# Multiproduct retailing and consumer shopping behavior: The role of shopping costs

Jorge Florez-Acosta and Daniel Herrera-Araujo*

**Abstract**

We empirically examine the role of shopping costs in consumer shopping behavior in a context of competing differentiated supermarkets that supply similar product lines. We develop and estimate a model of demand in which consumers can purchase multiple products from multiple stores in the same week, and incur transaction costs of dealing with supermarkets. We compare predicted substitution patterns by our model with those derived from a similar model without shopping costs. We find that the latter predicts a larger proportion of multistop shoppers and overestimates own-price elasticities and product markups. Further, we simulate a situation in which a supermarket delists one of the products while rivals keep selling it. The supermarket strategically decreases the price of products in the same category while it increases the prices of products in other categories to keep the value of the basket constant. As a consequence, the supermarket loses some demand to rivals and bears a decrease of about 3.4% of weekly revenue. In the absence of shopping costs, only the price of substitutes change and losses are lower.

**Keywords:** Supermarket competition, market power, multistop shopping, shopping costs, product delisting.

**JEL Codes:** D12, L13, L81.

**Acknowledgments:** Special thanks to Bruno Jullien, Pierre Dubois, and James Hammitt for their guidance and support. We thank Patrick Rey, François Salanié, Andrew Rhodes, Michelle Sovinsky, Martin Peitz, Rachel Griffith, Andrea Pozzi, Helena Perrone, Juan Carranza, Tiago Pires, Alessandro Iaria and Laura Lasio for their helpful comments and discussions. This work has also benefited from comments and suggestions by seminar participants at the Paris School of Economics, ParisTech Telecom, INRA-SMART, CREST, the Institute for Fiscal Studies, Universidad del Rosario, Banco de la República, and conference participants at LACEA-LAMES meetings, the Econometric Society World Congress, the European Economic Association World Congress, the EARIE Conference, the IIO Conference, and the ENTER MaCCI-TSE Competition Economics Workshop in Mannheim. We are grateful to the *Institut National de la Recherche Agronomique*, INRA, for providing us with access to their data. All errors are our own.

*Florez: Department of Economics, Universidad del Rosario, Calle 12C 4-59, Bogotá, Colombia (e-mail: jorge.florez@urosario.edu.co); Herrera: Université Paris-Dauphine, LEDA (CGEMP), Place du Maréchal de Lattre de Tassigny, 75016, (e-mail: daniel.herrera@dauphine.psl.eu).

# 1 Introduction

The modern grocery retail industry is dominated by a small number of powerful large-scale supermarket chains[1] that attempt to entice customers to favor one-stop shopping through aggressive non-price strategies such as: 1) the proliferation of superstores with huge floor areas (20,000+ sq. meters) that offer a large product range (200,000+ different brands)[2], 2) the joint location with suppliers offering parallel services (e.g. shopping malls, beauty salons, restaurants, car wash facilities, gas stations, and playgrounds for children.), 3) the promotion of private labels (PLs), which makes supermarkets less dependent on branded products (the so-called national brands –NBs) and induces consumer loyalty, and 4) the increasing emphasis on strategies that induce consumer retention and reinforce store loyalty, such as loyalty programs.

In such a context, it becomes really important to understand the role consumers play in the way a set of differentiated supermarkets offering similar product lines interact. In particular, whether customers prefer concentrating purchases with a single supermarket or sourcing multiple store chains in the same period, the transaction and opportunity costs associated to shopping activities (so-called *shopping costs*),[3] and how these costs and shopping patterns shape the way consumers substitute across products and supermarkets are key features that determine how effective are supermarket price and non-price strategies and the exercise of market power. Precisely, a number of theoretical papers on multiproduct retailing shows that allowing customers incur heterogeneous shopping costs in the analysis of multiproduct demand and supply may change policy conclusions dramatically.[4]

Our contribution in this article is to study the role of shopping costs in explaining consumer choice of multiple products and multiple shopping locations, and in the measurement of supermarkets' market power. To this end, we develop a multiple-discrete choice model in the context of competition between supermarkets that offer the same product line to the same customers. Consumers can purchase baskets of products from either a single store (*one-stop shopping*) or

[1] For instance, in 13 member countries of the European Union (including France, the UK, Germany, Sweden, Finland, and the Netherlands), the five leading retailers in each country have a combined market share of more than 60% (European Commission, 2014).

[2] In Europe, for instance, between 2000 and 2011 the number of hypermarket outlets increased by 72% and the sales area increased by 46%, while the corresponding figures for supermarket outlets were 10% and 26%, respectively (European Commission, 2014).

[3] See Klemperer (1992); Klemperer and Padilla (1997); Armstrong and Vickers (2010); Chen and Rey (2012, 2013).

[4] If customers face positive shopping costs, competition with homogeneous product lines tend to soften competition because consumers will stick with a single retailer as the benefit from visiting an additional supplier need not compensate the cost. By contrast, if product lines are differentiated, retailers may be tempted to undercut prices to make one-stop shoppers source several separate suppliers (Klemperer, 1992). Further, costly shopping may lead to the introduction of too many varieties of products with respect to the social optimum given that more consumers prefer to concentrate purchases with the retailer supplying a wider product range and save on shopping costs (Klemperer and Padilla, 1997). Moreover, below-cost pricing practices appear to be exploitative rather than predatory once shopping costs are accounted for. Large retailers can take advantage of the fact that consumers have heterogeneous shopping costs and adopt loss-leading strategies to price discriminate between one- and multi-stop shoppers. From this perspective, it is more profitable to keep rivals in the market so long as some customers prefer favoring multi-stop shopping (Chen and Rey, 2012). Finally, in a context 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).

multiple stores (*multistop shopping*) during a given period.[5,6] Our key modeling strategy is to explicitly account for this observed heterogeneity by introducing consumer transaction costs related to shopping. Following Klemperer (1992), we define *shopping costs* in a comprehensive way as all of the consumer's real or perceived costs of using a supplier.[7] These may include transportation costs and opportunity costs related to time spent parking, selecting the products in the store, and waiting in line at the checkout; they may as well account for the taste for shopping (Chen and Rey, 2012).

Our general empirical strategy is to estimate basket-level demand using standard techniques from the discrete-choice literature, along with simulated methods. We specify the utility of each product as a function of observed and unobserved product and store characteristics, as well as parameters to be estimated. On every shopping occasion, each consumer faces idiosyncratic shopping costs that increase with the number of supermarkets visited. Each consumer weighs up the extra benefits of dealing with an additional store against the additional costs involved. If benefits exceed costs, the individual will visit an additional supermarket. Otherwise, she will make all her purchases at a single place. The total utility of a basket of products is the sum of the product-specific utilities minus the shopping costs. To consistently estimate the parameters of the model, we have to deal with a challenge: shopping costs vary across individuals and are unobserved (by the econometrician). We deal with this by decomposing shopping costs into two components: a mean shopping cost, which is common to all consumers, and an idiosyncratic deviation to the mean cost which depends on both observed demographic characteristics and a random shock, which is known to consumers and assumed to follow a known parametric distribution. This shock captures all individual (unobserved) characteristics that cause individual costs to differ from the average shopping cost.

Once the parameters of the model are estimated, we perform two exercises that allow us to assess the relative importance of accounting explicitly for shopping costs in predicting reasonable substitution patterns in a multistop shopping environment. In a first exercise, we take our estimated model and simulate a scenario in which shopping costs fall to zero, i.e. we assume that consumers no longer incur positive shopping costs. In a second exercise, we estimate an

[5]Previous papers have developed demand models for multiple products. Some examples are Hendel (1999), who develops a multiple-discrete choice model to explain how firms choose multiple multiple alternative brands of personal computers. Dubé (2004) applies Hendel's model to the case of carbonated soft drinks given that, according to the evidence provided by the author, consumers commonly buy multiple alternatives on each shopping trip. Gentzkow (2007) develops a flexible framework in which similar products can be either substitutes or complements. None of these studies incorporate consumer transaction costs into the choice problem. Wildenbeest (2011) sets out a search cost model in which consumers are interested in a basket of products and care about the total price of the basket. They must purchase the entire basket at a single store (i.e. one-stop shopping). Finally, Thomassen et al. (2017) develop a model of demand in which consumers make discrete–continuous choices over multiple categories. Consumers can purchase from up to two stores in each period and incur a choice-specific fixed cost of shopping.

[6]In our data we observe this heterogeneity throughout the year. On average, in a given week barely 55% of households concentrate purchases at a single store, whereas the remaining households make purchases at two (around 30%), three (around 10%) and four (<5%) rival stores (see figure 1 in Appendix 2.2).

[7]Klemperer (1992) distinguishes among consumer costs in the following way: "(...) a 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.

alternative specification that does not include any shopping costs, and compare the results with those obtained from our preferred specification with shopping costs.

Furthermore, we use our model and estimates to empirically examine a related question in which shopping costs play a key role: how do consumers and rival supermarkets react when a supermarket removes a particular product from its shelves? This situation is similar to *product delisting*, a phenomenon mainly studied by the marketing literature (see, for example, Davies (1994)) but that has recently brought the attention of competition authorities. Delisting of products can happen purely for commercial reasons[8] or it can be used by the supermarket in a strategic way to impose restraints on manufacturers. In any case, this practice can entail losses for the delisting supermarket if a high proportion of its customers face high shopping costs, as they may be tempted to switch to rival stores in order to get the whole basket they are looking for.[9] We simulate a large price increase in one of the products sold by a given supermarket so that it becomes prohibitively costly for consumers, while the same product continues being supplied by competing supermarkets at observed prices. We measure the net effect of such price increase on demand and supply by allowing supermarkets to adjust prices to a new equilibrium. We do this under two scenarios: one in which consumers face positive shopping costs and an alternative scenario in which shopping costs are zero.[10]

Perhaps the biggest limitation of our approach is the dimensionality problem that arises when estimating demand for both baskets of products and multiple shopping locations. In our data set, we observe households that purchase up to 275 different products from up to nine separate grocery stores in the same week.[11] Estimating a demand system with such a huge choice set is infeasible. We deal with this as follows. First, we restrict our focus to three categories of products that are staple food items, among the most frequently purchased, and usually subject to unit demand. The categories that best meet these criteria in our data are yogurt, biscuits, and refrigerated desserts. Second, we aggregate brands to the category level so that we end up with a reduced set of composite products. To evaluate the demand-side effects of product delisting, we allow for two alternatives of yogurt, namely, the leading national brand (NB) in France in 2005, and a composite yogurt "brand" that includes all of the remaining alternatives (both other NBs and private labels –PL). Therefore, consumers have a set of four products

[8]Such as low gross and net profit margins, low sales volume, low customer flow, insufficient shelf space, high handling and storage costs, a large increase in the wholesale price, a reduction in the number of suppliers, or significant intrabrand competition (Davies, 1994).

[9]For instance, in 2009 in Belgium, a request for a price rise by Unilever triggered the delisting of 300 of Unilever's products by Delhaize, one of the largest supermarket chains in that country. Both parties ended up being hurt: Delhaize lost 31% of its customers to rivals and among those who remained, 47% substituted other brands for Unilever's products. See http://in.reuters.com/article/delhaize-unilever-idINLG51937220090216.

[10]We are implicitly assuming in our analysis that wholesale prices and vertical bargaining are exogenous. However, competition authorities have documented that product delisting is often used strategically by supermarkets to impose restraints on manufacturers' optimal pricing (see, for example, OFT (1997), OECD (1998, 2008), European Commission (1999), and FTC (2001)). In this context, accounting for upstream firms' best responses to product delisting is key to fully capture the effects on both downstream and upstream equilibrium pricing, and demand. A structural model of both vertical and horizontal relationships along with a model of multiproduct demand with shopping costs is necessary. Such a model is out of the scope of this paper and is currently an ongoing project in our agenda.

[11]On average, a household purchases baskets containing 24 different products from two separate stores each week.

from which they can choose at most three: one of the two alternatives of yogurt, biscuits, and desserts.[12] Finally, we focus on a reduced set of three supermarket chains that were the leading grocery retailers in France in 2005 based on market share. The remaining stores in our data set, along with the no purchase of the included goods option are left as part of the outside option.[13]

We obtain several interesting results. First, from descriptive regressions, we find a significant relationship between the number of supermarkets visited by a household in a week and household characteristics that are a proxy for the opportunity cost of time. Second, our structural model allows us to retrieve consumer total shopping costs, which we estimate to be 37 euro cents per store visited, on average. This cost includes a fixed shopping cost of 33 euro cents per store visited and a transport cost of 4 euro cents per trip to a store located at the average distance from consumer location. Third, when we unilaterally set shopping costs to zero, our simulations indicate a higher frequency and intensity of shopping by customers: in the absence of shopping costs all consumers would visit at least one store every week with positive probability and, in particular, the probability of doing multistop shopping is similar to that of doing one-stop shopping; once shopping costs are accounted for, the predicted probabilities of both one- and multistop shopping are lower, and consumers are less likely to visit a supermarket on a week-to-week basis. Fourth, when we compare the substitution patterns predicted by our model with those of an alternative specification under the assumption that shopping costs are zero, we find that shopping costs reinforce the complementarities between product categories that emerge when customers are allowed to purchase baskets of products. Such complementarities can be thought of as the "economies of scope" of buying related products from a single supermarket, as discussed by Klemperer (1992). In fact, the cross-price elasticities predicted by a model with shopping costs are larger than those of the alternative specification, suggesting that shopping costs make those economies of scope more valuable to consumers.

Further, when we simulate the delisting of a product by one supermarket, we find that supermarkets strategically decrease the price of substitutes to encourage intra-store substitution. Conversely, we find that prices of complement products increase and offset the decrease in the substitute product, which suggests that an optimal strategy for supermarkets is to keep the overall value of the basket constant and retain one-stop shoppers. By contrast, in a scenario in which shopping costs are unilaterally set to zero the price of the close substitute still decreases but the prices of complement product remain unchanged. When consumers do not face fixed shopping costs, the economies of scope of purchasing a basket of products at a single store are

[12] The choice of yogurt as the category offering two alternatives is arbitrary.

[13] Of course, our data set indicates that consumers often purchase/visit more than the included products/stores. We take this into account in our empirical strategy. Following Gentzkow (2007), we assume that every basket includes a maximization over the included as well as the excluded (both observed and unobserved) products. In line with this, if a consumer is observed to have purchased two of the inside goods at the included supermarkets, it may be the case that those were the only products she purchased, or it may be that she purchased additional products and visited stores other than those included. Regardless, we assume that the utility of baskets that contain inside products purchased from any of the included locations is greater than that of alternative baskets containing any other combination of inside products. However, it is important to note that labeling some products in the outside option does not change the interpretation of substitution patterns among baskets containing included products. See Section 4.4 for a detailed discussion.

no longer present because inter-store substitution for each individual product is now possible at a lower cost (the transport cost only), consumers need not leave their usual store and, as a consequence, supermarkets lose less demand to rivals as compared to our base case. Hence, the delisting supermarket tries to keep its market share on yogurts unaffected by making the substitute product more attractive to consumers but leaving the prices of other products unchanged. Last, we find that, on average, a supermarket loses more revenue when consumers incur positive shopping costs (about 3% of total revenue) than in a context of zero shopping costs (about 1% of total revenue).

This paper relates to a growing body of empirical literature that models consumer choice problems accounting explicitly for opportunity costs associated with shopping activities. This literature has focused on two types of costs, namely, search costs (e.g. Hortaçsu and Syverson (2014), Hong and Shum (2006), Koulayiev (2014), Moraga-Gonzalez, Sandor and Wildenbeest (2013), Kim and Bronnenberg (2010), Wildenbeest (2011), De los Santos, Hortaçsu and Wildenbeest (2012), Honka (2014), and Dubois and Perrone (2015)) and switching costs (e.g. Shy (2002), Viard (2007), and Honka (2014)).[14] Less attention has been paid to shopping costs. Brief (1967) models consumer shopping patterns in a Hotelling framework, and basically estimates transportation costs to account for consumer shopping costs.[15] Aguiar and Hurst (2007) evaluate how households substitute time for money by optimally combining shopping activities with home production. Customers incur a time cost of shopping that is explicitly accounted for.

Finally, in analyzing multiproduct and multistore choice with shopping costs, our paper is closely related to that of Thomassen et al. (2017). They study pricing by grocery stores in the context of competition between specialized stores and supermarkets. To this end, they develop a model of demand for multiple products in which some consumers purchase from a single store, whereas others visit at most two stores in each period. To rationalize this heterogeneity, they introduce a choice-specific fixed cost to the utility function of a consumer. Our approach, which we developed contemporaneously and independently, differs from theirs in several important ways. First, our primary focus is on the role of shopping costs in predicting consumer substitution and shopping patterns. Second, we use our model to explore the effects of product delisting, and though we add structure to the supply side in order to it in a more realistic way, our main focus is on the demand side effects of such a practice. Third, our empirical analysis is oriented by a model that is in line with the theoretical literature on multiproduct retailing with shopping costs (see, in particular, Chen and Rey (2012, 2013)). In our setting, the number of stores visited by a consumer is endogenously determined by a stopping rule involving the extra utility and extra costs involved in visiting an additional store. This enables us to empirically identify the distribution of shopping costs. Last, but not least, while we are interested in analyzing

[14]As noted by Klemperer and Padilla (1997), shopping costs differ from switching costs in that the latter derive from the economies of scale from repeated purchases of a product while the former are associated with economies of scope of buying related products.

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

competition between supermarket chains of similar size and characteristics that supply the same product range to customers, they focus on competition between small specialized stores and large supermarkets.

The rest of the paper is organized as follows. Section 2 presents the data and a preliminary analysis of consumer shopping behavior based on descriptive statistics and reduced-form regressions. Section 3 outlines our structural model of multiproduct demand and consumer shopping behavior in the presence of shopping costs, as well as a supply model of supermarket oligopolistic competition. Section 4 describes our empirical strategy, discusses our identification strategy, reports the estimation results and describes the role of shopping costs in predicting substitution patterns and supermarket markups and marginal costs by comparing the market equilibrium with shopping costs with the counterfactual in which shopping costs are set to zero. We then present a rubustness analysis in which shopping patterns are defined in a different way, which gives rise to an alternative way of quantifying shopping costs. Finally, section 5 presents and discusses results of our counterfactual simulations with product delisting. Finally, section 6 concludes and discusses possible directions for further research.

# 2 Grocery retailing, shopping patterns and opportunity cost of time

## 2.1 Data overview

Data on household purchases were obtained from the Kantar Worldpanel database. This is homescan data relating to grocery purchases made by a representative sample of 10,000 randomly selected households in France during 2005. These data were collected by household members using scanning devices.[16] The data set contains information on 352 grocery product categories from approximately 90 grocery stores including supermarket chains, hard discounters, and specialized stores. An entry in the data set records the purchase of a specific product from a given store on a particular date. Further, the data set includes information on household characteristics.

We supplement the homescan data with information on supermarket characteristics from the Atlas LSA 2005. This includes information by store format and type (regular and hard-discount stores) on aspects including the store's location, sales area, number of checkouts, and number of parking bays. We merge both data sets using household data, the name of the retailer, the zip code of the consumer's residence, and the floor area of the outlet. We follow Dubois and Jódar-Rosell (2010) and compute distances between a household residence and a store using zip codes; with this distances in hand, we determine the outlet of each supermarket chain that is closest to the consumer's dwelling and include only one outlet per retailer (in case there are several stores meeting this criterion); further, we keep in the consumer's choice set stores located

[16] The Kantar Worldpanel is a continuous panel database that commenced in 1998. Most households that comprise the panel have been randomly sampled since 1998. Every year, new randomly selected households are added to the panel, either to replace other households that rarely report data or to increase the sample size.

within 20 kilometers from the household location.

In the descriptive analysis presented in this section (and in Appendix 2.2), we use the full data on household purchases of all products recorded by the consumers in the sample, which gives a data set of approximately 9.5 million observations. In the structural estimation section, we will focus on a reduced set of products and stores for reasons that we discuss bellow.

## 2.2 Household characteristics and shopping patterns

Table 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[17] being 49 years old, with around 2,366 € monthly income and at least one car. Only half of the households in the sample reported having internet access at home which partially explains why internet purchases are not important in our data. As for storage capacity and home production, 77% of households have storage rooms at home and 67% an independent freezer in addition to a refrigerator. Further, about 36% of households reported producing vegetables at home which, along with the fact that nearly 30% of the households are located outside urban areas, can be a reason for the observed low frequency of shopping of some households.

Table 1: Summary statistics for household characteristics

| Variable | Mean | Median | Sd | Min | Max |
|---|---|---|---|---|---|
| **Demographics** Household size | 3.00 | 3 | 1.39 | 1 | 9 |
| Income (€/month) | 2,366 | 2,100 | 1,115 | 150 | 7,000 |
| Expenditure on groceries (€/month) | 261 | 239 | 140 | 0.16 | 1200 |
| Baby (=1 if yes) | 0.18 | 0 | 0.38 | 0 | 1 |
| Household head's age | 49.26 | 47 | 14.31 | 17 | 98 |
| Lives in city | 0.74 | 1 | 0.44 | 0 | 1 |
| Car (= 1 if yes) | 0.93 | 1 | 0.24 | 0 | 1 |
| Internet access at home | 0.51 | 1 | 0.50 | 0 | 1 |
| **Storage capacity** | | | | | |
| Independent freezer | 0.67 | 1 | 0.47 | 0 | 1 |
| Freezer capacity $> 150L$ | 0.57 | 1 | 0.50 | 0 | 1 |
| Storage room at home | 0.77 | 1 | 0.42 | 0 | 1 |
| Vegetables production at home | 0.36 | 0 | 0.48 | 0 | 1 |

*Source:* Kantar Worldpanel database 2005. Authors' calculations.

Table 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 tends to do between one and two trips per week to the same store. The average number of days between shopping occasions is 5 days. The preferred store type remains the regular supermarket over the hard discounter: only a 18,6% of weekly visits to grocery stores are done to the latter type. Larger store formats are preferred by consumers: on average, the two most frequently visited store formats are Supermarkets and Hypermarkets with 54% and 42%

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

share on total visits per week, respectively. Convenience stores, the smaller shops supplying a reduced product range generally at higher prices, receives the lower number of visitors per week with 3.9%. Although convenience stores have the advantage of being within walking distance from households location, as opposed to hypermarkets that are located far from city centers, the preference for larger store formats can be explained by several factors such as bulk shopping, lower prices, more intensive sales and promotional activities and a larger product range.

Table 2: Summary statistics for household shopping patterns

| **Variable** | **Mean** | **Median** | **Sd** | **Min** | **Max** |
|---|---|---|---|---|---|
| No. Trips to same grocery store/week | 1.41 | 1 | 0.76 | 1 | 7 |
| No. separate grocery stores visited/week | 1.65 | 1 | 0.83 | 1 | 9 |
| Days between visits | 5.40 | 4 | 6.07 | 1 | 188 |
| Visits to Hard discounters (% of total/week) | 18.58 | 0 | 38.89 | 0 | 100 |
| **Visits by format (% of total/week)** | | | | | |
| Hypermarket | 41.84 | 35.21 | 34.52 | 0 | 100 |
| Supermarket | 53.81 | 57.03 | 34.39 | 0 | 100 |
| Convenience | 3.92 | 0 | 12.90 | 0 | 100 |

*Source:* Kantar Worldpanel database 2005. Authors' calculations.

## 2.3 The nature of multistop shopping

Scanner data on supermarket purchases by households is particularly well suited to investigate multistop shopping behavior, because supermarket chains with similar characteristics (size, variety of store formats, products and brands carried and location) often offer similar product ranges to consumers who could opt for concentrating purchases with their preferred supermarket chain. Yet, we observe that multistop shopping as a widespread pattern among households in our data. Figure 1 shows the distribution of households by the average number of supermarkets visited in a week. Of the total number of households we observe, about 47% visit more than one supermarket per week, on average. According to theory, this observed heterogeneity in shopping patterns can be rationalized by the fact that for some consumers is more costly to deal with multiple supermarkets.

To empirically check this, we regress the number of separate supermarkets visited in a week on on a set of household characteristics that proxy for households' time constraints. We expect to find that more time-constraint consumers (i.e. with higher opportunity cost of time) tend to visit less supermarkets than less time-constraint ones. We add some controls for household storage capacity (housing type, presence of a storage room and/or independent freezer, and the size of the largest freezer) that help to rationalize (at least in part) the frequency of shopping. Further, we include supermarket, region, and week dummies in all regressions. Table 3 shows the results. Coefficients are basically of the expected sign and statistically significant.

We find evidence suggesting that a household's ability to source multiple supermarkets is dependent on time constraints and distance to the stores. Higher-income households and those

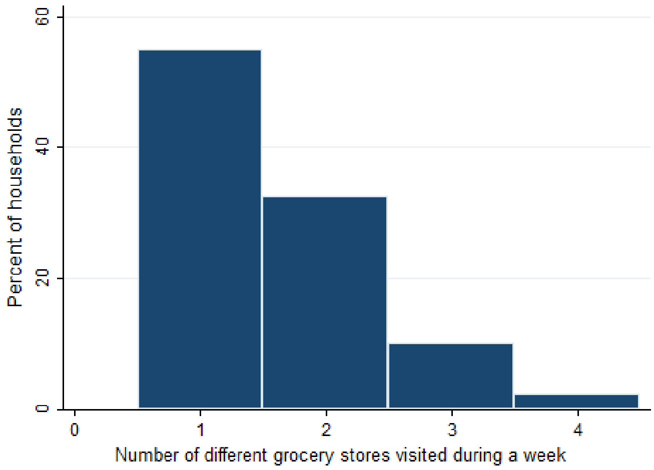



Figure 1: Distribution of households by average number of stores visited per week in 2005

*Notes:* The observed distribution has a longer tail than that displayed in the figure as some households visited up to nine separate supermarket chains per week. However, 99.4% of the observations involved between one and four stops.
*Source:* Kantar Worldpanel database 2005. Autors' calculatons.

with babies often visit less supermarkets on average, presumably because of a greater opportunity cost of time. Internet access reduces the number of shops visited, as people can shop online and use home delivery services, which might involve savings in terms of both transport costs and time.

Table 3: Number of stores visited per week[a]

| | OLS | | Poisson | |
|---|---|---|---|---|
| Variable | Coeffs. | Std. errors | Coeffs. | Std. errors |
| Hypermarket[b] | -0.080 | 0.028 | -0.051 | 0.018 |
| Supermarket[b] | -0.056 | 0.025 | -0.034 | 0.015 |
| Hard discounter (=1 if yes) | 0.198 | 0.024 | 0.130 | 0.016 |
| HH head's age | 0.003 | 0.000 | 0.002 | 0.000 |
| Log income | -0.021 | 0.012 | -0.014 | 0.008 |
| HH size | 0.061 | 0.006 | 0.041 | 0.004 |
| Car (=1 if yes) | -0.021 | 0.026 | -0.014 | 0.018 |
| Lives in urban areas (=1 if yes) | 0.064 | 0.014 | 0.042 | 0.010 |
| Lives in an apartment (=1 if yes) | 0.035 | 0.016 | 0.023 | 0.011 |
| Lives on a farm (=1 if yes) | -0.102 | 0.038 | -0.072 | 0.028 |
| Baby (=1 if yes) | -0.065 | 0.016 | -0.043 | 0.011 |
| Internet access at home (=1 if yes) | -0.012 | 0.013 | -0.009 | 0.009 |
| Home production (=1 if yes)[c] | -0.017 | 0.013 | -0.012 | 0.009 |
| Constant | 1.339 | 0.114 | 0.289 | 0.077 |
| $R^2$/Pseudo $R^2$ | 0.0641 | | 0.0074 | |

*Source:* Kantar Worldpanel database 2005. Authors' calculations.
*Notes:* [a]Standard errors are clustered by household. All specifications include the average distance between household location and the stores visited in a week, and controls for household storage capacity, and store, region, and week fixed effects.
[b] Proportion of visits to the respective store format to the total number of visits to stores that week.
[c] A household scores 1 if it grows vegetables at home and 0 otherwise.

The coefficients for housing type and store format suggest some patterns related to the physical structure of cities in France and store locations. In France, as a result of zoning regulations that limit store size depending on the zone of the city, larger store formats must be located farther away from the city centers. Accordingly, stores can be categorized in three formats depending on size and location: Hypermarkets are the stores with the largest sales areas and product range, are often only reachable by car and are located at a considerable distance from rival stores. Supermarkets are medium-sized stores with a fairly varied assortment that are generally closer to the city center than hypermarkets. And convenience stores, which are small stores widely present in the downtown areas and are easily reachable, but are generally small and only offer a limited range of products (mostly staples), which makes them suitable for top-up trips.[18]

We included dummy variables for two of the three store formats, with hypermarkets and supermarkets each taking a value of 1 if that was the format visited. The coefficients obtained are negative and significant in both cases, and consistent with economic intuition: given that hypermarkets and supermarkets are larger than convenience stores and carry a larger product range, consumers who source them need to make less shopping trips than those patronizing convenience stores because the larger stores make bulk shopping possible. The coefficient of distance from the home location to the store shows a positive correlation with the number of stores visited in all regressions. We interpret this as people making top-up trips to convenience stores during the week, but going to a hypermarket or supermarket to do a bulk shop.

As for housing type, our results suggest that people living in apartments, which are more likely to be in or closer to downtown areas, tend to visit a larger number of stores than those who live in houses. Meanwhile, those who live on farms visit less shops than families living in smaller types of accommodation. An alternative interpretation that is consistent with this result is related to household storage capacity. Households with lower storage capacity, i.e. those living in apartments, need to visit shops more often, and thus are more likely to be multistop shoppers.

# 3 The model

## 3.1 Consumer choice model

There are $I$ consumers in the market indexed by $i = 1, \ldots, I$ with idiosyncratic valuations of grocery products indexed by $k = 1, \ldots, K$. Suppose there are three store chains in the market indexed by $r \in \{A, B, C\}$ that supply the same products to all consumers.[19] Customer $i$ purchasing product $k$ from store $r$ in period $t$ derives a net utility of $\overline{v}_{ikrt}$, which is a function of the price of the product and other characteristics.[20]

[18] Officially, store formats are sorted according to their sales areas: hypermarkets have a sales area of 2500m$^2$ or more, supermarkets are between 400 m$^2$ and 2500 m$^2$, and convenience stores are less than 400 m$^2$

[19] Assuming that all consumers have access to the same product range might appear unreasonable. However, this helps us to reduce dimensionality issues in the estimation of the model. An extension of the model would be to relax this assumption and allow for heterogeneous choice sets.

[20] For now, we do not specify a functional form for the product-level utility, as it is not necessary for setting out the model. We will assume a parametric specification at the empirical implementation stage in Section 4.

Consumers have unit demand for each product class and can purchase one, two, or three products in the same period. Let $\mathcal{B}$ be the set of all exclusive and exhaustive baskets. Baskets with multiple products may be purchased from a single store (*one-stop shopping*) or from multiple stores (*multistop shopping*). A consumer favors multistop shopping if her shopping costs are sufficiently small, otherwise she will optimally make her purchases from a single store.

In the formulation of the model, we focus on the fixed component of the total shopping costs that may account for the consumer's taste for shopping. From now on, we will refer to this fixed cost as "shopping costs" and denote it as $s_i$. Transport costs, which are an important component of the total cost of shopping, are accounted for by including distance to stores as an additive term to the utility function of a basket of products (see below). Accordingly, shopping costs are assumed to be independent of store characteristics (e.g. size, facilities, location) and time invariant. Furthermore, we assume that $s_i$ is randomly drawn from a continuous distribution function $G(\cdot)$ and positive density $g(\cdot)$ everywhere. Finally, we assume that consumers are well informed regarding prices and product characteristics. Therefore, consumers do not need to engage in costly search to gather information about prices and product quality.

Consumer $i$ is supposed to exhibit optimal shopping behavior. This implies that she makes an optimal choice involving two elements: whether to be a one- or multistop shopper, and which stores to visit for each of the products she wants to buy. Roughly speaking, the choice set of consumer $i$ will be restricted by the number of stores she can visit given her shopping costs, so that her choice will consist of selecting the mix of products and stores that maximize the overall value of the desired basket. In line with this, a three-stop shopper who can visit all stores and wants the three products will select the best product–store combination from the alternatives existing in the market within each category. A two-stop shopper will select the mix of two stores maximizing the utility of the desired basket from all possible product–store combinations. Her final basket will consist of the best of the two alternatives in each product category. Finally, a one-stop shopper will pick the store offering the largest overall value of the whole basket of products.

Formally, let $D_{ir}$ for all $r \in \{A, B, C\}$ denote the distance traveled by consumer $i$ from his household location to store $r's$ location, and $\tau$ denote a parameter that captures the consumer's valuation of the physical and perceived costs of traveling that distance. We define the utility net of transport costs of a shopper who is able to visit only one of the three stores in the market as follows:[21]

$$v_{it}^{1} = \max\left\{\sum_{k=1}^{K}\overline{v}_{ikAt} - \tau D_{iA}, \sum_{k=1}^{K}\overline{v}_{ikBt} - \tau D_{iB}, \sum_{k=1}^{K}\overline{v}_{ikCt} - \tau D_{iC}\right\}. \tag{1}$$

Similarly, the net utility of a two-stop shopper is given by

[21] Note that the utilities below depend on the vector of all prices of products sold by the 3 stores in the market, which we denote by $\mathbf{p}_t$. However, we omit this for ease of presentation.

$$v_{it}^2 = \max\left\{\sum_{k=1}^{K}\max\{\overline{v}_{ikAt},\overline{v}_{ikBt}\} - \tau(D_{iA}+D_{iB})\ ,\right.$$
$$\sum_{k=1}^{K}\max\{\overline{v}_{ikAt},\overline{v}_{ikCt}\} - \tau(D_{iA}+D_{iC}), \quad (2)$$
$$\left.\sum_{k=1}^{K}\max\{\overline{v}_{ikBt},\overline{v}_{ikCt}\} - \tau(D_{iB}+D_{iC})\right\}.$$

Finally, the net utility of a consumer who is able to visit all three stores is given by

$$v_{it}^3 = \sum_{k=1}^{K}\max\left\{\overline{v}_{ikAt},\overline{v}_{ikBt},\overline{v}_{ikCt}\right\} - \sum_{r\in\{A,B,C\}}\tau D_{ir}. \quad (3)$$

Note that the expressions in (1), (2), and (3) are particular cases of a more general utility function in which, conditional on shopping costs, an $n$-stop shopper is selecting the subset of stores that maximizes the overall utility of her desired basket. 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 precisely the number of stores in the market, which is why she does not need to maximize over subsets of supermarkets.[22]

Suppose $v_{it}^1 - s_i > 0$ such that all consumers will visit at least one supermarket in each period. To determine the number of stops to be made, consumer $i$ weighs the extra utility of undertaking $n$-stop shopping with the extra costs involved, taking into account the fact that the total cost of shopping increases with the number of stores visited. Let $\delta_{it}^2 \equiv v_{it}^2 - v_{it}^1$ and $\delta_{it}^3 \equiv v_{it}^3 - v_{it}^2$ be the incremental utilities that consumer $i$ derives from visiting, respectively, two stores rather than one and three stores rather than two. A consumer might as well consider visiting either one or three stores in which case her incremental utility will be given by $v_{it}^3 - v_{it}^1 = \delta_{it}^2 + \delta_{it}^3$.

Consumer $i$ will optimally decide to undertake three-stop shopping only if the net utility derived from visiting three stores is greater than that from either one- or two-stop shopping. Formally,

$$v_{it}^3 - 3s_i \geqslant \max\{v_{it}^2 - 2s_i, v_{it}^1 - s_i\}.$$

Rearranging, the optimal stopping rule for a three-stop shopper is given by

$$s_i \leqslant \min\left\{\delta_{it}^3, \frac{\delta_{it}^2 + \delta_{it}^3}{2}\right\}. \quad (4)$$

Similarly, consumer $i$ will optimally decide to undertake two-stop shopping if and only if

$$v_{it}^2 - 2s_i \geqslant \max\{v_{it}^1 - s_i, v_{it}^3 - 3s_i\}.$$

[22] The general expression of the utility and choices of an $n$-stop shopper are described in Appendix A.

Hence, consumer $i$ will undertake two-stop shopping as long as

$$\delta_{it}^3 < s_i \leqslant \delta_{it}^2. \tag{5}$$

Finally, consumer $i$ will optimally decide to undertake one-stop shopping if and only if

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

from which we can derive the optimal stopping rule for a one-stop shopper as follows:

$$s_i > \max\left\{\delta_{it}^2, \frac{\delta_{it}^2 + \delta_{it}^3}{2}\right\}. \tag{6}$$

In general, the optimal stopping rule for consumer $i$ indicates that she will choose the mix of stores that maximizes her utility conditional on the extra shopping cost being at most the extra utility obtained from visiting additional stores. Note that equations (4), (5), and (6) imply that

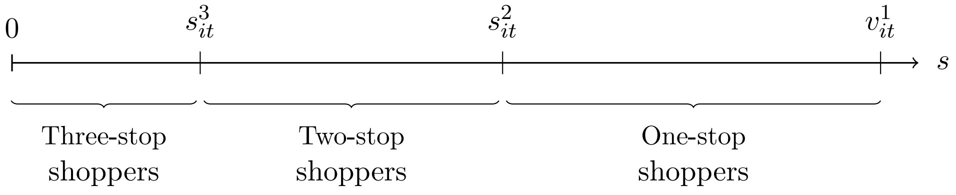
$$\delta_{it}^3 < \frac{\delta_{it}^2 + \delta_{it}^3}{2} < \delta_{it}^2. \tag{7}$$

Therefore, the highest possible shopping costs for any consumer able to undertake multistop shopping at either two or three stores, respectively, in equilibrium are given by the following critical cutoff points:

$$\begin{aligned} s_{it}^2 &= \delta_{it}^2, && \text{for two-stop shopping, and} \\ s_{it}^3 &= \delta_{it}^3, && \text{for three-stop shopping.} \end{aligned} \tag{8}$$

Note that these cutoff points depend on the period of purchase. The subscript $t$ was added because it depends on utilities that may vary with time. The derived cutoffs for the distribution of shopping costs in (8) indicate that for given shopping costs, consumers only care about the marginal utility of visiting an additional store in making their final decision on how many stores they should visit. Moreover, one-, two-, and three-stop shopping patterns arise in equilibrium and will be defined over the entire support of $G(\cdot)$ (see Figure 2).[23]

Figure 2: One-, two-, and three-stop shopping




[23]Note that the kind of behavior according to which a shopper evaluates extreme choices such as visiting all shops rather than only one does not appear to be relevant here.

### 3.2 Aggregate demand

Let $\mathcal{B}_2, \mathcal{B}_3 \in \mathcal{B}$ be subsets of baskets involving two- and three-stop shopping, respectively. The aggregate demand for product $k$ supplied by store $r$ is given by

$$\begin{aligned} q_{krt}(\mathbf{p}_t) = & \left[G(v_{it}^1(\mathbf{p}_t)) - G\left(s_{it}^2(\mathbf{p}_t)\right)\right] P_{irt}^1(\cdot) \\ + & \left[G\left(s_{it}^2(\mathbf{p}_t)\right) - G\left(s_{it}^3(\mathbf{p}_t)\right)\right] \prod_{\{b \in \mathcal{B}_2 \mid kr \in b\}} P_{irt}^2(\cdot) \\ & + G\left(s_{it}^3(\mathbf{p}_t)\right) \prod_{\{b \in \mathcal{B}_3 \mid kr \in b\}} P_{irt}^3(\cdot), \end{aligned} \tag{9}$$

where $\mathbf{p}_t$ is the $K * 3 \times 1$ vector of prices of the products sold by the 3 stores in the market, $P_{irt}^1$ is the probability that a one-stop shopper decides to shop at store $r$, $P_{irt}^2$ is the probability that a two-stop shopper chooses store $r$ as one of the two stores that she will visit, and $P_{irt}^3$ is the probability that a three-stop shopper decides to select a basket $b$ including product $kr$. All of these probabilities are known by shoppers.[24]

The own- and cross-price elasticities of demand are given by the standard formula $\eta_{krht} = \frac{\partial q_{krt}}{\partial p_{jht}} \frac{p_{jht}}{q_{krt}}$ for all $j \in \{1, \dots, K\}$, $h \in \{A, B, C\}$. It is important to note that a price change may affect not only the market shares per type of shopper but also the shopping cost cutoff values given that they depend on utilities. As a consequence, the distribution of shoppers between one-, two-, and three-stop shopping groups changes. In fact, an increase in product $k$'s price at store $r$ reduces the indirect utility of consumer $i$ visiting store $r$. Therefore, she may consider making less stops and purchasing a substitute for this product from a rival store, say $h$, as the gain in utility from visiting an additional store may not be sufficient to offset the extra shopping cost.

### 3.3 Supply

Recall that we assumed that all supermarkets supply the same product line to consumers consisting of $k = 1, \dots, K$ products. Given that *ex ante* homogeneous products are *ex post* differentiated because they are sold by different supermarkets, let $J = K * 3$ denote the total number of products existing in the market ($K$ products sold by each of the 3 supermarkets in the market) and $L_r$ the specific product line supplied by supermarket $r$. The profits of supermarket $r$ are given by

$$\Pi_{rt} = \sum_{k \in L_r} (p_{krt} - mc_{krt}) M_{kt} q_{krt}(\mathbf{p}_t),$$

where $M_{kt}$ is the size of the market for product $k$ in market $t$ and $q_{krt}$ is the market share of product $k$ sold by supermarket $r$ as defined by equation (9).

We suppose that supermarkets compete in prices. Assuming that there exist a Nash-Bertrand

[24] These probabilities are functions of observable characteristics and parameters. However, for the sake of simplicity, we do not specify this dependence at this stage. We defer these details to the empirical section below.

equilibrium in pure strategies, and that the price vector that supports it has strictly positive entries, the price $p_{krt}$ of some product $k$ sold by $r$ must satisfy the following first-order condition (FOC) derived from $r$'s profit maximization problem:

$$q_{krt}(\mathbf{p}_t) + \sum_{j \in L_r} (p_{jr't} - mc_{jr't}) \frac{\partial q_{jr't}(\mathbf{p}_t)}{\partial p_{krt}} = 0, \text{ for all } r' \in \{A, B, C\}.$$

The FOCs yield a system of $J$ equations that implies price-cost margins for each product in supermarket $r$'s product line. In order to write this system in matrix notation, let $\mathbf{S}_t$ be a $J \times J$ matrix containing market shares responses to changes in retail prices, with entry $S(j,k) = -\frac{\partial q_{jr't}}{\partial p_{krt}}$ for $j, k = 1, \ldots, J$. Further, let $\mathbf{\Omega}_r$ be a matrix of dimension $J \times J$ with $j$th entry equal to 1 if products $j, k$ are in $r$'s product line and zero otherwise. The system of equations writes as:

$$\mathbf{q}_t - (\mathbf{\Omega}_r \mathbf{S}_t)(\mathbf{p}_t - \mathbf{mc}_t) = \mathbf{0},$$

where $\mathbf{q}_t, \mathbf{p}_t$ and $\mathbf{mc}_t$ are $J$-dimensional vectors of market shares, prices and marginal costs. Solving for the margins, we obtain:

$$(\mathbf{p}_t - \mathbf{mc}_t) = (\Omega_r \mathbf{S})^{-1} \mathbf{q}_t,$$

# 4 Empirical implementation and results

## 4.1 Shopping period

We define a shopping period as a week in which a household is recorded as making grocery purchases. During a week, a household that concentrates its purchases in a single supermarket chain can either make one visit to a store in the chain, make several visits to the same store, or it can visit several stores in the same chain. Regardless, we define this household as a one-stop shopper as long as it is observed to only deal with a single supermarket chain during the week. Conversely, when we see the household purchasing products from stores in competing chains during the week, we define it as a multistop shopper.

## 4.2 Products and stores

Our demand model allows shoppers to buy several different products in the same week and assumes that shoppers are making a series of multiple-discrete decisions regarding which products to buy as part of a desired basket of products from a set of mutually exclusive and exhaustive alternatives. This choice set includes baskets of either one product or multiple products that can be purchased from either one store or several stores. When a household is observed to have made no purchases in a given shopping period, we define it as having opted for the outside option.

In our data set, households are observed to purchase up to 275 different products from up to 9 separate grocery retailers during a given shopping period. On average, a household purchases baskets containing 24 products from two stores each week. Estimating a demand system for such a large choice set is infeasible. Thus, we focus on a reduced set of three product categories and three stores that represent the most frequently purchased products and stores most frequently visited by French households. In particular, we include yogurt, biscuits, and refrigerated desserts, given that they meet several criteria that make our empirical exercise consistent with our structural model (see Table 4). First, they are staples, because most French households are heavy consumers of products from these categories (they are typically consumed every day by the average French household), and are not stored for very long, so stockpiling is not a first order concern. Second, these categories are not close substitutes, which ensures that we can observe sufficient variation in shopping patterns, as consumers may tend to concentrate their purchases from the same category in a particular store, but might want to diversify purchases from other categories across stores. Third, the expenditure of on only these three categories corresponds to 8.5% of all grocery product expenditures in our sample. Finally, customers tend to consume one serving of a product from these categories at a time, which makes it convenient for a demand model that relies on a unit demand assumption (see Table 4 for details of how we define servings).

Table 4: Characteristics of the selected categories

| | **Yogurt** | **Biscuits** | **Desserts** |
|---|---|---|---|
| Position among 352 products[a] | 2 | 3 | 6 |
| Serving size (in grams)[b] | 125 | 30 | 80 |
| Mean price (euro cents/serving) | 26.27 | 9.77 | 45.42 |
| Mean consumption[c] | 9 | 12 | 8 |
| Days between purchase | 8 | 10 | 10 |

Source: Kantar Worldpanel database 2005. Authors' calculations.
*Notes:* [a] Positions of the selected products in a ranking of the 352 products we observe, based on the number of purchases in 2005.
[b] Servings are defined according to the most frequently purchased serving of each product.
[c] Average number of servings purchased per household-week.

To capture consumers' responses when a product is delisted, we allow for two mutually exclusive alternatives in the yogurt category, one being the leading NB in France in 2005 and the other being a composite "brand" that consists of the remaining varieties (both other NBs and PLs) available in the market. Concerning the other two categories, in each case we treat purchases of all brands as if they were purchases of a single general brand. Therefore, consumers face a set of four products from which they can choose at most three: one of the two yogurt alternatives, biscuits, and desserts.[25]

Regarding stores, we restrict our attention to the three leading supermarket chains in France based on national market share in 2005. These chains are present throughout the country and account for nearly 60% of groceries sales made by store chains in France (excluding hard

[25] The choice of yogurt as the category with two alternatives is arbitrary. Our results are robust to the selection of the product category containing two options.

discounters) in 2005. The remaining stores observed in our data are included in the outside option along with the no purchase of the included goods option (the interpretation of the outside good in this context is discussed below). Thus, we have four *ex ante* homogeneous products that are available at three stores of similar size. This is consistent with our modeling framework of oligopolistic competition with differentiated product lines, where customers can visit multiple stores in the same shopping period to increase variety. In this context, a basket is a collection of product–store items, and given that there are four products, three stores, and baskets that can consist of one, two, or three items, we end up with a choice set of 112 mutually exclusive alternatives.[26]

### 4.3 Empirical specification of the utility

We empirically specify product-level utility as a function of observed and unobserved product and store characteristics, and time fixed effects. We allow consumer heterogeneity to enter the model through the price coefficient, which is a function of observed and unobserved household characteristics. Formally, let the utility of consumer $i$ from purchasing product $k$ from store $r$ at time $t$ be given by

$$\overline{v}_{ikrt} = -\alpha_i p_{krt} + \mathbf{x}_{kr}\boldsymbol{\beta}_1 + \xi_{kr} + \phi_t, \tag{10}$$

where $p_{krt}$ is the price of product $k$ at store $r$, $\mathbf{x}_{kr}$ is a vector of observed product-store characteristics, $\xi_{kr}$ is a vector of unobserved product-store characteristics, $\phi_t$ are time fixed-effects, $\boldsymbol{\beta}_1$ is a vector of parameters common to all households, and $\alpha_i$ is an individual-specific coefficient that captures the valuation of the price.

The mean valuation of the observed product-store characteristics is not separately identified from that of the unobserved characteristics. These are captured jointly by including product-store dummies.

We model the distribution of consumers' tastes for prices as a function of unobserved demographic characteristics as follows:

$$\alpha_i = \alpha + \sigma^\alpha \nu_i, \qquad \nu_i \sim N(0,1), \tag{11}$$

where $\alpha$ captures the mean (across consumers) valuation of the price of product $k$ sold by $r$ at $t$, $\nu_i$ is a random variable that captures unobserved household attributes that influence consumer choices, and $\sigma^\alpha$ is a scaling parameter.

Further, we assume that individual shopping costs are a parametric function of a common shopping cost across all consumers, $\varsigma$, which can be thought of as the minimum cost every consumer bears as a result of the need to engage in shopping, and an idiosyncratic deviation from this mean that consist of observed, $\mathbf{d}_i$, as well as unobserved household characteristics, $\eta_i$, which rationalizes the observed heterogeneity in shopping patterns across individuals. This yields

[26] We consider cases where consumers purchase only one alternative of the two types of yogurts available.

$$s_i = \varsigma + \mathbf{d}_i \boldsymbol{\pi} + \sigma^s \eta_i, \quad \eta_i \sim N(0,1), \tag{12}$$

where $\boldsymbol{\pi}$ is a vector of coefficients measuring the change in costs with household characteristics and $\sigma^s$ is a scaling parameter.

In line with our modeling framework, we empirically define the utility a $n$-stop shopper ($n \in \{1,2,3\}$) derives from purchasing basket $b$ of products as

$$u_{ibt} = v_{ibt}^n - ns_i + \varepsilon_{ibt} \tag{13}$$

where $v_{ibt}^n$ is the overall utility of basket $b$ net of transport costs as defined by equations (1) through (3) above, $s_i$ is the individual shopping cost, and $\varepsilon_{ibt}$ is an idiosyncratic basket-level shock to utility.

Note that equation (13) along with equations (10) and (12) fully specify the utilities of one- and multistop shoppers as a function of price, product characteristics, distance to stores, and individual transaction costs of shopping. Thus, our utility accounts for both the vertical and horizontal dimensions of consumers' valuations of products. The vertical differentiation is captured by product-store characteristics, while the horizontal differentiation is captured by distance, which varies across store formats and zip codes, and shopping costs.

Finally, we normalize the utility net of shopping costs of the basket containing only excluded products (which we denote as basket $b = O$) to zero. Thus, it is modeled as a function of an individual random shock to utility, $u_{iOt} = \varepsilon_{iOt}$.

### 4.4 The outside option

Recall that to keep our problem involving multiproduct and multistore choices tractable, we restricted the choice set to all baskets resulting from the mix of up to three stores and up to three products from the four alternatives available (two brands of yogurt, biscuits, and refrigerated desserts). The remaining purchases of both included and excluded products at excluded stores and excluded products at included stores, as well as unobserved purchases and visits to unobserved sellers, are treated as outside products.

In this context, the interpretation of the outside option differs from that in a standard discrete-choice model of demand for a single product. As pointed out by Gentzkow (2007), in a model that allows consumers to choose multiple products simultaneously, every choice involves a maximization over all excluded alternatives, unlike the case of a standard multinomial model, where only the utility from good 'zero' is implicitly maximized over all excluded products.

To see what this means in our case, take, for instance, a household that registered purchases of biscuits and desserts from two of the included supermarkets (say, A and C) in a given week. If this was the only grocery shopping activity by that household in that week, we would interpret the household as a two-stop shopper that purchased a basket of two products in two separate shopping locations and conclude, according to our structural model, that its overall utility net of shopping costs was larger than that attainable through either of the two alternatives (one- and three-stop shopping). However, it may be that this household purchased yogurt and/or some excluded products at excluded stores, or that it also purchased some excluded products from the included stores. In any case, the interpretation is that this household was better off choosing a basket obtained by shopping at supermarkets A and C, and possibly an outside option, rather than choosing a basket obtained from any combination of products and stores that included supermarket B.

In line with this, we must include a caveat related to the interpretation of the shopping patterns we observe in our final data set. If a consumer is observed to have made purchases at two of the three included stores, we interpret this as meaning that she was able to make two visits in addition to visits she might have made to any excluded stores. In this sense, the number of store visits by a household that possibly also purchased products at excluded stores is interpreted in the context of this paper as the number of additional visits the household made to included stores.

### 4.5 Identification

Equation (8) shows that we can identify critical cutoff points of the distribution of shopping costs if we are able to both observe the optimal shopping patterns of one- and multistop shoppers and identify the parameters of the product-specific utilities involved in the computation of the $n$th cutoff point. For each individual, we need to identify both the utility of her actual choice, say a basket implying two stops, and the utility she would have derived had she chosen any basket involving alternative shopping patterns (either one- or three-stop shopping). To do this, we exploit the panel structure of our data. We observe sufficient cross-sectional and time variation in terms of choices of products and stores that allow us to identify the mean utility parameters. In particular, we are able to separately identify the price coefficient from the mean utility thanks to the observed variation in the price of the same product. Thus, the predicted probabilities vary as a result of this variation in prices, which generates sufficient moments for identification.

Fixed shopping costs are identified from the observed week-to-week variation in the shopping behavior of each household, e.g. a household undertaking one-stop shopping one week can be observed undertaking multistop shopping the following week. Week-to-week variation is necessary but not sufficient for identification; variation in terms of the set of products purchased from each store is also needed to enable the separation of shopping costs from mean product–store utility parameters. Further, to capture the component of shopping costs that varies with time and stores, we control for household characteristics that account for time constraints. As in Dubois and Jódar-Rosell (2010), we assume that the distance between household location and nearby stores is the same for all households in the same zip code.[27] The inclusion of distances to stores is useful for two reasons: they capture part of the horizontal dimension of consumers' preferences for product characteristics, and they allow us to identify the disutility of transport. By adding this information to the model, along with the unit demand assumption, the remaining

[27] Due to data limitations, we do not pinpoint the precise locations of either households or retailers, but merely use zip codes. As a consequence, we are not able to compute exact distances for each household.

variation in shopping costs across consumers can be interpreted as a pure idiosyncratic shopping cost that is constant across stores and time periods, which is consistent with our modeling framework.

A concern in the identification of demand models is the potential endogeneity of prices. There are two sources for this endogeneity commonly discussed in the literature. On the one hand, firms often react to changes in consumers' valuations for (unobserved) product characteristics by changing prices, which is the reason why prices may be correlated with the error term of the model. On the other hand, the observed variation in prices of the same product across markets may be due to market specific demand shocks that are not independent of price. We try to exploit the panel structure of our data to control for unobservables by including brand and supermarket fixed effects. Moreover, to control for possible market-level shocks, we include time fixed effects. However, given that these controls may not fully account for all the exogenous variation in prices, we estimate our demand model by applying a control function approach for price. In a first stage, we regress the price of each brand on Hausman instruments (Hausman (1997)), which we compute as the prices of the same brand in other regions of France, excluding that where the price to be instrumented was observed.

Finally, the identification of aggregate demand requires the computation of the proportions of one-, two-, and three-stop shoppers, which in equation (9) are defined as the differences between the distribution of shopping costs $G(\cdot)$ evaluated at two different cutoff points. Given our setup, we are able to compute these values from the empirical distribution of one-, two-, and three-stop shoppers that we observe in our data.

### 4.6 Estimation

To estimate the parameters of our model, we use the data set described in Section 2. The sample we use consists of local areas where we observe households undertaking one-, two-, and three-stop shopping and purchasing at least one unit of one of the included products. From the data meeting this criteria, we randomly sample 31 households. Given that we allow for random coefficients of price and shopping costs, our choice probabilities do not have a closed-form solution. Thus, we use simulated methods to compute them. The details of the estimation method we use are as follows.

Let $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}_1', \tau, \boldsymbol{\pi}', \sigma^{\alpha}, \sigma^{s}, \varsigma)'$ be a vector containing all parameters to be estimated. A consumer who wishes to buy a basket of products, denoted as $b$, faces a choice set $\mathcal{B}$ of mutually exclusive and exhaustive alternatives consisting of combinations of products and stores. The basket she chooses is such that she obtains the highest possible utility net of shopping costs. This maximizing behavior defines the set of unobserved charcteristics leading to the choice of alternative $b$ as

$$A_{ibt} = \left\{(\varepsilon_{ibt}, \nu_i, \eta_i) | v_{ibt}^{n} - ns_i > v_{ijt}^{m} - ms_i \;\forall\; m \in \{1,2,3\}, j \in \mathcal{B}\right\},$$

where $n$ and $m$ correspond to the number of stores visited to purchase basket $b$ and basket $j$, respectively, $v^n_{ibt}$ corresponds to the utility derived by consumer $i$ from basket $b$ at time $t$ net of shopping costs. We assume that the random shocks to utility, $\varepsilon_{ibt}$, are distributed i.i.d. type I extreme value. Integrating over $\varepsilon_{ibt}$ yields the closed-form choice probability of alternative $b$, at time $t$, as a function of the characteristics of products and retailers:

$$Q_{bt}(\boldsymbol{\delta}, \mathbf{p}, \mathbf{d}_i, \boldsymbol{\eta}, \boldsymbol{\nu}; \boldsymbol{\theta}) = \frac{\exp(v^n_{ibt} - ns_i)}{1 + \sum_{j \in \mathcal{B}} \exp(v^m_{ijt} - ms_i)}. \tag{14}$$

As each consumer makes a sequence of $T$ choices, we index $H$ as the set of all possible values our data takes, i.e. all sequence of baskets at all choice occasions during our period of observation. The probability of observing consumer $i$ making a sequence of choices $h \in H$ is:

$$P_h(\boldsymbol{\delta}, \mathbf{p}, \mathbf{d}_i, \boldsymbol{\eta}, \boldsymbol{\nu}; \boldsymbol{\theta}) = \prod_{t=1}^{T} Q_{\hat{b}t}(\boldsymbol{\delta}, \mathbf{p}, \mathbf{d}_i, \boldsymbol{\eta}, \boldsymbol{\nu}; \boldsymbol{\theta}), \tag{15}$$

where $\hat{b}$ denotes the actual basket chosen at each corresponding choice occasion. Given the matrix of observable characteristics, $\mathbf{X}$ $(= \{\boldsymbol{\delta}, \mathbf{p}, \mathbf{d}_i\})$, and the $T$-dimensional vector of observed choices for each consumer, $\boldsymbol{h}$, a natural way to estimate $\boldsymbol{\theta}$ is by maximizing the log-likelihood function:

$$L(\mathbf{X}, \mathbf{h}; \boldsymbol{\theta}) = \sum_i \ln \int_{\nu, \eta} P_h(\boldsymbol{\delta}, \mathbf{p}, \mathbf{d}_i, \boldsymbol{\eta}, \boldsymbol{\nu}; \boldsymbol{\theta}) dF(\boldsymbol{\eta}, \boldsymbol{\nu}; \boldsymbol{\theta}). \tag{16}$$

However, the integral over unobservables $\nu, \eta$, does not have a closed-form solution. We use the Simulated Maximum Likelihood (SML) method (see Lerman and Manski (1981), and Pakes and Pollard (1989)) to overcome this problem. As the SML method requires the number of simulation draws, $S$, to approach infinity with $\sqrt{S/I} = O(1)$, we use 100 draws in our simulations. The SML estimator is given by:

$$\hat{\boldsymbol{\theta}}_{SML} = \arg\max_{\boldsymbol{\theta}} \left\{ \sum_i \ln \left[ \frac{1}{S} \sum_s P^s_h(\boldsymbol{\delta}, \mathbf{p}, \mathbf{d}_i, \boldsymbol{\eta}, \boldsymbol{\nu}; \boldsymbol{\theta}) \right] \right\}.$$

### 4.7 Empirical estimates

We use simulated maximum likelihood (SML) to estimate our demand model. We report estimates of the utility parameters, distance and shopping costs, according to four specifications in Table 5. Columns three and four are analogous to columns one and two, respectively, but they include a control function for price. In columns two and four we allow observed household characteristics to enter the utility model through interactions with the number of stores visited in a week. All regressions include product, store and time fixed-effects.

Results are as expected: demands are downward sloping with statistically significant estimates for price in all specifications. The estimate for the mean valuation of distance is negative and statistically significant, which means that the mean valuation of a basket of products is lower the farther a store is from a customer's location. The estimate for the number of supermarkets visited in a week is negative and statistically significant, which we interpret as consumers facing, on average, a positive fixed-cost of dealing with multiple stores. Moreover, the estimates for the standard deviations of the coefficients of price and number of visits are significant and larger than one in both cases. This indicates that unobserved household characteristics are important in explaining observed heterogeneity in household choices and shopping patterns. The estimates of the interactions of the number of different stores visited in a week with observed household characteristics are statistically insignificant in most cases.

Table 5: Estimates for the utility parameters and shopping costs[a]

| | Uncorrected | | With control function | |
|---|---|---|---|---|
| Covariate | (1) | (2) | (3) | (4) |
| **Means** | | | | |
| Price (€/basket) | -4.19 | -3.72 | -12.79 | -12.77 |
| | [-2.01 , -6.00] | [-1.91, -5.66] | [-4.55, -20.78] | [-3.96, -20.25] |
| No. of visited stores | -3.84 | -3.83 | -4.05 | -4.16 |
| | [-2.93, -4.67] | [-2.96, -4.56] | [-2.99 , -4.98] | [-3.26, -4.97] |
| Distance (km) | -0.06 | -0.06 | -0.06 | -0.06 |
| | [-0.05, -0.07] | [-0.05, -0.07] | [-0.05, -0.08] | [-0.05, -0.07] |
| No. of visited stores × hh's head age | | -0.01 | | 0.01 |
| | | [-0.04, 0.03] | | [-0.02, 0.04] |
| No. of visited stores× Car | | 0.91 | | -0.07 |
| | | [-0.87, 3.35] | | [-1.33, 1.38] |
| No. of visited stores × log of income | | 0.91 | | 0.84 |
| | | [0.12, 1.61] | | [-0.11, 1.92] |
| No. of visited stores × Household size | | -0.06 | | -0.30 |
| | | [-0.41, 0.36] | | [-0.64, 2.00] |
| **Standard deviations** | | | | |
| Price | 2.74 | 2.84 | 2.58 | 2.79 |
| | [2.34, 3.37] | [2.13, 3.46] | [2.12, 3.12] | [2.07, 3.51] |
| No. of visited stores | 1.26 | 1.42 | 1.41 | 1.47 |
| | [1.09, 1.81] | [0.98, 1.86] | [1.06, 1.83] | [0.95, 1.82] |
| Observations | 252,267 | 252,267 | 252,267 | 252,267 |

*Notes:* [a]The 95% confidence intervals are given in square brakets. All regressions include product, store and time (months) fixed effects.
[b]A basket can consist of unit servings of one, two or three products: dessert (80g), biscuit (30g) and, yogurt (125g).

Our preferred specification is the full model with a control function for price (column (4) of Table 5). We use the results of this specification to obtain measures in euros of the average shopping and transport costs by dividing the estimated coefficient of number of supermarket visited in the first case, and distance, in the second case, by the estimate of the price coefficient. We report the results in Table 6. The average fixed shopping cost a consumer incurs when sourcing a supermarket is 33 euro cents. The transport cost per store, taking the average distance of 7.2 km between consumer's dwelling and a store, is 4 euro cents per trip. Summing up, the total shopping costs (fixed shopping costs plus transport costs) the consumer incurs for visiting a store is, on average, 36 euro cents. This costs are nearly 1% of the mean total expenditure on groceries made by a household on one supermarket in a week.

Similarly, we translate into euros measures of the shopping cost cutoffs computed as differences in utilities between one, two and three-stop shopping according to our model. The

Table 6: Mean fixed shopping cost, mean distance cost and total shopping cost at the supermarket level

| | Euro cents[a] | % of mean weekly expenditure[b] |
|---|---|---|
| Mean shopping cost | 33 | 0.82 |
| Mean transport cost[c] | 4 | 0.09 |
| Mean total shopping costs | 37 | 0.94 |

*Notes:* [a] To translate estimates into euros, we divide the absolute value of each coefficient by the absolute value of the mean price coefficient.
[b] The average weekly expenditure of a household on groceries per supermarket is 39.5 euros. [c] This figure is obtained by multiplying 0.005 euro cents, which is the transport cost per kilometer, by 7.2km which is the average distance to a supermarket.

threshold between zero and one-stop shopping is 4.2 €. Customers with shopping costs beyond that threshold find it more costly to visit one store in a given week than its benefits, which prevent them to grocery shopping every week. This corresponds to a proportion of shoppers of 60.4% as predicted by our model (see bottom panel of Table 7). One-stop shoppers are those whose shopping costs lie between 1.2 € and 4.2 €, and comprise 38.4% of all customers. Two-stop shoppers (1.2%) are those whose shopping costs are small or even negative. These should lie between -1 € and 1.2 €. Finally, we estimate the maximum shopping costs necessary to rationalize the lower proportion of three-stop shoppers (0.008%) as negative: -1 €. Consistent with our model, we interpret a negative shopping cost as a measure of a strong taste for shopping.

Table 7: Average implied shopping cost cutoffs (across periods and consumers –in euros) and distribution of shoppers (in percentages)[a]

| Number of stops | Mean shopping cost cut-offs[b] (in euros) | Distribution of shoppers (%) |
|---|---|---|
| Zero - stop shopping | —— | 60.4 |
| | | [52.5, 66.7] |
| One - stop shopping | 4.179 | 38.4 |
| | [1.087, 6.639] | [32.1, 45.6] |
| Two - stop shopping | 1.232 | 1.2 |
| | [-2.323, 4.244] | [0.6, 3.0] |
| Three-stop shopping | -1.022 | 0.008 |
| | [-4.982, 2.394] | [0.001, 0.05] |

*Notes:* [a] 95% confidence intervals are given in square brackets.
[b] To translate estimates into euros, we divide the absolute value of each coefficient by the absolute value of the price coefficient.

**Elasticities**

We report the averages of the estimated own- and cross-price elasticities in Table C.2.1 of the Appendix. Own-price elasticities are negative and most are larger than $-1$. Cross-price elasticities intra-store are positive for products in the same category (yogurts) and negative otherwise. We interpret this as complementarity between categories in the same store. Further, cross-price elasticities are positive for the same product across stores. When the price of a product rises in a given supermarket, the demand for products in different categories in that supermarket decreases, whereas the demand for all products in competing stores increases. This result is a consequence of two things we observe in our data, namely, the large number of consumers who purchase products from the three categories we consider in the same week, and the large proportion of one-stop shoppers. According to our framework, due that these consumers find it very costly to make an additional stop for purchasing just one product, they prefer substituting stores instead and buy the entire basket at a competing location where its total value is lower.

### 4.8 The role of shopping costs

Do shopping costs matter when it comes to explaining consumer shopping behavior? We answer this question with the help of two exercises. On the one hand, we take our estimated model and simulate a scenario in which shopping costs fall to zero, i.e. we assume that consumers no longer incur positive shopping costs. On the other hand, we estimate an alternative specification that does not include any shopping costs, and compare the results with those obtained from our preferred specification with shopping costs.

In the first exercise, we compare the predicted probabilities of being a zero-, one-, or a multistop shopper obtained from our estimated model (baseline) with a counterfactual scenario in which all consumers face zero shopping costs. The results are reported in Table 8. In the absence of shopping costs, consumers choose the outside option with smaller probability. Shopping costs introduce frictions that deter consumers from purchasing the included products on a regular basis, which is consistent with the theory and shows the importance of accounting for such costs in a model of multistop shopping. Further, a scenario with zero shopping costs predicts a larger proportion of multistop shoppers, nearly 41% as opposed to 1.2% in our estimated model with shopping costs. Note that removing shopping costs need not translate into a situation in which all consumers are multistop shoppers. In fact, most shoppers optimally choose to visit a single store. This might be the result of unobserved idiosyncratic valuations of product-store combinations that yield a greater utility through concentrating purchases at a single store (and that are captured by the error term of the model).

In the second exercise, we compare estimated substitution patterns and implied markups and marginal costs from two alternative specifications of the demand model, namely, one with shopping costs (our base model) and another one without shopping costs. We use the semi-elasticities obtained under each model and compute the ratio between the two as an indicator of estimation bias (assuming that our base model is the correct specification). Table 9 shows the average ratios in two panels: the top panel presents the ratios of within-retailer own- and cross-price semi-elasticities averaged across retailers, while the bottom panel reports the ratios

Table 8: Probability of visiting a given number of stores with and without shopping costs[a]

| Number of stops[b] | Baseline | Zero shopping costs | Precentage change |
|---|---|---|---|
| 0 | 0.604 | 0.015 | -98% |
| | [0.525, 0.667] | [0.0065, 0.032] | |
| 1 | 0.384 | 0.554 | 44% |
| | [0.321, 0.456] | [0.351, 0.701] | |
| 2 | 0.012 | 0.406 | 3,174% |
| | [0.006, 0.0296] | [0.265, 0.561] | |
| 3 | 0.00008 | 0.0245 | 30,525% |
| | [0.00001, 0.0005] | [0.0085, 0.092] | |

*Notes:* [a] 95% confidence intervals are given in square brackets.
[b] If the consumer visited stores other than the three included in our sample, the number of stops in this table should be interpreted as "additional" stops.

of inter-retailer cross-price semi-elasticities averaged across supermarkets.[28]

All of the entries on the main diagonal of Panel A are positive and greater than one, which means that the estimated own-price semi-elasticities under the no-shopping-costs scenario are, on average, biased upwards as compared to those obtained when shopping costs are accounted for. Similarly, all of the entries in the main diagonal of panel B are positive and larger than one. Moreover, cross-price elasticities between "yogurt" and "yogurt NB" are also positive and larger than one for both intra- and inter store cases. Such figures suggest that the absence of shopping costs in a model of multistop shopping tends to overestimate cross-price semi-elasticities of the close substitutes of a product. Our interpretation of this result is that in the absence of shopping costs, consumers are more sensitive to a price change because they are able to substitute products across stores at no additional costs.

Conversely, a model without shopping costs appears to predict lower cross-price semi-elasticities for products in different categories both intra- and inter store. This suggests that estimates of substitution patterns between product categories are biased downwards when shopping costs are not accounted for in a model of multistop shopping. Furthermore, the magnitudes of the cross-category price-semi-elasticity ratios are in most cases lower than 0.2. This means that the magnitudes of this type of semi-elasticities obtained from a model without shopping costs are very small with respect to those obtained from our preferred specification. Now, both models predict the same direction for consumer substitution patterns. When the price of a product, say, biscuits changes, the demand for the remaining products decreases. This suggests that cross-category price semi-elasticities are capturing the economies of scope of concentrating purchases with a single store. In this case, shopping costs make those economies of scope more valuable as it is more costly for consumers to substitute a single product and dealing with multiple supermarkets rather than switching stores.

[28]We report mean estimated own- and cross-price elasticities for each specification in Tables C.2.1 and C.2.2 of the Appendix.

Table 9: Average ratios of own- and cross-price semi-elasticities estimated from two alternative models[a]

| Product | Yogurt | Yogurt NB | Biscuits | Desserts |
|---|---|---|---|---|
| **Panel A: within supermarket** | | | | |
| Yogurt | 1.47 | 2.73 | 0.15 | 0.18 |
| Yogurt NB | 2.45 | 1.63 | 0.15 | 0.18 |
| Biscuits | 0.14 | 0.15 | 1.55 | 0.14 |
| Dessert | 0.17 | 0.19 | 0.14 | 1.48 |
| **Panel B: across supermarkets** | | | | |
| Yogurt | 2.70 | 2.84 | 0.15 | 0.31 |
| Yogurt NB | 2.64 | 2.78 | 0.15 | 0.30 |
| Biscuits | 0.15 | 0.15 | 3.17 | 0.17 |
| Dessert | 0.32 | 0.33 | 0.23 | 3.03 |
| **Outside option** | 9.99 | 10.86 | 11.25 | 10.85 |

*Notes:* [a]Ratios are computed by dividing the estimated semi-elasticity of the model without shopping costs with the corresponding semi-elasticity of the model with shopping costs.
[b]An entry in the panel "Within supermarket" corresponds to the average across supermarkets of intra-store own- and cross-price semi-elasticity ratios. An entry in the sub-panel "Across supermarkets" corresponds to the average across retailers of inter-store cross-price semi-elasticity ratios.

Next, we report average prices of the observed equilibrium as well as average implied markups and marginal costs according to the two specifications in in Table 10. The markups predicted by our base model are lower, on average, as compared to those predicted by a model without shopping costs. This is in part explained by the lower cross-price elasticities, and for the higher market shares that the model without shopping costs predicts.

### 4.9 Robustness

In this section we test our estimates of shopping costs to one potential concern. Our identification strategy for the identification of shopping costs comes from the definition of one- and multistop shopping, which is in line with the theory literature that considers a consumer as a multistop shopper when she is dealing with more than one supermarket in the same shopping period (see Klemperer (1992)). Our empirical interpretation of this definition does not consider the intensity of the relationship of a consumer with each supermarket in the same shopping period as being multistop shopping. For example, if we observed that a consumer used only one supermarket chain in a given week, we would interpret this as that consumer being a one-stop shopper, even if during that week we observed that the consumer made multiple trips to the stores of that supermarket.

A concern with our identification strategy is that additional trips to the same store may imply positive costs to the consumer, even if she is dealing with a single supermarket. If this is so, our estimates of shopping costs may be biased upwards because we are counting less stops per individual, on average. In order to check if our results are sensitive to the definition of shopping

Table 10: Average prices and implied marginal costs and markups according to two alternative models

| | | With shopping costs | | With shopping costs | |
|---|---|---|---|---|---|
| Product | Price* (Euro cents) | Marginal cost (Euro cents) | Markup (%) | Marginal cost (Euro cents) | Markup (%) |
| **Supermarket 1** | | | | | |
| Yogurt | 26.83 | 21.55 | 19.68 | 17.76 | 33.82 |
| Yogurt NB | 25.34 | 20.08 | 20.75 | 17.42 | 31.23 |
| Biscuits | 9.83 | 4.72 | 51.98 | 0.52 | 94.74 |
| Desserts | 49.44 | 44.91 | 9.15 | 39.32 | 20.45 |
| **Supermarket 2** | | | | | |
| Yogurt | 27.34 | 21.37 | 21.84 | 17.66 | 35.40 |
| Yogurt NB | 27.32 | 21.34 | 21.89 | 18.62 | 31.84 |
| Biscuits | 10.14 | 4.18 | 58.74 | 0.34 | 96.67 |
| Desserts | 48.52 | 42.97 | 11.44 | 38.11 | 21.45 |
| **Supermarket 3** | | | | | |
| Yogurt | 31.45 | 23.70 | 24.65 | 20.27 | 35.55 |
| Yogurt NB | 26.02 | 18.45 | 29.12 | 15.42 | 40.73 |
| Biscuits | 11.62 | 5.66 | 51.27 | 1.85 | 84.08 |
| Desserts | 52.48 | 47.05 | 10.35 | 43.19 | 17.70 |

*Notes:* * Correspond to the averages of the prices observed in the data.

patterns, we re-estimate the demand model with a more flexible definition of multistop shopping which will count every trip to a store made by an individual as a stop no matter how many supermarkets is she dealing with. Results are displayed in table 11 (the columns of this table are analogous to those of Table 5). The estimates for price and shopping costs are similar to those obtained in table 5.

We translate estimates of shopping costs obtained from the two alternative definitions of shopping patterns into euros by dividing the mean estimate of the number of stores visited/trips by the estimate of the price coefficient for each specification. We report the results in Table 12. The estimate of the shopping cost per trip made to a store is 24 euro cents. In order to compare this figure with the one obtained in our main results, we need to compute shopping costs at the store level because consumers in our data make, on average, more than one trip to a given store in a week. We do this by multiplying the cost per trip (24 euro cents) by the average number of trips made to each store in a week (which is 1.41 trips, see Table 2), which yields a mean shopping cost of 34 euro cents of dealing with each supermarket. This figure is almost identical to the estimate we obtained with our original definition multi-stop shopping. Our main results are, therefore, unaffected by the definition of shopping patterns.

Table 11: Estimates for the utility parameters and shopping costs using an alternative definition for shopping costs[a]

| Covariate | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Means** | | | | |
| Price (€/basket[b]) | -3.13 | -3.55 | -12.15 | -12.92 |
| | [-1.06, -5.05] | [-0.97, -5.48] | [-4.00, -19.41] | [-4.92, -19.84] |
| No. of trips to stores | -3.29 | -2.99 | -2.77 | -3.09 |
| | [-2.63, -3.99] | [-2.43, -3.64] | [-2.21, -3.30] | [-2.42, -3.82] |
| Distance (km) | -0.06 | -0.06 | -0.06 | -0.06 |
| | [-0.05, -0.08] | [-0.06, -0.08] | [-0.05, -0.08] | [-0.05, -0.08] |
| No. of visited stores × hh's head age | | 0.01 | | 0.02 |
| | | [-0.02, 0.03] | | [-0.05, 0.04] |
| No. of visited stores× Car | | 0.77 | | 0.43 |
| | | [-0.25, 1.86] | | [-0.42, 1.32] |
| No. of visited stores × log of income | | 0.72 | | 1.00 |
| | | [-0.19, 1.94] | | [0.30, 1.89] |
| No. of visited stores × Household size | | 0.20 | | 0.03 |
| | | [-0.10, 0.48] | | [-0.28, 0.33] |
| **Standard deviations** | | | | |
| Price | 2.21 | 2.55 | 3.01 | 2.45 |
| | [1.53, 2.88] | [1.71, 2.97] | [2.00, 3.30] | [1.50, 3.03] |
| No. of trips to stores | 1.43 | 1.37 | 1.44 | 1.48 |
| | [1.00, 1.94] | [0.79, 2.05] | [1.09, 1.85] | [0.94, 2.04] |
| Observations | 252,267 | 252,267 | 252,267 | 252,267 |

*Notes:* [a]The 95% confidence intervals are given in square brakets. All regressions include product, store and time (months) fixed effects.
[b]A basket can consist of unit servings of one, two or three products: dessert (80g), biscuit (30g) and, yogurt (125g).

Table 12: Estimated mean shopping costs according to two alternative definitions of shopping patterns (in euro cents)

| | Definition of shopping patterns | |
|---|---|---|
| | Preferred | Alternative |
| Trip level | —— | 24[a] |
| Supermarket level | 33[a] | 34[b] |

*Notes:* [a] Mean shopping costs for each case are obtained, respectively, using results in column 4 of our main results (Table 5) and column 4 of the robustness results (Table 11).
[b] This figure is obtained as the product of the mean shopping cost at the trip level and the mean number of trips to a store observed in our data (which is 1.41, see Table 2).

# 5 Application to product delisting

We apply our models of demand and supply to study the effects of product delisting by a particular supermarket on consumer shopping behavior. We are interested in how consumers substitute products and stores when one of the products is not available any longer at their preferred shopping location. To capture this demand-side effects in a flexible and realistic way, we take into account that rival supermarkets may react to product delisting by adjusting their prices. Thus, we use our model of supermarket competition in order to capture the strategic reactions of rivals when a supermarket removes a product from its stores' shelves.

There is, however, an important limitation to our approach. In our model of supply, we implicitly assume that wholesale prices are exogenous and do not adjust when a product is delisted by one supermarket. In other words, we are assuming that upstream firms do not respond to product delisting by a downstream counterpart. Of course, this is a very restrictive and unrealistic assumption given that, as we pointed out in the introduction, delisting can be used strategically as a vertical restraint imposed by supermarkets on manufacturers in order to obtain better terms of trade. We, however, make this assumption for the sake of simplicity. A complete study of the effects of product delisting would consider how both rival stores and manufacturers optimally adjust prices in the same period. This is out of the scope of this paper and we leave it for future research.[29]

## 5.1 Counterfactual simulations

Retail stores often use the threat of delisting either a product or a range of products to obtain better deals with manufacturers. Although the supermarket can benefit from this practice, it also entails some losses, as consumers may be tempted to switch to rival stores when a product they desire is unavailable at their usual store. For instance, in 2009 in Belgium, a request for a price rise by Unilever triggered the delisting of 300 of Unilever's products by Delhaize, one of the largest supermarket chains in that country. Both parties ended up being hurt: Delhaize lost 31% of its customers to rivals and among those who remained, 47% substituted other brands for Unilever's products.[30] In the UK, a similar dispute between Tesco, the largest supermarket chain in the country, and *Premier Foods* resulted in the delisting of a number of the suppliers' products in 2011, resulting in a 1% fall in sales and an 18% fall in the value of Premier Foods' shares.[31]

If consumers find it very costly to visit alternative stores (e.g. because of strong feelings of loyalty, very large shopping costs, or head-to-head competition between stores), the delisting of a product will only hurt the manufacturer, making the threat of delisting an effective bargaining strategy for retailers. However, when consumers find it optimal to undertake multistop shopping, the delisting of a product can also hurt the delisting store as a result of a reduction in demand from shoppers who either continue to undertake one-stop shopping at a competing store or visit an additional store if their shopping costs are sufficiently low.

To assess the effects of product delisting on consumer shopping behavior and supermarket

[29] In a sepparate project we are currently working on, we are developing a model of vertical relations with multiple common agencies in which we consider the bargaining process between manufacturers and retailers that will allow us to account for the implications of product delisting on both the vertical and the horizontal dimensions, using our framework on multiproduct demand with shopping costs. With this model, we hope to obtain empirical evidence to find out if supermarkets delist products for genuine commercial reasons or as a strategy to exploit upstream firms.

[30] See http://in.reuters.com/article/delhaize-unilever-idINLG51937220090216.

[31] See http://www.telegraph.co.uk/finance/8684844/Premier-Foods-crumbled-by-Tesco-bust-up.html

competition, we perform a counterfactual experiment in which we simulate a large increase in the price of NB yogurt in supermarket 1 so that it becomes prohibitively expensive.[32] We use our preferred demand estimates and the supply model introduced above to compute prices and market shares of the resulting counterfactual equilibrium.

Table 13 reports the prices and price changes of each store under the baseline scenario and when supermarket 1 delists yogurt NB.[33] We compare two distinct cases: (1) no delisting versus delisting under the estimated shopping costs; (2) no delisting versus delisting assuming that no shopping costs. For each situation, except for baseline (column 1), we recompute equilibrium prices and market shares.

Under the case with shopping costs, supermarket 1 decreases the price of the composite yogurt category by 0.23% after delisting a yogurt NB. The decrease in prices might be associated to the removal of the upward pricing pressure caused by the removal of a substitute product, the yogurt NB. That is, prior to the delisting, supermarket 1 internalized the business stealing effect caused by an increase in prices on the yogurt composite good leading to higher prices on the overall yogurt category. After the delisting, supermarket 1 incentives to raise prices by internalizing the business stealing effect disappear resulting in a decrease of the price for the yogurt composite brand. Conversely, the delisting of the yogurt NB increases prices of biscuits and desserts by 0.57% and 0.10%, respectively. Even though, yogurts, biscuits and desserts are a priory nor substitutes nor complements, consumers shopping costs create complementarities between product categories. The complementarities between product categories create a downward pricing pressure on the categories sold by a supermarket. That is, supermarkets internalize the negative effects on the demand for biscuits and desserts when increasing the price of yogurts. After delisting the yogurt NB, the downward pressure pricing effect no longer exists resulting in an increase for the prices of desserts and biscuits categories. Competitors increase the prices of all categories, suggesting a shift of consumer purchasing behavior their favor.

Under the case without shopping costs, prices are higher. Under the no shopping costs scenario, consumers increase their likelihood of purchasing all product categories in a single week. This implies a larger demand leading to higher prices in equilibrium. After delisting the yogurt NB, supermarket 1 decreases the price of the composite yogurt category by 0.32%. Nevertheless, the delisting of the yogurt NB by supermarket 1 is not followed by an increase in the prices of the biscuits and dessert categories. As shopping costs are no longer creating the complementarity between the product categories, supermarket 1 no longer has the inter-category downward pricing pressure. This finding is in line with the role of shopping costs on creating inter-category complemetarities. The decrease in the price of the yogurt composite brand is not enough to prevent the consumers shopping behavior shift to other retailers. As a result of the overvall increased demand, competitors prices increase for all categories.

[32] We believe that a situation in which a product is so expensive that nobody can afford to buy it is equivalent to a situation in which the product is no longer available at that store.

[33] This simulation is similar to what happened in 2009 between Costco and Coca-Cola. After a dispute regarding prices, Costco decided to delist Coca-Cola products. (http://www.reuters.com/article/cocacola-costco-idUSN1020190520091210).

Table 13: Prices and price changes when supermarket 1 delists yogurt NB*

| Scenario/ Product | Prices with shopping costs | | | Prices without shopping costs | | |
|---|---|---|---|---|---|---|
| | Baseline (euro cents) | Delisting (euro cents) | Change (%) | Baseline (euro cents) | Delisting (euro cents) | Change (%) |
| **Supermarket 1** | | | | | | |
| Yogurt | 26.83 | 26.77 | -0.23 | 31.16 | 31.06 | -0.32 |
| Yogurt NB | 25.34 | —– | —– | 29.63 | —– | —– |
| Biscuits | 9.83 | 9.89 | 0.57 | 14.63 | 14.63 | 0.00 |
| Desserts | 49.44 | 49.48 | 0.10 | 54.34 | 54.34 | 0.00 |
| **Supermarket 2** | | | | | | |
| Yogurt | 27.34 | 27.35 | 0.02 | 30.21 | 30.24 | 0.10 |
| Yogurt NB | 27.32 | 27.33 | 0.02 | 30.15 | 30.18 | 0.10 |
| Biscuits | 10.14 | 10.14 | 0.02 | 13.56 | 13.57 | 0.08 |
| Desserts | 48.52 | 48.53 | 0.01 | 51.72 | 51.73 | 0.01 |
| **Supermarket 3** | | | | | | |
| Yogurt | 31.45 | 31.46 | 0.04 | 34.57 | 34.62 | 0.16 |
| Yogurt NB | 26.02 | 26.03 | 0.04 | 29.11 | 29.16 | 0.18 |
| Biscuits | 11.62 | 11.62 | 0.04 | 15.46 | 15.48 | 0.09 |
| Desserts | 52.48 | 52.49 | 0.02 | 55.54 | 55.55 | 0.02 |

*Notes:* *Supermarket 1 delists yogurt NB while its rivals continue to stock the product.

Similarly, Table 14 reports the market shares and the market shares changes of each store under the baseline scenario and when supermarket 1 delists yogurt NB. Under the case with shopping costs, supermarket 1 decrease in the price of the composite yogurt category is matched with an increase in demand of 1.4%. Nevertheless, we observe a decrease in market shares for Biscuits and Desserts in supermarket 1. Two compounding effects cause the decrease in market shares: (1) the increase in prices due to the removal of the downward pricing pressure generated by the delisting of the Yogurt NB; (2) consumers shopping costs, which make one- or two-stops shoppers decide to switch stores. The combination of both effects make that the decrease in prices in the composite yogurt category not sufficient to compensate one- and two-stop shoppers to keep patronizing supermarket 1. Instead, consumers switch to either supermarket 2 or 3 leading to an increase in market shares for all products offered in these supermarkets. As expected, the increase in market shares is larger for yogurt NB than for the composite yogurt category. The biscuit and desserts categories also experience an increase in market shares across supermarkets 2 and 3.

Under the case without shopping costs, the decrease in the price of yogurt NB by supermarket 1 is followed by an increase in the composite brand yogurt of 2.4%. As consumers no longer have costs to patronize other supermarkets, the increase in market shares for the composite brand yogurt is 1 percentage point larger than the scenario where shopping costs are positive. Further more, supermarket 1 market shares for biscuit and dessert categories decrease by 0.02% and 0.15%, respectively. The difference in the magnitudes of the market share reductions across shopping cost counterfactuals clearly reflects the complementarities generated by the shopping costs across seemingly unrelated categories. In line with our previous results, market shares for all other products across supermarkets increase even more than in the case with shopping costs.

Table 14: Market shares and changes in market shares when supermarket 1 delists yogurt NB*

| Scenario/ | With shopping costs | | | Without shopping costs | | |
|---|---|---|---|---|---|---|
| Product | Baseline | Delisting | % change | Baseline | Delisting | % change |
| **Supermarket 1** | | | | | | |
| Yogurt | 3.81 | 3.86 | 1.38 | 19.95 | 20.43 | 2.43 |
| Yogurt NB | 0.24 | 0 | -100 | 1.32 | 0 | -100 |
| Biscuits | 3.45 | 3.37 | -2.55 | 19.87 | 19.87 | -0.02 |
| Desserts | 3.56 | 3.49 | -2.71 | 19.03 | 19.01 | -0.15 |
| **Supermarket 2** | | | | | | |
| Yogurt | 7.87 | 7.89 | 0.19 | 16.11 | 16.27 | 0.98 |
| Yogurt NB | 0.36 | 0.36 | 0.22 | 0.86 | 0.87 | 1.10 |
| Biscuits | 7.48 | 7.51 | 0.17 | 18.97 | 19.02 | 0.22 |
| Desserts | 7.04 | 7.05 | 0.19 | 15.01 | 15.04 | 0.16 |
| **Supermarket 3** | | | | | | |
| Yogurt | 14.79 | 14.83 | 0.26 | 27.51 | 27.72 | 0.75 |
| Yogurt NB | 0.72 | 0.72 | 0.27 | 1.52 | 1.53 | 0.81 |
| Biscuits | 11.47 | 11.51 | 0.24 | 20.99 | 21.04 | 0.19 |
| Desserts | 9.51 | 9.53 | 0.25 | 13.37 | 13.39 | 0.16 |
| **Outside option** | 59.54 | 59.64 | 0.17 | 2.91 | 2.94 | 1.47 |

*Notes:* *Supermarket 1 delists yogurt NB while its rivals continue to stock the product.

In line with our analysis, Table 15 reports the probabilities of visiting each store under the two scenarios. The upper panel reports the results allowing for positive shopping costs, while the lower panel reports on the results when shopping costs are zero. There are three rows within each panel: the first row corresponds to the baseline scenario, second row corresponds to the delisting scenario, and the third row shows the percentage difference between the predicted probabilities displayed in the preceding rows.

Results are consistent between panels. They suggest that when supermakert 1 delists the yogurt NB, the probability of being visited by consumers decreases, while it increases for rival supermarkets. Nevertheless, the probability of being visited decreases more sharply when shopping costs are higher. These results suggest that the mix of shoppers that visit supermarket 1 change. That is, one- and two-shoppers are more likely to switch stores entirely when shopping costs are larger decreasing the probability of visiting supermarket 1. When consumers are allowed not to incur any shopping costs, they are more likely to visit supermarket 1 given its lower prices for the composite yogurt brand.

Finally, we report implied product markups and weekly revenues for all supermarkets under the base and the counterfactual scenarios. We report the results in Tables 16 and 17. Overall, our results suggest that a supermarket hurts itself when it delists one product while competing

Table 15: Predicted probabilities of visiting stores 1, 2, and 3 under different scenarios when supermarket 1 delists yogurt NB*

| Scenario | Baseline | Delisting | % change |
|---|---|---|---|
| **With shopping costs** | | | |
| Supermarket 1 | 0.0740 | 0.0716 | -3.12 |
| Supermarket 2 | 0.1347 | 0.1351 | 0.24 |
| Supermarket 3 | 0.2196 | 0.2204 | 0.33 |
| **Without shopping costs** | | | |
| Supermarket 1 | 0.4549 | 0.4496 | -1.17 |
| Supermarket 2 | 0.3948 | 0.3965 | 0.43 |
| Supermarket 3 | 0.4813 | 0.4834 | 0.43 |

*Notes:* *Supermarket 1 delists yogurt NB while its rivals continue to stock the product.

rivals keep supplying it. Consistent with our previous results, the delisting supermarket suffers more in a situation in which shopping costs are positive because one-stop shoppers who desire to purchase the missing product would switch stores and purchase the whole basket at a rival store.

Table 16: Margins and margin changes when supermarket 1 delists yogurt NB*

| Scenario/ | Margins with shopping costs | | | Margins without shopping costs | | |
|---|---|---|---|---|---|---|
| Product | Baseline | Delisting | % Change | Baseline | Delisting | % Change |
| **Supermarket 1** | | | | | | |
| Yogurt | 19.68 | 19.49 | -0.95 | 30.83 | 30.61 | -0.73 |
| Yogurt NB | 20.75 | ——- | -100 | 32.24 | ——- | -100 |
| Biscuits | 46.04 | 46.35 | 0.67 | 63.76 | 63.76 | 0.00 |
| Desserts | 10.33 | 10.42 | 0.86 | 18.42 | 18.43 | 0.02 |
| **Supermarket 1** | | | | | | |
| Yogurt | 21.84 | 21.86 | 0.06 | 29.27 | 29.34 | 0.24 |
| Yogurt NB | 21.89 | 21.90 | 0.07 | 29.21 | 29.28 | 0.24 |
| Biscuits | 54.77 | 54.78 | 0.01 | 66.17 | 66.20 | 0.04 |
| Desserts | 12.27 | 12.28 | 0.07 | 17.70 | 17.71 | 0.07 |
| **Supermarket 1** | | | | | | |
| Yogurt | 24.65 | 24.68 | 0.12 | 31.45 | 31.56 | 0.34 |
| Yogurt NB | 29.12 | 29.15 | 0.11 | 36.62 | 36.74 | 0.31 |
| Biscuits | 46.76 | 46.79 | 0.05 | 59.99 | 60.03 | 0.06 |
| Desserts | 11.35 | 11.37 | 0.15 | 16.22 | 16.24 | 0.10 |

*Notes:* *Supermarket 1 delists yogurt NB while its rivals continue to stock the product.

The structure of table 17 mimicks that of Table 14. We assume that the market size for each product is equivalent to the average number of servings consumed by an individual in a week multiplied by the french population in 2005.[34]

[34]According to our data, on average, an individual use to consume 3 servings of yogurt, 4 servings of biscuits and 3 servings of desserts in a week.

In all cases, removing one product from its shelves leads to a decrease in revenues while rivals' revenues increase. The revenue decrease for supermarket 1 under positive shopping costs is 3.4%, while without shopping costs is 1.4%. The 2 percentage point difference comes from the outcome of two opposing effects. An extensive margin effect, where consumers decide to patronize a different supermarket due to high enough shopping costs. An intensive margin effect, where consumers previously purchasing a single product, purchase a second (or third) due to a decrease in the prices of baskets that include composite brand yogurt. The fact that the loss in revenue is attenuated in the without shopping costs scenario implies that the intensive margin effect has a stronger under this scenario than in the scenario where shopping costs are positive.

Table 17: Change in weekly revenues under different scenarios when supermarket 1 delists NB yogurt (millions of euros)

| Scenario | Baseline | Delisting | % change |
|---|---|---|---|
| **With shopping costs** | | | |
| Supermarket 1 | 19.48 | 18.84 | -3.37 |
| Supermarket 2 | 39.09 | 39.17 | 0.20 |
| Supermarket 3 | 67.89 | 68.08 | 0.28 |
| **Without shopping costs** | | | |
| Supermarket 1 | 120.73 | 119.05 | -1.39 |
| Supermarket 2 | 94.06 | 94.53 | 0.50 |
| Supermarket 3 | 125.38 | 126.06 | 0.55 |

*Notes:* *Supermarket 1 delists NB yogurt while its rivals continue to stock the product. These numbers were obtained under the assumption that the market size for each product was equivalent to the average number of servings consumed by an individual in a week multiplied by the french population in 2005. According to our data, on average, an individual use to consume 3 servings of yogurt, 4 servings of biscuits and 3 servings of desserts in a week. 95% confidence intervals are given in square brackets.

# 6 Conclusions

In this article, we develop and estimate a model of multiproduct demand for groceries in which customers with different shopping costs can choose between visiting one store or multiple stores in a given shopping period to empirically examine the role individual shopping costs in consumer shopping behavior and supermarket pricing. In our framework, each consumer faces idiosyncratic costs that increase with the number of supermarkets visited on every shopping occasion. Each consumer weighs up the extra benefits of dealing with an additional supermarket against the additional costs involved. If benefits exceed costs, the individual will visit an additional supermarket. Otherwise, she will make all her purchases at a single place. We obtain that the mean (across consumers) total shopping costs are 37 euro cents per store visited. This cost includes a fixed shopping cost of 33 euro cents per store visited and a transport cost of 4 euro cents per trip to a store located at the average distance from consumer location.

We compare our results with those obtained from a situation in which shopping costs are zero, and confirm that accounting explicitly for the observed heterogeneity in shopping patterns through positive shopping costs is important from an empirical standpoint. For instance, when we unilaterally set shopping costs to zero, the model predicts that all consumers would visit at least one store every week with positive probability (which is not consistent with what we observe in our data); conversely, once shopping costs are accounted for, the predicted probabilities of both one- and multistop shopping are lower, and consumers are less likely to visit a supermarket on a week-to-week basis. Our results suggest also that a model of multiproduct choice that does not account for shopping costs tend to overestimate, on average, the implied own-price elasticities, market shares, and markups. On the other hand, it underestimates cross-price elasticities, which implies that the economies of scope of buying related products at a single store are reinforced by shopping costs as it is now more costly to customers to substitute products across stores (i.e. multistop shopping).

Further, we use our model of demand and a model of supply to empirically examine how consumers and rival supermarkets react when one supermarket removes a particular product from its shelves. We simulate a large price increase in one of the products sold by a given supermarket so that it becomes prohibitively costly for consumers, while the same product continues being supplied by competing supermarkets at observed prices. We measure the net effect of such price increase on demand and supply by allowing supermarkets to adjust prices to a new equilibrium. We do this under two scenarios: one in which consumers face positive shopping costs and an alternative scenario in which shopping costs are zero. We find that the delisting supermarket strategically decrease the price of similar products to encourage intra-store substitution. Conversely, we find that prices of complement products increase and offset the decrease in the substitute product, which suggests that an optimal strategy for supermarkets is to keep the overall value of the basket constant and retain one-stop shoppers. By contrast, in a scenario in which shopping costs are unilaterally set to zero, the price of the close substitute still decreases but the prices of complement product remain unchanged. On average, a supermarket loses more revenue when consumers incur positive shopping costs (about 3% of total revenue) than in a context of zero shopping costs (about 1% of total revenue).

There are several interesting avenues for future empirical research that can incorporate our demand framework. One is related to the implications for vertical bargaining and equilibrium pricing at both upstream and downstream levels when supermarkets use product delisting strategically to increase their bargaining power vis-à-vis manufacturers. A model that accounts for the vertical dimension of the supply of groceries along with our model of multiproduct and multistore choice with shopping costs be a natural step toward understanding an overall picture of the consequences of product delisting, even though it represents a complex and challenging task. A second avenue concerns below-cost pricing strategies by supermarkets (such as loss leaders). 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 to offer 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 were eliminated would help to clarify this issue and motivate changes in RBC laws that currently are quite inflexible.

## References


**Aguiar, Mark, and Erik Hurst.** 2007. "Life-cycle prices and production." *The American Economic Review*, 97(5): 1553–1559.

**Armstrong, Mark, and John Vickers.** 2010. "Competitive non-linear pricing and bundling." *The Review of Economic Studies*, , (77): 30–60.

**Brief, Richard.** 1967. "Measuring Shopping Costs." *Journal of Industrial Economics*, 15(3): 237–241.

**Chen, Zhijun, and Patrick Rey.** 2012. "Loss leading as an exploitative practice." *The American Economic Review*, 102(7): 3462–3482.

**Chen, Zhijun, and Patrick Rey.** 2013. "Competitive Cross-Subsidization." *IDEI Working Papers*, , (808).

**Davies, Gary.** 1994. "The delisting of products by Retail Buyers." *Journal of Marketing Management*, , (10): 473–493.

**De los Santos, Babur, Ali Hortaçsu, and Matthijs Wildenbeest.** 2012. "Testing models of consumer search using data on web browsing and purchasing behavior." *The American Economic Review*, 102(6): 2955–2980.

**Dubé, Jean-Pierre.** 2004. "Multiple discreteness and product differentiation: demand for carbonated soft drinks." *Marketing Science*, 23(1): 66–81.

**Dubois, Pierre, and Helena Perrone.** 2015. "Price Dispersion and informational frictions: Evidence from supermarket purchases."

**Dubois, Pierre, and Sandra Jódar-Rosell.** 2010. "Price and Brand Competition between Differentiated Retailers: a Structural Econometric Model." *IDEI Working Paper Series*, , (64): 42.

**European Commission.** 1999. *Buyer power and its impact on competition in the food retail distribution sector of the European Union.* European Commission.

**European Commission.** 2014. "The economic impact of modern retail on choice and innovation in the EU food sector." http://ec.europa.eu/competition/publications/KD0214955ENN.pdf (accessed July 17, 2015).

**Florez, Jorge, and Daniel Herrera.** 2017. "Multiproduct retailing and buyer power: the effects of product delisting on consumer shopping behavior." *PSE Working Papers*, 2017(07).

**FTC.** 2001. *Report on the Federal Trade Commission Workshop on Slotting Allowances and Other Marketing Practices in the Grocery Industry.* Federal Trade Commission.

**Gentzkow, Matthew.** 2007. "Valuing new goods in a model with complementarity: online newspapers." *The American Economic Review*, 97(3): 713–744.

**Hausman, Jerry.** 1997. "Valuation of new goods under under perfect and imperfect competition." In *The Economics of New Goods.* , ed. R. Gordon and T. Bersnahan, Chapter 5, 207–248. University of Chicago Press.

**Hendel, Igal.** 1999. "Estimating multiple-discrete choice models: an application to computerization returns." *The Review of Economic Studies*, , (66): 423–446.

**Hong, Han, and Matthew Shum.** 2006. "Using price distributions to estimate search costs." *RAND Journal of Economics*, 37(2): 257–275.

**Honka, Elizabeth.** 2014. "Quantifying search and switching costs in the U.S. auto insurance industry." *RAND Journal of Economics*, 45(4): 847–884.

**Hortaçsu, Ali, and Chad Syverson.** 2014. "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*, 119(2): 403–456.

**Kim, Jun, Paulo Albuquerque, and Bart Bronnenberg.** 2010. "Online demand under limited consumer search." *Marketing Science*, 29(6): 1001–1023.

**Klemperer, Paul.** 1992. "Equilibrium product lines: competing head-to-head may be less competitive." *The American Economic Review*, 82(4): 740–755.

**Klemperer, Paul, and Jorge Padilla.** 1997. "Do firms' product lines include too many varieties?" *RAND Journal of Economics*, 28(3): 472–488.

**Koulayiev, Sergei.** 2014. "Search for differentiated products: identification and estimation." *RAND Journal of Economics*, 45(3): 553–575.

**Lerman, Steven, and Charles Manski.** 1981. "On the Use of Simulated Frequencies to Approximate Choice Probabilities." In *Structural Analysis of Discrete Data with Econometric Applications.* , ed. Charles Manski and Daniel McFadden, Chapter 7, 305–319. MIT Press.

**Moraga-Gonzalez, Jose Luis, Zsolt Sandor, and Matthijs Wildenbeest.** 2013. "Semi-Nonparametric estimation of consumer search costs." *Journal of Applied Econometrics*, 28(7): 1205–1223.

**Nevo, Aviv.** 2001. “Measuring market power in the ready-to-eat cereal industry.” *Econometrica*, 69(2): 307–342.

**OECD.** 1998. “Buying power of multiproduct retailers.” *Policy Roundtables*, 326.

**OECD.** 2005. “Resale below cost.” *Policy Roundtables*, 231.

**OECD.** 2008. “Monopsony and buyer power.” *Policy Roundtables*, 328.

**OFT.** 1997. “Competition in Retailing.” *Research Papers*, , (13).

**Pakes, Ariel, and David Pollard.** 1989. “Simulation and the asymptotics of optimization estimators.” *Econometrica*, 57(5): 1027–1057.

**Shy, Oz.** 2002. “A quick-and-easy method for estimating switching costs.” *International Journal of Industrial Organization*, , (20): 71–87.

**Thomassen, Oyind, Howard Smith, Stephan Seiler, and Pasquale Schiraldi.** 2017. “Multi-Category Competition and Market Power: A model of Supermarket Pricing.” *American Economic Review*, 107(8): 2308–2351.

**Viard, Brian.** 2007. “Do switching costs make markets more or less competitive? The case of 800-number portability.” *RAND Journal of Economics*, 38(1): 146–163.

**Wildenbeest, Matthijs.** 2011. “An empirical model of search with vertically differentiated products.” *RAND Journal of Economics*, 42(4): 729–757.

# Appendix

## 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 baskets 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

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

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 she is not maximizing over several subsets of stores in equation (3).

# B Data

Three products are taken into the analysis, fresh desserts, yogurt and biscuits, 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 (e.g. different flavors and fruit contents). However, following Nevo (2001), we claim that an individual normally consumes one unit of either product at a time: yogurt (125 grams per portion), biscuits (30 grams per serving), and one serving of dessert (28 grams per serving), 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, the variable "price" is created in the following way: (1) compute the sum of expenditures per product-retailer pair over local markets, defined at the department at each month; (2) compute the sum of number of servings of each product-pair per market; (3), divide the total expenditure over total quantity per market to obtain a common unit price; (4) If no price information is available at the market level, take average unit prices of the product-retailer pair across other local department within the same period and use it. Due to data limitations, we do not account for neither manufacturers' nor stores' promotional activities of any kind.

# C Demand elasticities

## C.1 Own- and cross-price elasticities from our main specification

Table C.2.1 shows mean own- and cross-price elasticities obtained from our main specification. Each entry $i$, $j$, where $i$ indexes rows and $j$ columns, gives the average elasticity of product category $i$ with respect to a change in price of $j$. For the three store chains we consider in our analysis, average estimated elasticities show similar patterns. As expected, we obtain negative own-price elasticities that in most cases are larger than $-1$.

An effect captured by our elasticities is the complementarity between categories in the same store. When the price of a product increases, the demand of products in other categories in that store decrease, whereas the demand for all products at competing stores rises. We interpret this result as being the consequence of two things: the large number of consumers who are observed to purchase the three categories in the same week, and the larger proportion of one-stop shoppers in our data. Consider, for example, an increase in the price of the yogurt NB in store 1. Given that all baskets containing that product in store 1 are more expensive now, demand of all consumers who desire to purchase that product will decrease. Some consumers will find it optimal to substitute Yogurt NB for its alternative and some others will prefer to source another store in order to purchase the Yogurt NB. In particular, if a one-stop shopper prefers the inter-store substitution of Yogurt NB, she must purchase all products at another supermarket as her shopping costs do not allow her to do multistop shopping. As a consequence, demand for all products (but the alternative yogurt) decreases at store 1 and increases at the rival stores.

## C.2 Own- and cross-price elasticities from an alternative specification without shopping costs

As a support for section 4.8 of this paper, and for the sake of comparison, we provide mean own- and cross-price elasticities estimated from an alternative specification in which shopping costs are assumed to be zero for all consumers. We re-estimate such a model and report the resulting demand elasticities in Table C.2.2.

Table C.2.1: Mean own and cross-price elasticities from a model with shopping costs (averages across periods and consumers)

| | | Supermarket 1 | | | | Supermarket 2 | | | | Supermarket 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Brand | Yogurt | Yogurt NB | Biscuits | Desserts | Yogurt | Yogurt NB | Biscuits | Desserts | Yogurt | Yogurt NB | Biscuits | Desserts |
| Supermarket 1 | Yogurt | **-2.9977** | 0.0095 | -0.4387 | -2.2668 | 0.2968 | 0.0165 | 0.0863 | 0.3737 | 0.8886 | 0.0347 | 0.2088 | 0.8861 |
| | Yogurt NB | 0.1662 | **-2.9899** | -0.4417 | -2.2365 | 0.2964 | 0.0165 | 0.0871 | 0.3666 | 0.8828 | 0.0346 | 0.2087 | 0.8644 |
| | Biscuits | -1.3518 | -0.0732 | **-1.1330** | -2.1000 | 0.2460 | 0.0137 | 0.1082 | 0.3486 | 0.7417 | 0.0293 | 0.2377 | 0.7909 |
| | Desserts | -1.3798 | -0.0731 | -0.4132 | **-5.1418** | 0.2543 | 0.0141 | 0.0822 | 0.5003 | 0.7913 | 0.0294 | 0.2013 | 1.1257 |
| Supermarket 2 | Yogurt | 0.1637 | 0.0094 | 0.0425 | 0.2050 | **-2.4907** | 0.0164 | -0.3025 | -1.2219 | 0.8725 | 0.0344 | 0.2034 | 0.8123 |
| | Yogurt NB | 0.1636 | 0.0094 | 0.0425 | 0.2051 | 0.2950 | **-3.3340** | -0.3025 | -1.2210 | 0.8725 | 0.0344 | 0.2033 | 0.8118 |
| | Biscuits | 0.1313 | 0.0076 | 0.0512 | 0.1814 | -0.8115 | -0.0448 | **-0.9861** | -1.0789 | 0.6919 | 0.0278 | 0.2362 | 0.6875 |
| | Desserts | 0.1461 | 0.0082 | 0.0421 | 0.2805 | -0.8633 | -0.0476 | -0.2840 | **-4.0503** | 0.7505 | 0.0280 | 0.1932 | 1.1107 |
| Supermarket 3 | Yogurt | 0.1657 | 0.0095 | 0.0428 | 0.2132 | 0.2946 | 0.0164 | 0.0842 | 0.3548 | **-2.5875** | 0.0344 | -0.3142 | -1.1198 |
| | Yogurt NB | 0.1628 | 0.0094 | 0.0430 | 0.2012 | 0.2932 | 0.0163 | 0.0861 | 0.3385 | 0.8675 | **-2.9496** | -0.3218 | -1.0729 |
| | Biscuits | 0.1408 | 0.0081 | 0.0507 | 0.1954 | 0.2479 | 0.0138 | 0.1061 | 0.3312 | -1.1146 | -0.0465 | **-1.0575** | -0.9962 |
| | Desserts | 0.1565 | 0.0087 | 0.0428 | 0.2945 | 0.2581 | 0.0144 | 0.0788 | 0.5057 | -1.0728 | -0.0419 | -0.2682 | **-4.2389** |
| | Outside Good | 0.0713 | 0.0043 | 0.0256 | 0.0783 | 0.1387 | 0.0078 | 0.0576 | 0.1534 | 0.3721 | 0.0165 | 0.1131 | 0.2959 |

*Notes:* Each entry $i$, $j$, where $i$ indexes rows and $j$ columns, gives the percentage change of demand for product category $i$ with respect to a percentage change in price of $j$. Own-price elasticities are highlighted in bold characters.

Table C.2.2: Mean own and cross-price elasticities from a model without shopping costs (averages across periods and consumers)

| | | Supermarket 1 | | | | Supermarket 2 | | | | Supermarket 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Brand | Yogurt | Yogurt NB | Biscuits | Desserts | Yogurt | Yogurt NB | Biscuits | Desserts | Yogurt | Yogurt NB | Biscuits | Desserts |
| Supermarket 1 | Yogurt | **-2.6019** | 0.0285 | -0.3169 | -1.4478 | 0.3128 | 0.0141 | 0.0930 | 0.3830 | 0.6441 | 0.0266 | 0.1548 | 0.5484 |
| | Yogurt NB | 0.5255 | **-2.8960** | -0.3195 | -1.4239 | 0.3126 | 0.0141 | 0.0940 | 0.3759 | 0.6402 | 0.0266 | 0.1551 | 0.5328 |
| | Biscuits | -0.9319 | -0.0509 | **-1.0297** | -1.3146 | 0.2523 | 0.0114 | 0.1210 | 0.3550 | 0.5150 | 0.0218 | 0.1850 | 0.4779 |
| | Desserts | -0.9463 | -0.0504 | -0.2920 | **-4.3293** | 0.2579 | 0.0117 | 0.0871 | 0.5293 | 0.5385 | 0.0214 | 0.1434 | 0.7341 |
| Supermarket 2 | Yogurt | 0.5268 | 0.0286 | 0.1448 | 0.6654 | **-2.8753** | 0.0142 | -0.3408 | -1.3914 | 0.6424 | 0.0268 | 0.1537 | 0.5151 |
| | Yogurt NB | 0.5269 | 0.0286 | 0.1448 | 0.6655 | 0.3143 | **-3.1790** | -0.3404 | -1.3930 | 0.6425 | 0.0268 | 0.1537 | 0.5155 |
| | Biscuits | 0.4119 | 0.0225 | 0.1821 | 0.5901 | -0.8939 | -0.0406 | **-1.1404** | -1.2276 | 0.4932 | 0.0212 | 0.1873 | 0.4315 |
| | Desserts | 0.4438 | 0.0236 | 0.1398 | 0.9346 | -0.9542 | -0.0435 | -0.3207 | **-4.7212** | 0.5141 | 0.0206 | 0.1395 | 0.7205 |
| Supermarket 3 | Yogurt | 0.5305 | 0.0286 | 0.1447 | 0.6801 | 0.3142 | 0.0142 | 0.0920 | 0.3668 | **-2.8811** | 0.0267 | -0.3586 | -1.1811 |
| | Yogurt NB | 0.5244 | 0.0285 | 0.1464 | 0.6460 | 0.3136 | 0.0142 | 0.0945 | 0.3510 | 0.6393 | **-3.0128** | -0.3654 | -1.1171 |
| | Biscuits | 0.4403 | 0.0239 | 0.1789 | 0.6239 | 0.2590 | 0.0117 | 0.1203 | 0.3426 | -1.2316 | -0.0525 | **-1.1648** | -1.0398 |
| | Desserts | 0.4770 | 0.0251 | 0.1414 | 0.9739 | 0.2654 | 0.0120 | 0.0847 | 0.5410 | -1.2435 | -0.0492 | -0.3185 | **-4.6154** |
| | Outside Good | 0.2539 | 0.0143 | 0.0971 | 0.3014 | 0.1572 | 0.0071 | 0.0680 | 0.1758 | 0.3014 | 0.0137 | 0.0986 | 0.2063 |

*Notes:* Each entry $i$, $j$, where $i$ indexes rows and $j$ columns, gives the percentage change of demand for product category $i$ with respect to a percentage change in price of $j$. Own-price elasticities are highlighted in bold characters.