retail price does not change from the previous period ($p_{jt} = p_{jt-1}$). Manufacturer f chooses not to change the price of product j at period

[9] For simplicity, I omit fixed fees and adjustment costs from the profit as they are constant at t.

t if

$$
\begin{aligned}
&(p_{jt-1} - \mu_{jt} - c_{jt}) s_{jt}(p_{jt-1}, p_{-jt}) M + \sum_{k \in G_f} (p_{kt} - \mu_{kt} - c_{kt}) s_{kt}(p_{jt-1}, p_{-jt}) M \\
&\qquad + \sum_{k \notin G_f} (p^*_{kt} - w^*_{kt} - c_{kt}) s_{kt}(p_t) M \\
&\geqslant (p^c_{jt} - \mu_{jt} - c_{jt}) s^c_{jt}(p^c_{jt}, p_{-jt}) M + \sum_{k \in G_f} (p_{kt} - \mu_{kt} - c_{kt}) s^c_{kt}(p^c_{jt}, p_{-jt}) M \\
&\qquad + \sum_{k \notin G_f} (p^*_{kt} - w^*_{kt} - c_{kt}) s^c_{kt}(p^c_{jt}, p_{-jt}) M - A^r_{jt}, \;\; k \neq j.
\end{aligned}
\tag{3.3.19}
$$

By rearranging terms, we obtain the following lower bound for the adjustment costs of a manufacturer that can exert RPM

$$
\begin{aligned}
A^r_{jt} \geqslant \underline{A^r_{jt}} = &\Big[(p^c_{jt} - \mu_{jt} - c_{jt}) s^c_{kt}(p^c_{jt}, p_{-jt}) - (p_{jt-1} - \mu_{jt} - c_{jt}) s_{jt}(p_{jt-1}, p_{-jt}) \\
&+ \sum_{k \in G_f} (p_{kt} - \mu_{kt} - c_{kt})(s^c_{kt}(p^c_{jt}, p_{-jt}) - s_{kt}(p_{jt-1}, p_{-jt})) \\
&+ \sum_{k \notin G_f} (p^*_{kt} - w^*_{kt} - c_{kt})(s^c_{kt}(p^c_{jt}, p_{-jt}) - s_{kt}(p_{jt-1}, p_{-jt}))\Big] M, \;\; k \neq j.
\end{aligned}
\tag{3.3.20}
$$

■ Two-part tariffs without RPM

Suppose manufacturers are not allowed to use RPM. Therefore, they make offers to the retailer consisting of a wholesale price and a fixed fee per product. They delegate the task of optimally setting retail prices to the retailer and, as a consequence, is the retailer who should solve the optimal repricing problem and bear the costs of adjustment in case she decides to do so.

□ Retailer problem

Retailer sets prices by maximizing profits, which are given by

$$
\Pi^r_t = \sum_{j=1}^{J} \left[(p_{jt} - w_{jt} - c_{jt}) s_{jt}(p) M - F_{jt} - \mathbb{1}_{\{p_{jt} \neq p_{jt-1}\}} A^r_{jt}\right]
$$

Provided that franchise fees and adjustment costs are constant at t, the first order conditions with respect to retail prices are the same as those under linear pricing and are given by (3.3.3). On the other hand, adjustment costs are not contractible and repricing decisions are not observable at the moment manufacturers make offers to the retailer. As a consequence, franchise fees are not contingent, i.e. they do not vary depending on whether prices are adjusted or not. This implies that fixed fees cancel out from the expression that compares actual and counterfactual profits in the optimal repricing problem. Hence, bounds in a context of simple two-part tariffs are exactly the same as in linear pricing and are given by equations (3.3.5) and (3.3.7).

□ Manufacturer problem

The problem of manufacturer f in a context of no RPM is to optimally set wholesale prices and franchise fees. She does this by maximizing (3.3.14) with respect to wholesale prices, given those of other manufacturers. The FOCs of this program, for all $i \in G_f$ write as

$$\sum_{k=1}^{J} \frac{\partial p_{kt}}{\partial w_{it}} s_{kt}(p_t) + \sum_{k \in G_f} \left[(p_{kt} - \mu_{kt} - c_{kt}) \sum_{j=1}^{J} \frac{\partial s_{kt}}{\partial p_{jt}} \frac{\partial p_{jt}}{\partial w_{it}} \right] + \sum_{k \notin G_f} \left[(p_{kt} - w_{kt} - c_{kt}) \sum_{j=1}^{J} \frac{\partial s_{kt}}{\partial p_{jt}} \frac{\partial p_{jt}}{\partial w_{it}} \right] = 0.$$

In matrix notation, f's FOCs are given by:

$$I_f P_w s(p_t) + I_f P_w S_p I_f (\gamma_t + \Gamma_t) + I_f P_w S_p (I - I_f)\gamma_t = \mathbf{0}$$

From this equation, I can derive an expression for the total margins of manufacturer $f \in \{1, \ldots, N\}$

$$I_f(\gamma_t + \Gamma_t) = -(I_f P_w S_p I_f)^{-1} \left[I_f P_w s(p_t) + I_f P_w S_p (I - I_f)\gamma_t \right],$$

plugging the expressions of both retail margins given by (3.3.4) and matrix P_w given by (3.3.11) into the previous equation yields an expression that is identified from data and estimates of the demand parameters.

3.3.2 Demand

I index consumers by $i = 1, 2, \ldots, I$. The conditional indirect utility consumer i derives from purchasing product j in market t writes as

$$u_{ijt} = x_j \beta_i - \alpha_i p_{jt} + \xi_j + \eta_t + \Delta\xi_{jt} + \epsilon_{ijt} \tag{3.3.21}$$

where x_j is a row vector containing K observable characteristics of product j that do not vary across markets, p_{jt} is the unit price of product j in market t, ξ_j captures the mean (across individuals and time) valuation of the unobserved (by the econometrician) product characteristics, η_t denotes time fixed effects that account for both unobserved determinants that vary with time, and time trends, and $\Delta\xi_{jt} = \xi_{jt} - \xi_j$ captures market-specific deviations from this mean under the assumption that in each market people value differently product characteristics. Finally, I allow individual heterogeneity enter the model through the standard additive separable mean-zero random shock ϵ_{ijt} and $K+1$ individual-specific parameters (α_i, β_i). These coefficients aim at capturing individual marginal valuations of price and product characteristics and are modelled as a function of observed and unobserved demographics, as follows:

$$\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \pi \text{income}_i + \Sigma v_i, \qquad v_i \sim N(0, I_{K+1})$$

where α and β are $K+1$ mean taste coefficients common to all individuals, π is a $(K+1)\times 1$ vector of coefficients that measure how valuations of product characteristics vary with individual income and Σ is a $(K+1)\times(K+1)$ scaling matrix to be estimated.

I define the "outside good" as any alternative brand or type of breakfast cereal, or any other product not included in the choice set; it too accounts for the no purchase option. Normalizing the mean utility to zero, the indirect utility derived from the outside option writes as $u_{i0t} = \epsilon_{i0t}$.

Following Nevo (2000), the utility in (3.3.21) can be expressed as the sum of a mean utility common to all consumers and an idiosyncratic deviation from this mean:

$$u_{ijt} = \delta_{jt}(p_{jt}, \xi_j, \eta_t, \Delta\xi_{jt}; \alpha, \beta) + \mu_{ijt}(p_{jt}, \text{income}_i, v_i; \pi, \sigma) + \epsilon_{ijt} \tag{3.3.22}$$

with $\delta_{jt} = x_j\beta - \alpha p_{jt} + \xi_j + \eta_t + \Delta\xi_{jt}$ and $[p_{jt}, x_j]' * (\pi\text{income}_i + \Sigma v_i)$. A key assumption of this model is that consumers choose at most one unit of the brand that gives the highest utility. Suppose that ϵ_{ijt} is distributed i.i.d. type I extreme value, then the aggregate market share of product j at period t as a function of mean utility levels of all the $J+1$ products, given the parameters, is given by:

$$s_{jt} = \int \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{J} \exp(\delta_{kt} + \mu_{ikt})} dF(\mu) \tag{3.3.23}$$

where $F(\cdot)$ denotes population distribution function.

3.4 Empirical implementation

The estimation of the model described in Subsection 3.3.2 is conducted following standard discrete choice methods —Berry, 1994, Berry, Levinsohn and Pakes (1995), Nevo (2000, 2001). In this Section, I give details on the data used for estimation, the estimation method and I discuss identification issues and how I deal with them.

3.4.1 The final data set

The model presented in the previous Section relies on tracking price changes on a week-to-week basis (recall that *DFF* sets prices on a weekly basis). Even though prices were reported with quite good regularity, information on some weeks for most stores is missing. In particular, there are three main interruptions: in May 1990 (4 weeks), in September-October 1994 (4 weeks) and between February and August 1995 (25 weeks). To circumvent this problem, I restrict the sample to stores with the largest number of price observations in the period comprised between 24 May 1990 and 14 September 1994, which gives 224 weeks. In this period, I observe the least number of consecutive weeks with missing price data for each store. The final sample includes, thus, 71 stores that represent the 85.5% of the total number of stores in the chain. Further, I aggregate prices and quantities across stores into three price zones: high-price (24 stores), medium-price (30 stores) and low-price (17 stores including 'Club-fighter' stores).

From 490 UPC observed in *Dominick's* database I keep the 22 leading based on the overall market share in the last quarter of the sample period. I define a product as one serving of a UPC of RTE cereal, according to the weight suggested by manufacturers which I assume is a good approximation to the true serving consumers have. Notice that the same brand can enter the database with different UPCs depending on specific characteristics. For example, different box sizes of Special K are coded as separate UPCs and may have different price schedules and promotional activity. Due to this, I treat different UPCs of the same cereal brand as separate products.

I define a market as a zone-week combination, which gives 672 markets. Product market shares are computed as the number of servings sold of each product in a market divided by the potential number of servings that can be sold in the Chicago area in a week. Following Nevo (2001), this potential is assumed to be one serving per capita per day. The market share of the outside alternative corresponds then to one minus the sum of market shares across products. Retail and wholesale prices per serving were computed as total dollar sales divided by the number of servings sold in a market.

Finally, I complement *Dominick's* database with data on brand characteristics

Table 3.1: Summary statistics of brands in the sample

Variable	Mean	Median	Std. Dev.	Min	Max
Serving weight (g)	32.73	29.5	8.84	27	58
Amounts per serving					
Calories	123.18	110	31.39	100	210
Caories from Fat	8.41	10	6.29	0	25
Sugar (g)	6.95	8	3.78	0	12
Fiber (g)	2.55	3	1.70	0	7
Protein (g)	3.05	2	2.70	1	10
Brands by segment (%)					
All family segment	31.82	—	—	—	—
Kids segment	31.82	—	—	—	—
Adult segment	36.36	—	—	—	—

Notes: Based on 14,784 observations. *Source:* Cereal boxes.

such as calories from fat, sugar, fiber and protein contents taken from cereal boxes and segment indicators (Kids, All-family and Adults).[10]

I exploit the panel structure of my data to control for product fixed-effects by including product dummy variables, which captures brand unobserved characteristics. Thanks to this, demand is identified without the need to characterize the supply side.[11]

3.4.2 Estimation

Estimation relies on the population moment conditions given by $E[h(z)'\rho(x, \theta_o)] = 0$, where $z_1, ..., z_M$ are a set of instrumental variables, ρ is a function of the parameters of the model and θ_o is the true value of the parameters (see subsection 3.4.3 below for a discussion on the instruments used for identification). A generalized method of moments estimator is obtained by solving the problem

$$\min_{\theta} \rho(\theta)' h(z) \hat{\Lambda}^{-1} h(z)' \rho(\theta), \tag{3.4.1}$$

[10] I classify cereals by segments following Nevo (2001) and categories available on manufacturers web sites. Kids segment: (Kellogg) Froot Loops, Frosted Flakes, Corn Pops, Apple Jacks; (General Mills) Golden Grahams, Honey Nut Cheerios; and (Quaker Oats) CapN' Crunch. Adults segment: (Kellogg) Special K; (General Mills) Total, Whole Grain Total; (Quaker Oats) Oat Squares; (Post) Grape-Nuts; and (Nabisco) Spoon Size Shredded Wheat. All-family segment: (Kellogg) Cocoa Krispies, Corn Flakes; (General Mills) Cheerios, Rice Chex; and (Post) Honey Comb.

[11] For a detailed discussion of the differences with BLP's procedure, see Nevo (2000a, 2001).

where $\hat{\Lambda}$ is a consistent estimator of $E[h(z)'\rho\rho' h(z)]$ and plays the role of the optimal weighting matrix in expression (3.4.1).

Now, according to the empirical framework described before, once product dummy variables are included, the error term of the model is $\Delta\xi_{jt}$ which can be computed as a function of the mean utilities δ_{jt}, the data and the parameters. Following Berry (1994), this computation requires solving first for δ_{jt} from the system of equations resulting from the match of observed and predicted market shares

$$s_{jt}(x, p_t, \delta_t; \pi, \sigma) = S_{jt} \tag{3.4.2}$$

where $s_{jt}(\cdot)$ is the predicted market share function defined in (3.3.23) and S_{jt} denotes the market share of product j observed in the data. As the system in (3.4.2) does not have a closed-form solution for the the mixed Logit case, it should be solved numerically. After inverting (3.4.2) in order to express δ_{jt} as an explicit function of the observed market shares, the error term in (3.4.1) writes as

$$\rho_{jt} = \delta_{jt}(x, p_t, S_t; \pi, \sigma) - (x_j\beta - \alpha p_{jt} + \xi_j + \eta_t)$$

The estimation of the parameters is performed using a non-linear search. To do this, I use the standard estimation algorithm proposed by BLP (1995) and improved by Nevo (2000) and Knittel and Metaxoglou (2012).[12]

3.4.3 Identification

There are two identification issues. One concerns the supply side and, in particular, the multiplicity of equilibria arising in the model of two-part tariffs with RPM. The other one is related to the demand specification and the endogeneity of prices.

As described in subsection 3.3.1.2, in the model of two-part tariffs with RPM there is a coordination problem due to there are more instruments –retail price that helps coordinate joint profits and wholesale price and fixed fees that help extract profit from the retailer (Rey and Vergé, 2010)– than targets. Empirically, this implies that we do not have enough equations to separately identify wholesale and retail markups. More importantly, if we were interested in estimating total margins, we still would not know which equilibrium of the infinitely many we are estimating. A particular equilibrium must be selected by imposing further restrictions on the problem. To deal with this, I assume that manufacturers set wholesale prices equal to marginal costs ($w^* = \mu$), which corresponds to a symmetric subgame perfect equilibrium where retail prices are at the monopoly level, retailers earn zero profit and upstream firms share equally the monopoly profit (Rey and Vergé, 2010). With

[12] For a detailed description and discussion of the estimation method, and the algorithm, see Knittel and Metaxoglou (2012).

this assumption, total markups $(\Gamma + \gamma)$ in equation (3.3.17) will be equal to retail markups as upstream margins are zero $(\Gamma = 0)$.

There are several important reasons to choose this equilibrium. First, Rey and Vergé (2010) show that there always exist such an equilibrium even if retail price responses are not 'well defined'. Second, because this is the only equilibrium in which manufacturers attain maximum profits. Last, because this is the only equilibrium which is robust to the introduction of solutions to the coordination problem.

A second identification problem stems from the correlation of retail prices with the local deviation of the mean valuation of product unobserved characteristics $\Delta\xi_{jt} \equiv \rho_{jt}(\theta)$, under the assumption that both firms and customers observe those characteristics and, consequently, their decisions account for these local deviations. This endogeneity issue requires finding a set of exogenous variables $z_1, ..., z_M$ that are correlated with the price but uncorrelated with the error term to be used as instrumental variables so that the mean independence assumption, on which the estimation of the model relies, is satisfied.

Different sets of instruments have been used in the IO literature of incomplete pass-through, namely, nominal prices of inputs (Bonnet et al., 2013; Hellerstein, 2008; Goldberg and Hellerstein, 2013) and nominal exchange rates in the case of a local market with imported products (Goldberg and Hellerstein, 2013). I try both types of instruments separately and altogether and select the one giving the best estimates.

The first set of IVs I use is monthly nominal prices of commodities used as inputs in the production of RTE cereals, namely, oil, wheat, corn, and sugar, and an employment cost index for total compensation of workers of the goods-producing industries in the U.S. Prices of inputs are valid instruments as they are correlated with retail prices since they make part of manufacturers' costs, which is reflected in wholesale prices which are, at the same time, components of retailer's marginal costs. On the other hand, it is very unlikely that input prices respond to local demand shocks for RTE cereals or retailer's promotional activities.

The second set of IVs is monthly nominal exchange rates. Even though the RTE breakfast cereals consumed by U.S. citizens are all domestic-produced goods and the U.S. has been historically a net exporter of agricultural products related to cereals production (excepting sugar and cocoa),[13] I claim that shocks affecting bilateral trade between the U.S. and its main international markets for agricultural products, such as corn and wheat, may affect domestic prices of goods using such products as inputs. Under this argument I use bilateral nominal exchange rates of India, South Korea, Colombia and Mexico, which have larger correlations with RTE cereals' retail prices as compared to those of commodities. The exogeneity of these

[13] According to the U.S. Department of Agriculture, U.S. agricultural exports have been larger than imports ever since 1960. See: http://www.ers.usda.gov/topics/international-markets-trade/us-agricultural-trade.aspx.

IVs follows the same argument as that of input prices: from the point of view of a specific local market, exchange rates seem to be randomly determined.

3.5 Results

3.5.1 Results from Logit regressions

As a first step, I perform Logit regressions both without IVs and with different sets of IVs, to get a preliminary idea on how prices are responding to controls and IVs. Results, given in Table 3.1, show that all estimates are significant and most are of the expected sign. The estimated marginal utility of price becomes more elastic as controls for brand and households characteristics are added to the model. This is also the case when IVs are included. Columns (4) and (6) show results of 2SLS regressions using commodity prices ('cost') as IVs, and columns (5) and (7) give results using exchange rates as IVs. Overall, both sets of instruments seem to have a similar power as in all cases first-stage R-squared and F-tests are large and of similar magnitudes. Last column in Table 3.1 uses both exchange rates and commodity prices as IVs, and as expected, the price coefficient increases but continues to be significant as well as all other estimates. First- and second-stage R-squared as well as F-test do not vary in an important way though, as compared to regressions with either set of IVs.

3.5.2 Results from the mixed-Logit model

Table 3.2 displays the results of a full random-coefficients Logit model with exchange rates IVs. Given that the regression includes brand dummy variables that capture both observed and unobserved product characteristics, mean coefficients of brand attributes are backed out using a minimum distance procedure (see Nevo, 2001).

Results show that, on average, consumers value negatively product characteristics such as calories from fat and sugar contents, which is somewhat expected and coincides with Logit predictions. However, the utility decreases with higher contents of healthy ingredients such as fiber and protein. Cereals in the Kids and Adults segments are preferred to those in the All-family segment. The estimated standard deviations (σ's) of price and Adults segment coefficients are not significant, suggesting that most of the heterogeneity is explained by income. All other estimated standard deviations are significant, although just that of Kids segment and Calories from fat are larger than one. This may indicate that, with the exception of these two, unobserved demographics have lower explanatory power than included demographics.

The estimate of log of Income (interacted with the constant), the only observed household characteristic included in the model, is negative and significant. The latter

Table 3.1: Results from Logit demand

	OLS			IV				
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price	-2.550	-32.563	-36.309	-47.503	-59.189	-60.183	-71.828	-65.367
	(0.357)	(0.693)	(0.641)	(9.114)	(11.424)	(10.568)	(13.588)	(10.681)
Constant	-6.930	-6.592	-20.668	-5.737	-5.068	-17.776	-16.365	-17.148
	(0.068)	(0.055)	(0.588)	(0.522)	(0.653)	(1.410)	(1.786)	(1.439)
Cal from Fat	-0.011	—	—	—	—	—	—	—
	(0.001)							
Sugar	-0.053	—	—	—	—	—	—	—
	(0.002)							
Fiber	-0.078	—	—	—	—	—	—	—
	(0.006)							
Protein	0.048	—	—	—	—	—	—	—
	(0.002)							
Kids	0.479	—	—	—	—	—	—	—
	(0.017)							
Adults	-0.242	—	—	—	—	—	—	—
	(0.021)							
Log of Median	—	—	2.052	—	—	2.013	1.993	2.004
Income			(0.067)			(0.075)	(0.085)	(0.079)
Av. hh size	—	—	-2.808	—	—	-3.226	-3.430	-3.317
			(0.059)			(0.200)	(0.253)	(0.204)
R^2	0.221	0.603	0.668	0.566	0.485	0.579	0.470	0.536
1st Stage R^2	—	—	—	0.901	0.901	0.906	0.906	0.906
1st Stage F-test	—	—	—	1,159	1,162	1,263	1,261	1,256
Instruments	—	—	—	cost	exchange	cost	exchange	cost,
	—	—	—		rates		rates	ex. rates

Notes: Dependent variable is $ln(S_{jt}) - ln(S_{ot})$. Based on 14,784 observations. All regressions include time dummy variables and, with the exception of column (1), all regressions include brand dummy variables. Asymptotically robust s.e. are reported in parentheses. All estimates are significant at 1% level.

indicates that larger income households prefer to have other alternatives at breakfast than the included RTE cereals (other brands, or totally different products). On the other hand, the estimate of the interaction of income with price is not significant, although the coefficient is positive suggesting that above-average income individuals are less price sensitive, which is in line with findings of previous literature (Nevo, 2001).

Table 3.2: Results from the Mixed Logit model[a]

Variable	Means (β's)	Std. Deviations (σ's)	Interactions with Log of Income
Price	-45.251***	0.024	0.135
	(0.031)	(0.026)	(0.117)
Constant[b]	2.052***	0.141***	-0.097***
	(0.038)	(0.050)	(0.014)
Cal from Fat[b]	-1.615***	1.017***	—
	(0.004)	(0.005)	
Sugar[b]	-0.175***	0.073***	—
	(0.005)	(0.010)	
Fiber[b]	-3.171***	0.867***	—
	(0.021)	(0.017)	
Protein[b]	-1.492***	0.761***	—
	(0.011)	(0.011)	
Kids[b]	6.131***	2.739***	—
	(0.062)	(0.040)	
Adults[b]	5.422***	0.081	—
	(0.054)	(0.073)	
MD R^2	0.933		
GMM objective	3.69E-19		

Notes: [a] Based on 14,784 observations. Except where noted, parameters were estimated using GMM. All regressions include brand and week dummies. Asymptotically robust s.e. are given in parentheses.
[b] Means estimated using a minimum-distance procedure.
***Significant at 1% level.

Table 3.3 presents the medians of the distribution of derived own- and cross-price elasticities, and standard deviations. Results show that demand is considerably elastic to changes in price: overall the median elasticity across brands and markets is -8.16, consumers are less elastic to changes in prices of Kellogg's cereals, whereas a huge median elasticity is obtained for Quaker's products. A similar pattern is observed with cross-price elasticities.

Table 3.3: Median own- and cross-price elasticities of RTE cereal brands by manufacturer[a]

	Own-price elasticity		Cross-price elasticity	
Manufacturer	Median[b]	Std. deviation	Median[c]	Std. deviation
General Mills	-9.047	16.627	0.139	0.341
Kellogg	-1.321	33.476	0.017	0.471
Nabisco	-14.084	22.527	0.421	0.687
Post	-5.468	14.885	0.142	0.562
Quaker	-44.531	144.507	1.029	7.868
All	-8.155	61.764	0.136	3.084

Notes: [a]Based on 14,784 observations.
[b]Each entry corresponds to the median of the elasticities across the 672 markets.
[c]Each entry was obtained as the mean of cross-price elasticities of each product with respect to all other prices, and then the median across the 672 markets by brand.

3.5.3 Implied margins and retail marginal costs

Once demand coefficients have been estimated, I can compute retail and wholesale margins from FOCs under each supply specification. According to the structural model, FOCs only hold for prices that adjusted at t. This has two implications: first, the system of FOCs can only be used to compute margins of products for which prices adjusted in the current period. Second, the calculation of margins from FOCs includes reaction matrices S_p and P_w which, as previously described, contain derivatives of market shares with respect to retail prices and retail prices with respect to wholesale prices, respectively. In a static framework it is assumed that firms set prices to satisfy static FOCs and, consequently, all prices vary from period to period. This implies that all entries of S_p and P_w are generally different from zero, ignoring the fact that some prices remain constant from previous period and that the derivatives measuring the effects of such prices on market shares as well as their reactions to changes in wholesale prices should be zero.

I follow Goldberg and Hellerstein (2013) and compute margins for each model of supply in a two-step procedure as follows. In a first step, I calculate markups using FOCs for those products for which prices changed at t only, setting entries of reaction matrices S_p and P_w to zero for prices that remained constant. Table 3.4 displays the results.

Under the linear pricing model, predicted retail markups range on average between 11% and 39% of the retail price, whereas wholesale markups range between 7% and nearly 20%. In all cases but one, average retail margins are larger than average wholesale margins. The largest average retail and total markups predicted by the model are those of Kellog's with a considerable gap between the retail margin and the wholesale margin. Quaker products show similar downstream and upstream margins on average. The average retail markup across all products, 29.22%, is not far from Nevo (2001)'s estimate of 35.8%, although the difference between these two numbers is not negligible. It may be explained by the fact that here I account for price rigidity by setting some marginal effects to zero as discussed previously, whereas Nevo (2001) takes prices as varying all the time.[14]

As for the two-part tariff with RPM supply model, recall that I estimate the equilibrium in which manufacturers set wholesale prices equal to marginal costs ($w = \mu$), implying that retail markups equal total markups and producers capture all profit from the retailer through fixed fees. The predicted markups under this supply model are slightly lower on average to retail markups computed under linear pricing. Overall, the average markup is 28.16% and the relative positions of manufactures in the ranking remains the same: Kellogg's markup is the largest with 38.7% and Quaker is the lowest with 8.1%. As compared to total margins under linear tariffs, RPM markups are remarkably lower with differences ranging between 7% to 21%. This is consistent with theory according to which two-part tariff contracts help solve the double marginalization problem emerging under linear pricing schemes and characterized by larger total margins and lower quantity as compared to the monopoly levels. This is so because manufacturers' profit do not depend on wholesale margins as they are able to extract it from retailer through fixed fees. As a consequence, lower total margins are expected.

Finally, the model of two-part tariffs without RPM predicts zero wholesale margins and, consequently, total markups of the industry are exactly equal to retail markups. These coincide with retail margins under linear pricing because the absence of RPM enables the retailer to set prices that maximize own profit. Formally speaking, retailer's FOCs under the two models are exactly the same. This result is consistent with the theory of single common agency with competing manufacturers, which is the case this paper considers given that the data set used contains information on a single retail chain only (DFF).[15] In fact, according to Rey and Vergé (2010) in the presence of a retail monopolist, simple two-part tariffs suffice for manufacturers to attain monopoly prices and profits. By supplying at cost, a manufacturer makes its rivals become residual claimants on the sales of all brands in the market, which turns out to be an incentive for them to supply at cost as well so as to keep retail prices at the monopoly level. As a result, upstream markups are

[14]The mean retail margin retrieved using all FOCs equals 32.6% which is closer to that presented by Nevo (2001).

[15]Recall that the data comes from DFF and contains information only on DFF's stores. Consequently, the restriction of a single retailer in the market is imposed by the data and not as an assumption of the structural model. A model of multiple common agency would be possible with a data base containing information on several competing retailers.

zero and manufacturers earn monopoly profits thanks to franchise fees.[16]

Table 3.4: Mean implied markups by manufacturer for several supply models (% of retail price)

				Two-part tariffs			
	Linear tariffs			with RPM	without RPM		
Producer	Retail	Wholesale	Total	Total	Retail	Wholesale	Total
General Mills	30.74	10.70	38.14	29.69	30.74	0.00	30.74
Kellogg	38.98	6.93	45.81	38.74	38.98	0.00	38.98
Nabisco	10.99	14.80	29.11	9.05	10.99	0.00	10.99
Post	24.34	19.42	43.24	22.57	24.34	0.00	24.34
Quaker	10.55	9.27	22.02	8.06	10.55	0.00	10.55
All	29.22	10.68	38.51	28.16	29.22	0.00	29.22

Notes: Each entry corresponds to the average markup across periods and products of the same manufacturer.

With the retail markups in hand, I compute retailer's total marginal costs as the difference between observed retail prices and retrieved markups for periods in which price changes are observed only. In a second step, I am able to predict the whole vector of retailer's marginal costs for each model h by assuming the following linear specification

$$C_{jt}^h = \varsigma_j^h + \lambda_h d_z + \phi_h AAC_{jt} + \eta_{jt}^h$$

where C_{jt}^h is the retrieved marginal cost using supply model h's FOCs in a first step, ς_j^h is an unknown product-specific parameter, d_z are price zone dummy variables, and AAC_{jt} is the average acquisition cost of product j at t reported by the retailer. Recall that AAC is a proxy of the wholesale price *Dominick's* pays to the manufacturer of each product, so it accounts for an important part of retailer's costs. Finally, η_{jt}^h is an unobservable (to the econometrician) random shock to costs. Assuming that $E[\eta_{jt}^h \,|\, \varsigma_j^h, AAC_{jt}] = 0$, the parameters of the model $(\varsigma_j^h, \lambda_h, \phi_h)'$ can be consistently estimated. Table 3.5 reports the results of regressions for linear tariffs and Two-part tariffs with RPM supply models. As expected, the AAC's coefficient is positive and significant in both regressions. Moreover, the large $R-$squared and $F-$statistic for the two regressions indicate that included covariates have considerable explanatory power.

Table 3.6 gives mean prices across products of each manufacturer as well as results on retrieved and fitted marginal costs. On average, costs range from 40 cents to 85

[16] Given these markups, the remaining of the paper presents results on two supply models: linear pricing and Two-part tariffs with RPM. I omit the third specification as it would give exactly the same results as the linear pricing model.

Table 3.5: Results from OLS regression of structurally retrieved marginal costs on determinants

Variable	Linear tariffs	Two-part tariffs (RPM)
Average acquisition cost	0.604***	0.615***
	(0.043)	(0.043)
Product FE	Yes	Yes
Price zone FE	Yes	Yes
Week dummies	Yes	Yes
R^2	0.932	0.933
F-test	2,891	3,090
Observations	3,766	3,766

Notes: Dependent variable is total retail costs retrieved from the structural model. Asymptotically robust s.e. are reported in parentheses. Results for the supply model of Two-part tariffs without RPM are the same as those from linear tariffs since the retailer's problem is identical.
***Significant at 1% level.

cents per serving for both linear tariffs and two-part tariffs with RPM. The largest costs are those of Quaker's products, followed by General Mills's products which have at the same time the highest average price per serving. By contrast, Kellogg's cereals have an average price similar to General Mills's but the lowest marginal costs (both retrieved and fitted) below 50 cents per serving on average.

3.5.4 Bounds for retail price adjustment costs

To derive bounds of retail price adjustment costs I use expressions (3.3.5) for upper bounds and (3.3.7) for lower bounds in the linear pricing case, and (3.3.18) for upper bounds and (3.3.20) for lower bounds in the two-part tariff with RPM case. Such expressions require observing some components and estimating other, of both actual and counterfactual profits with the exception of fixed fees in the two-part tariffs case as they cancel each other out. The method I will describe below is similar for all the supply specifications I consider, what makes the difference is the terms entering the respective expression.

The computation of upper bounds is the simplest of the two. Recall that expressions for these bounds are obtained by comparing current profits of products for which prices changed in the current period with the counterfactual profit the retailer would have got had she left the price of that product unchanged (Case 1). The vector of counterfactual prices obtains directly by replacing current prices at t

Table 3.6: Mean retireved and fitted retailer total marginal costs according to distinct supply models (averages accross products and markets, in US dollar cents per serving)

	Mean	Total marginal costs[a]			
	retail	Linear tariffs		Two-part tariffs with RPM	
Manufacturer	price	Retrieved[b]	Fitted[c]	Retrieved[b]	Fitted[c]
General Mills	11.07	6.97	7.17	7.09	7.28
Kellogg	11.01	4.42	4.80	4.47	4.83
Nabisco	6.67	5.97	6.11	6.10	6.24
Post	8.15	5.61	5.56	5.75	5.70
Quaker	8.94	8.12	7.18	8.28	7.37
All	10.39	5.82	5.90	5.92	5.98

Notes: [a]Total marginal cost includes the wholesale price per product plus the retailer's marginal cost of distribution.
[b]Based on 3,766 observations.
[c]Based on 14,784 observations.

by prices at $t-1$ for those products for which prices changed at t. Counterfactual market shares are then computed using equation (3.3.23), the counterfactual price vector and estimated coefficients of the demand model.

The expression for lower bounds comes from the comparison of actual profits of products for which prices remained equal from previous period with counterfactual profits the retailer would have got had she decided to adjust the price of that product at t and bear the repricing costs. These prices should be optimal, i.e. they should satisfy current FOCs. I compute the vector of counterfactual prices using equations (3.3.4) for the linear pricing case, and (3.3.17) for the two-part tariffs case. Once prices have been computed, counterfactual market shares are obtained in the same way as for upper bounds, using the expression of the predicted market shares (3.3.23), counterfactual prices and demand estimates.

Table 3.7 gives results of implied upper and lower bounds of retail adjustment costs averaged across products of the same manufacturer and markets ('price zone'-week). By manufacturer, average lower bounds range from US \$18.22 to US \$156.95 in the linear tariffs case and from US \$16.43 and US \$141.27 for the two-part tariffs case. Upper bounds are quite similar for the two supply models and range from US \$66 to US \$2511. The very high mean upper bounds are for Quaker, for which the model predicts very large values for some periods. However, the median upper bound for this brand is US \$30. Notice that with the exception of Quaker, lower bounds under RPM are lower than those under linear pricing. This suggest that, if two-part tariffs with RPM is the true vertical conduct of this industry, linear tariffs give a biased estimate of retail price adjustment costs.

Overall, DFF is willing to adjust the price of one of its products if it obtains an extra profit of at least US \$98.84. On the other hand, the retailer may pay at most US \$447 for adjusting the price of one product, on average. As pointed out by Goldberg and Hellerstein (2013), the magnitudes of these bounds do not give a lot of information. To have an idea of their relative importance, I compute the share of the sum of each lower and upper bounds of adjustment costs on retailer's total revenue for the period considered here (224 weeks). It corresponds to 6.5% and 10.14% in the RPM case.

Table 3.7: Adjustment costs bounds for several supply models (averages across products and markets, in US dollars)

	Linear tariffs		Two-part tariffs RPM	
Producer	Lower bound[a]	Upper bound[b]	Lower bound[a]	Upper bound[b]
General Mills	79.13	247.13	26.23	247.28
Kellogg	156.95	218.66	141.27	219.10
Nabisco	18.22	66.74	18.21	67.13
Post	36.71	115.71	16.43	116.27
Quaker	91.84	2,508.12	297.90	2,511.28
All	109.34	446.07	98.84	446.70
Share on total revenue[c]	7.27%	10.12%	6.52%	10.14%

Notes: [a]Based on 3,746 observations.
[b]Based on 10,971 observations.
[c]Computed as the sum of lower (upper) bounds across all products and markets, divided by total revenue over the 224 weeks.

3.6 Conclusion and future research

This paper aims at investigating the role played by the vertical structure of the industry in the degree of retail price stickiness. To do that, I develop a structural model of vertical relationships between manufacturers and retailers in which three possible vertical conducts can take place: linear pricing, simple two-part tariffs, and two-part tariffs with resale price maintenance (RPM). To account for the fact that retail prices are rigid, I include costs of repricing in the profit function of the retailer. At each period, the retailer weighs benefits and costs of adjusting the price of a product and makes and optimal decision. Using data on prices and sales of ready-to-eat breakfast cereals from a large supermarket chain in Chicago, I estimate demand and retrieve margins and marginal costs under each supply specification. With this in hand, I am able to quantify bounds of repricing costs, i.e. fixed costs that prevent the retailer from adjusting the price of some products at each period and that explain why observed price of RTE cereals remain at a given level for several weeks.

An exploratory analysis of the data shows that cereal prices responses to changes in input prices is very low. Surprisingly, retail prices do not always respond to changes in wholesale prices, which makes plausible the hypothesis of manufacturers using market power to maintain retail prices at a certain level. The estimation of the structural models of demand and supply allows the computation of lower and upper bounds of retail price adjustment costs. In a context of two-part tariffs with RPM, average lower bounds are smaller than those obtained under alternative models of supply. Results suggest that if manufacturers and retailers actually sign two-part tariffs contracts with RPM, the specification of the supply side as linear pricing results in biased adjustment costs. On the other hand, average upper bounds are similar across.

A full assessment of the relative importance of price adjustment costs and other sources of incomplete pass-through such as markup adjustment, and the role two-part tariffs and RPM in it, requires counterfactual simulations. A future version of this research work will include such an analysis. Finally, there is at least one interesting avenue for future research and is related to the case in which the retailer has large enough bargaining power so that she can impose restraints to manufacturers and avoid costs shocks to be passed-through to retail prices. Bonnet and Dubois (2015) provide conditions for identification of models of vertical relations with buyer power that can be extended to the analysis of incomplete pass-through and retail price stability.

Bibliography

1. Aguiar, Mark and Hurst, Erik (2007), "Life-cycle prices and production", *The American Economic Review*, Vol. 97, No. 5, pp. 1533-1559.

2. Amemiya, Takeshi (1974), "The nonlinear Two-Stage Least Squares Estimator", *Journal of Econometrics*, No. 2, pp. 105-110.

3. _______ (1977), "The maximum likelihood and the nonlinear three-stage least squares estimator in the general nonlinear simultaneous equation model", *Econometrica*, Vol. 45, No. 4, pp. 955-968.

4. Armstrong, Mark and Vickers, John (2010), "Competitive non-linear pricing and bundling", *The Review of Economic Studies*, No. 77, pp. 30-60.

5. Berges-Sennou, Fabian (2006), "Store loyalty, bargaining power and the private label production issue", *European Review of Agricultural Economics*, Vol. 33, No. 3, pp. 315-335.

6. Berges-Sennou Fabian, Bontems Phillip and Requillart V. (2009), "L'impact économique du développement des marques de distributeurs", *Working paper*, Toulouse School of Economics, 30 p.

7. Bernheim, Douglas and Whinston, Michael (1985), "Common marketing agency as a device for facilitating collusion", *RAND Journal of Economics*, Vol. 16, No. 2, pp. 269-281.

8. Berry, Steven (1994), "Estimating discrete-choice models of product differentiation", *The RAND Journal of Economics*, Vol. 25, No. 2, pp. 242-262.

9. Berry, Steven, Levinsohn, James and Pakes, Ariel (1995), "Automobile prices in market equilibrium", *Econometrica*, Vol. 63, No. 4, pp. 841-890.

10. _______ (1999), "Voluntary export restraints on automobiles: evaluating a trade policy", *The American Economic Review*, Vol. 89, No. 3, pp. 400-430.

11. Bettendorf, L. and Verboven, Frank (2000), "Incomplete transmission of coffee bean prices: evidence from The Netherlands", *European Review of Agricultural Economics*, Vol. 27, No. 1, pp. 1-16.

12. Bierlaire, Michel, Bolduc, Denis and McFadden, Daniel (2008), "The estimation of Generalized Extreme Value models from choice-based samples", *Transportation Research Part B: Methodological*, vol. 42, No.. 4, pp. 381-394.

13. Bolton, R., Kannan, P. and Bramlett, M. (2000), "Implications of loyalty program membership and service experiences for customer retention and value", *Journal of the Academy of Marketing Science*, Vol. 28, No. 1, pp. 95-108.

14. Bonnet, Céline and Dubois, Pierre (2010), "Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance", *RAND Journal of Economics*, Vol. 41, No. 1, pp. 139-164.

15. _______ (2015), "Identifying two part tariff contracts with buyer power: empirical estimation on food retailing", *Toulouse School of Economics Working Papers*, No. 575.

16. Bonnet, Céline, Dubois, Pierre, Villas-Boas, Sofia, and Klapper, Daniel (2013), "Empirical evidence on the role of nonlinear wholesale pricing and vertical restraints on cost pass-through", *The Review of Economics and Statistics*, Vol. 95, No. 2, pp. 500-515.

17. Caillaud, Bernard and De Nijs, Romain (2011), "Strategic loyalty reward in dynamic price discrimination", *Paris School of Economics Working Papers*, No. 2011-29, 38 p.

18. Caminal, Ramon and Matutes, Carmen (1990), "Endogenous switching costs in a duopoly model", *International Journal of Industrial Organization*, Vol. 8, No. 3, pp. 353-373.

19. Campa, Jose and Goldberg, Linda (2006), "Pass through of exchange rates to consumption prices: what has changed and why?", *NBER working Paper Series*, No. 12547, 40 p.

20. Chamberlain, Gary (1987), "Asymptotic efficiency in estimation with conditional moment restrictions", *Journal of Econometrics*, No. 34, pp. 305-334.

21. Chen, Yongmin and Pearcy, Jason (2010), "Dynamic pricing: when to entice brand switching and when to reward consumer loyalty", *The RAND Journal of Economics*, Vol. 41, No. 4, pp. 674-685.

22. Chen, Zhijun and Rey, Patrick (2012), "Loss leading as an exploitative practice", *The American Economic Review*, Vol. 102, No. 7, pp. 3462-3482.

23. _______ (2013), "Competitive Cross-Subsidization", *IDEI Working Papers*, No. IDEI-808.

24. Chevalier, Judith, Kashyap, Anil, and Rossi, Peter (2003), “Why Don’t Prices Rise during Periods of Peak Demand? Evidence from Scanner Data”, *American Economic Review*, Vol. 93, No. 1, pp. 15-37.

25. Chintagunta, Pradeep, Bonfrer, André and Song, Inseong (2002), “Investigating the effects of store-brand introduction on retailer demand and pricing behavior”, *Management Science*, Vol. 48, No. 10, pp. 1242-1267.

26. Crémer, Jacques (1984), “On the economics of repeat buying”, *The RAND Journal of Economics*, Vol. 15, No. 3, pp. 396-403.

27. Dubé, Jean-Paul, Fox, Jeremy and Su, Che-Lin (2012), “Improving the numerical performance of BLP static and dynamic aggregate discrete choice random coefficients demand estimation”, *Econometrica*, Vol. 80, No. 5, pp. 2231-2267.

28. Dubois, Pierre and Perrone, Helena (2010), “Price Dispersion and Search Costs: The Roles of Imperfect Information and Product Differentiation”, *unpublished working paper*.

29. _______ and Jódar-Rosell, Sandra (2010), “Price and Brand Competition between Differentiated Retailers: a Structural Econometric Model”, *IDEI Working Paper Series*, No. 604, 42p.

30. Eichenbaum, Martin, Jaimovich, Nir and Rebelo, Sergio (2011), “Reference prices, Costs, and Nominal Rigidities”, *American Economic Review*, No. 101, pp. 234-262.

31. Engel, Charles (2002), “Expenditure switching and exchange-rate policy”, *NBER Macroeconomics Annual*, Vol. 17, pp. 231-272.

32. European Commission (2011), “The impact of private labels on the competitiveness of the European food supply chain”, available at: http://ec.europa.eu/enterprise/sectors/food/documents/studies/comments_privlab_en.htm.

33. _______ (2014), *The economic impact of modern retail on choice and innovation in the EU food sector*, http://ec.europa.eu/competition/publications/KD0214955ENN.pdf.

34. Fudenberg, Drew and Tirole, Jean (2000), “Customer poaching and brand switching”, *The RAND Journal of Economics*, Vol. 31, No. 4, pp. 634-657.

35. Goldberg, Pinelopi and Hellerstein, Rebecca (2013), “A structural approach to identifying the sources of local currency price stability”, *Review of Economic Studies*, No. 80, pp. 175-210.

36. Goldberg, Pinelopi and Verboven, Frank (2005), “Market integration and convergence to the Law of One Price: evidence from the European car market”, *Journal of International Economics*, Vol. 65, No. 1, pp. 49-73.

37. Hausman, Jerry (1996), “Valuation of new goods under perfect and imperfect competitio”, *The Economics of New Goods*, Studies in Income and Wealth, Vol. 58, ed. by T. Bresnahan and R. Gordon, National Bueau of Economic Research, Chicago.

38. Hellerstein, Rebecca (2008), “Who bears the cost of a change in the exchange rate? Pass-through accounting for the case of beer”, *Journal of international Economics*, No. 76, pp. 14-32.

39. Hellerstein, Rebecca and Villas-Boas, Sofia (2010), “Outsourcing and Pass-through”, *Journal of International Economics*, No. 81, pp. 170-183.

40. Hong, Han and Shum, Matthew (2006), “Using price distributions to estimate search costs”, *RAND Journal of Economics*, Vol. 37, No. 2, pp. 257-275.

41. Honka, Elisabeth (2012), “Quantifying search and switching costs in the U.S. auto insurance industry”, *unpublished working paper*.

42. Hortaçsu, Ali and Syverson, Chad (2004), “Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 Index Funds”, *The Quarterly Journal of Economics*, Vol. 119, No. 2, pp. 403-456.

43. IGD (2015), *Key Industry Facts*, http://www.igd.com/About-us/Media/Key-industry-facts/.

44. Jorgenson, Dale and Laffont, Jean-Jacques (1974), “Efficient estimation of nonlinear simultaneous equations with additive disturbances”, *Annals of Economic and Social Measurement*, 3/4, pp. 615-640.

45. Jullien, Bruno and Rey, Patrick (2007), “Resale price maintenance and collusion”, *RAND Journal of Economics*, Vol. 38, No. 4, pp. 983-1001.

46. Kehoe, Patrick and Midrigan, Virgiliu (2012), “Prices are sticky after all”, *Research Department Staff Report*, No. 413, 41 p.

47. Kelejian, Harry (1971), “Two-stage least squares and econometric systems linear in parameters but nonlinear in endogenous variables”, *Journal of the American Economic Association*, Vol. 66, No. 334, pp. 373-374.

48. Klemperer Paul (1987a), “Markets with consumer switching costs”, *The Quarterly Journal of Economics*, Vol. 102, No. 2, pp. 375-394.

49. _______ (1987b), "The competitiveness of markets with switching costs", *The RAND Journal of Economics*, Vol. 18, No. 1, pp. 138-150.

50. Klemperer, Paul (1992), "Equilibrium product lines: competing head-to-head may be less competitive", *The American Economic Review*, Vol. 82, No. 4, pp. 740-755.

51. Klemperer, Paul and Padilla, Jorge (1997), "Do firms' product lines include too many varieties?", *RAND Journal of Economics*, Vol. 28, No. 3, pp. 472-488.

52. Knittel, Christopher and Metaxoglou, Konstantinos (2012), "Estimation of random coefficient demand models: two empiricists' perspective", unpublished manuscript.

53. Lal, Rajiv and Bell, David (2003), "The impact on frequent shopper programs in grocery retailing", *Quantitative Marketing and Economics*, No. 1, pp. 179-202.

54. Lederman, Mara (2007), "Do enhancements to loyalty programs affect demand? The impact of international frequent flier partnerships on domestic airline demand", *RAND Journal of Economics*, Vol. 38, No. 4, pp. 1134-1158.

55. Leibtag, Ephraim, Nakamura, Alice, Nakamura, Emi and Zerom, David (2007), "Cost pass -through in the U.S. coffee industry", *USDA Economic Research Reports*, No. 38, 28 p.

56. Lewis, Michel (2004), "The influence of loyalty programs and short-term promotions on customer retention", *Journal of Marketing Research*, Vol. 41, pp. 281-292.

57. Manski, Charles and Lerman, Steven (1977), "The estimation of choice probabilities from choice based samples", *Econometrica*, Vol. 45, No. 8, pp. 1977-1988.

58. McFadden, Daniel (1989), "A method of simulated moments for estimation of discrete response models without numerical integration", *Econometirca*, Vol. 57, No. 5, pp. 995-1026.

59. _______ (1999), "Computing willingness-to-pay in random utility models", in J. Moore, R. Riezman, and J. Melvin, eds., *Trade, Theory and Econometrics: Essays in Honour of John S. Chipman*, Routledge, London, pp. 253?274.

60. Midrigan, Virgiliu (2011), "Menu costs, multiproduct firms and aggregate fluctuations", *Econometrica*, Vol. 79, No. 4, pp. 1139-1180.

61. Moraga-Gonzalez, Jose Luis, Sandor, Zsolt and Wildenbeest, Matthijs (2011), “Semi-Nonparametric estimation of consumer search costs”, *unpublished working paper.*

62. Nakamura, Emi and Zerom, Dawit (2010), “Accounting for incomplete pass-through”, *Review of Economic Studies*, No. 77, pp. 1192-1230.

63. Nasser, Sherif, Turcic, Danko and Narisimhan, Chakravarthi (2013), “National Brand’s Response to Store Brands: Throw In the Towel or Fight Back?”, *Marketing Science*, Vol. 32, No. 4, pp. 591-608.

64. Nevo, Aviv (2000a), “A practitioner’s guide to estimation of random-coefficients Logit models of demand”, *Journal of Economics and Management Strategy*, Vol. 9, No. 4, pp. 513-548.

65. _______ (2000b), “Mergers with differentiated products: the case of the ready-to-eat cereal industry”, *RAND Journal of Economics*, Vol. 31, No. 3, pp. 395-421.

66. _______ (2001), “Measuring market power in the ready-to-eat cereal industry”, *Econometrica*, Vol. 69, No. 2, pp. 307-342.

67. Nevo, Aviv and Wolfram, Catherine (2002), “Why do manufacturers issue coupons? An empirical analysis of breakfast cereals”, *RAND Journal of Economics*, Vol. 33, No. 2, pp. 319-339.

68. Newey, Whitney (1990), “Efficient instrumental variables estimation of nonlinear models”, *Econometrica*, Vol. 58, No. 4, pp. 809-837.

69. OECD (2005), “Resale below cost”, *Policy Roundtables*, 231p.

70. Pakes, Ariel, Porter, J., Ho, Kate and Ishii, Joy (2011), “Moment inequalities and their application”, Mimeo.

71. _______ and Pollard, David (1989), “Simulation and the asymptotics of optimization estimators”, *Econometrica*, Vol. 57, No. 5, pp. 1027-1057.

72. Parks, Richard (1974), “The demand and supply of durable goods and durability”, *The American Economic Review*, Vol. 64, No. 1, pp. 37-55.

73. Peltzman, Sam (2000), “Prices rise faster than they fall”, *Journal of Political Economy*, Vol. 108, No. 3, pp. 466-502.

74. Raju, Jagmohan, Sethuraman, Raj and Dhar, Sanjay (1995), “The introduction and performance of store brands”, *Management Science*, Vol. 41, No. 6, pp. 957-978.

75. Rey, Patrick and Vergé, Thibaud (2010), "Resale price maintenance and interlocking relationships", *The Journal of Industrial Economics*, Vol. 58, No. 4, pp. 928-961.

76. Reynaert, Mathias and Verboven, Frank (2014), "Improving the performance of the random coefficients demand models: the role of optimal instruments", *Journal of Econometrics*, Vol. 179, No. 1, pp. 83-98.

77. Schemalensee, Richard (1978), "Entry deterrence in the Ready-to-Eat Breakfast cereal Industry", *The Bell Journal of Economics*, Vol. 9, No. 2, pp. 305-327.

78. Schiraldi, Pasquale, Seiler, Stephan and Smith, Howard (2011), "Supermarket choice with multistore shopping: measuring the effect of format regulation", *unpublished working paper.*

79. Slade, Margaret (1998), "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices", *Review of Economic Studies*, Vol. 65, No. 1, pp. 87-107.

80. Soberman, David and Parker, Phillip (2006), "The economics of quality-equivalent store brands", *International Journal of Research in Marketing*, No. 23, pp. 125-139.

81. Steenkamp, Jan-Benedict and Geyskens, Inge (2013), "Manufacturer and Retailer Strategies to Impact Store Brand Share: Global Integration, Local Adaptation, and Worldwide Learning", *Marketing Science*, forthcoming.

82. Stigler, George (1961), "The Economics of Information", *The Journal of Political Economy*, Vol. 69, No. 3, pp. 213-225.

83. Train, Kenneth (2009), *Discrete Choice Methods with Simulation*, Second edition, Cambridge University Press.

84. Wildenbeest, Matthijs (2011), "An empirical model of search with vertically differentiated products", *RAND Journal of Economics*, Vol. 42, No. 4, pp. 729-757.

85. Wooldrige, Jeffrey (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

86. Zhou, Jidong (2012), "Multiproduct search", *unpublished working paper.*

Résumé

Cette thèse est composée de trois essaies portant sur l'analyse empirique de la grande distribution et le comportement d'achat des consommateurs. Le premier chapitre est dedié à l'étude des programmes de fidélité des supermarchés et leur impact sur la demande de marques de distributeur (MDD). Souvent les supermarchés lient les avantages fidélité à l'achat en marques de distributeurs, quelles sont les motivations des supermarchés à faire cela? C'est la question que cet étude cherche à répondre d'un point de vue empirique. Je travail sur des donnés extraites d'un panel représentatif des consommateurs concernant les achats des ménages français, et l'utilisation d'une méthode structurelle d'estimation de demande. Les résultats sont conformes aux faits: les MDD sont des produits moins préférés vis-à-vis les marques nationales (MN) de même qualité. Cependant, la carte fidélité a, en effet, un impact positive sur le choix du consommateur: ceux qui portent une carte fidélité ont une probabilité supérieur de choisir des MDD que ceux qui ne l'ont pas. Par ailleurs, l'impact d'un programme de fidélité sur la demande des MDD est moins importants chez les détenteurs de plusieurs cartes.

Le deuxième chapitre, co-écrit avec Daniel Herrera Araujo, vise à mesurer les coûts d'achat des consommateurs à partir des données de panel concernant les achats des ménages en France. Quand l'analyse économique tient compte des coûts d'achat, que rationalisent l'hétérogénéité observée en nombre des enseignes visitées par les consommateurs, les conclusions de politique publique peuvent changer remarquablement. Nous identifions les coûts d'achat du consommateur dans le cadre d'un modèle structurel de demande en plusieurs enseignes ainsi qu'en plusieurs produits, qui lie le choix optimale du nombre de supermarchés à visiter (un seule ou plusieurs) aux coûts d'achat. Nous estimons les paramètres du modèle et mesurons le coût d'achat total moyen en 18,7 € par enseigne visitée. Deux quantités y sot compris: le coût fixe moyen, 1,53 € et le coût de transport moyen 17,1 € par visite.

Le troisième chapitre porte sur l'analyse empirique du rôle des tarifs binômes et la fixation du prix de vente (RPM, d'après l'expression anglo-saxonne Resale Price Maintenance) dans la stabilité des prix de vente. Il est largement reconnu dans la littérature économique que la transmission incomplète des chocs de coûts aux prix de vente est expliquée par l'ajustement des marges ainsi que les coûts d'ajustement des prix. Les relations entre fournisseurs et distributeurs et le RPM peuvent renforcer la rigidité des prix. Je présente un modèle structurel de relations verticales dont des tarifs binômes peuvent être adoptées ainsi que le RPM. Ce modèle tient compte de la rigidité des prix de vente à travers des coûts fixes d'ajustement des prix qui sont ajoutés au profit du détaillant. En utilisant des données concernant les ventes de marques de céréales pour le petit déjeuner dans une grande chaîne des supermarchés au Chicago, j'estime la demande, récupère les marges et calcule les limites supérieure et inférieure de l'intervalle que contienne les vrais coûts d'ajustement. Les résultats obtenus montrent que ces coûts représentent en moyenne entre 1.6% et 3% des revenus totales du distributeur par an.

Abstract

This dissertation consists of three essays on the empirical analysis of grocery retailing and consumer shopping behavior. The first chapter focuses on supermarket loyalty programs and their impact on the demand for private labels. Supermarkets often link loyalty rewards to private label purchases, What are supermarkets' motivations to do this? This empirically examines this link using scanner data on grocery purchases of French households and structural methods of demand estimation. Results are consistent with the industry lore: private labels are less valued products relative to quality-equivalent national brands. However, members of loyalty programs have a larger valuation of private labels than non-members. Moreover, the more prone to subscribe to LPs a customer is, the larger her sensitivity to a price increase and the weaker the expected effects on the demand for private labels.

The second chapter, joint with Daniel Herrera Araujo, is inspired by a number of theory papers showing that when shopping costs, that rationalize the observed heterogeneity in consumer shopping patterns, are introduced in economic analysis, policy conclusions can change dramatically. We structurally identify consumer shopping costs using scanner data on grocery purchases of French households. We present a model of demand for multiple stores and products consisting of an optimal stopping problem in terms of individual shopping costs. This rule determines whether to visit one or multiple stores at a shopping period. We then estimate the parameters of the model and recover the distribution of shopping costs. We quantify the total shopping cost in 18.7 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 1.53 € and mean total transport cost of 17.1 € per trip.

The third chapter empirically examines the role of nonlinear contracts between manufacturers and retail stores, and Resale Price Maintenance (RPM) on nominal price stability. According to the literature the incomplete transmission of costs shocks into retail prices is explained by the existence of markup adjustment and price adjustment costs. The vertical conduct of the industry and the use of RPM can introduce further price stickiness or reinforce it. I present a structural model of vertical relations allowing for two-part tariffs and RPM, and accounting explicitly for retail price rigidity by including fixed costs of price adjustment in retailer's profit function. Using micro data on sales of breakfast cereals from a large supermarket chain in Chicago, I estimate demand, retrieve margins, and compute bounds for retail price adjustment costs. I find that these costs lie between 1.6% and 3% of its total revenue a year, on average.