Working Papers SMART - LERECO

# Spatial competition in the French supermarket industry

Stéphane TUROLLA

**Working Paper SMART – LERECO** N°12-01

February 2012

UMR INRA-Agrocampus Ouest **SMART** (Structures et Marchés Agricoles, Ressources et Territoires)
UR INRA **LERECO** (Laboratoires d'Etudes et de Recherches en Economie)

*Les Working Papers SMART-LERECO ont pour vocation de diffuser les recherches conduites au sein des unités SMART et LERECO dans une forme préliminaire permettant la discussion et avant publication définitive. Selon les cas, il s'agit de travaux qui ont été acceptés ou ont déjà fait l'objet d'une présentation lors d'une conférence scientifique nationale ou internationale, qui ont été soumis pour publication dans une revue académique à comité de lecture, ou encore qui constituent un chapitre d'ouvrage académique. Bien que non revus par les pairs, chaque working paper a fait l'objet d'une relecture interne par un des scientifiques de SMART ou du LERECO et par l'un des deux éditeurs de la série. Les Working Papers SMART-LERECO n'engagent cependant que leurs auteurs.*

*The SMART-LERECO Working Papers are meant to promote discussion by disseminating the research of the SMART and LERECO members in a preliminary form and before their final publication. They may be papers which have been accepted or already presented in a national or international scientific conference, articles which have been submitted to a peer-reviewed academic journal, or chapters of an academic book. While not peer-reviewed, each of them has been read over by one of the scientists of SMART or LERECO and by one of the two editors of the series. However, the views expressed in the SMART-LERECO Working Papers are solely those of their authors.*

# Spatial competition in the French supermarket industry

Stéphane TUROLLA

*INRA, UMR1302 SMART, F-35000 Rennes, France*
*Agrocampus Ouest, UMR1302 SMART, F-35000 Rennes, France*

**Auteur pour la correspondance / Corresponding author**

**Stéphane TUROLLA**

INRA, UMR SMART
4 allée Adolphe Bobierre, CS 61103
35011 Rennes cedex, France
Email: Stephane.Turolla@rennes.inra.fr
Téléphone / Phone: +33 (0)2 23 48 54 00
Fax: +33 (0)2 23 48 53 80

# Spatial competition in the French supermarket industry


## Abstract

This paper challenges the conventional wisdom on the competitive grocery retail sector in France. To that end, I develop a structural model of spatial competition that accounts for (i) market geography on consumers' preferences, and (ii) differences in their shopping list. The demand estimates are used to recover stores' price-cost margin under alternative pricing strategies. I select the best pricing model by applying non-nested tests and show that retailers noticeably distort their offer in highly concentrated markets. Finally, I perform counterfactual experiments to quantify the expected gain of an additional store on consumer welfare and retail prices.



**Keywords:** spatial competition, structural model, discrete choice model, differentiated products, supermarket industry

**JEL classifications:** C35, L13, L81


# Concurrence spatiale dans le secteur de la grande distribution française


## Résumé

Cet article questionne l'idée habituellement convenue d'un secteur de la grande distribution française concurrentiel. Pour répondre à cette question, cet article développe un modèle structurel de concurrence spatiale qui tient compte (i) de la répartition spatiale des acteurs du marché sur les préférences des consommateurs et (ii) des différences en matière de biens achetés entre consommateurs. L'estimation des coefficients du modèle de demande permet de déterminer les marges brutes des magasins sous différentes hypothèses de stratégie de tarification. Par la suite, je détermine le modèle de tarification le plus pertinent à partir d'une procédure de tests statistiques non-emboités et je montre que les distributeurs distordent sensiblement leur offre commerciale dans les marchés fortement concentrés. Enfin, je procède à des expériences contrefactuelles afin d'évaluer les gains attendus sur le bien-être du consommateur et sur les prix associé à l'ouverture d'un magasin.



**Mots-clefs :** concurrence spatiale, modèle structurel, modèle à choix discret, produits différenciés, grande distribution

**Classifications JEL :** C35, L13, L81

**Spatial Competition in the French Supermarket Industry**

# 1 Introduction

Fifteen years ago, the French grocery retail sector underwent major regulatory changes that affected the nature and the intensity of the competition among retailers. In response to the increasing power of large retailers, the French Parliament passed the Galland Act and the Raffarin Act in the summer of 1996 to restore the balance of power with manufacturers and small independent stores.[1] These regulations constitute the climax of government intervention in this sector in the name of protecting consumer interests and the different forms of retail activity. Fifteen years later, the accumulation of reforms indicates the limited effectiveness of these laws.[2] Basically, it is alleged that the Galland Act shifted the bargaining process from "upfront margins" to "hidden margins" at the expense of the retail price, while the Raffarin Act is accused of having secured the rent of incumbents by establishing important barriers to entry. Luc Chatel, the Secretary of State for Industry and Consumer Affairs, claimed: "*The Galland Act [...] as the Raffarin Act [...] pursue laudable goals. Twelve years later, it is clear that they have not achieved these goals*" (see *Le Figaro*, May 26, 2008). Besides, many expert reports and academic studies have emphasized that the enforcement of these acts has generated important distorting effects that have dampened competition (see, e.g., Commission Hagelsteen [2008] or Allain, Chambolle, and Vergé [2008] for a review of the complaints against these regulations). In particular, these laws have been charged with having an inflationary effect on retail prices. Looking at the Eurostat figures, it is striking to note that food product prices increased 5.9% faster than the consumer price index from 1996 to 2009 but remained stable in the Eurozone and even decreased for some countries during this period (e.g., Germany, Netherlands).

The price-raising effect of resale-below-cost laws, such as the Galland Act, has generated a large body of work.[3] In contrast, the effects induced by the Raffarin Act and, more generally, the impact of market concentration on retailers' local monopoly power remain largely unexplored. This issue is particularly topical, as the last reform of the Galland Act in 2008 did not produce

[1] On the one hand, the Galland Act was principally dedicated to preventing retailers from engaging in below-cost pricing by defining *clearly* the below-cost price threshold. Technically, this resulted by excluding the conditional and deferred rebates (i.e. the so-called "hidden margins") from the invoice such that they could not be deducted from the final price. On the other hand, the Raffarin Act was enacted to reinforce the control at the market entrance. To protect small independent stores from the growing pressure of mass distribution, especially the entry of German mass discounters (i.e., Aldi and Lidl), the legislator toughened the planning system by extending the administrative authorization (i.e., a prerequisite for granting a building permit) to stores with selling areas over 300 $m^2$ (1,500 $m^2$ under the previous regulation).

[2] The definition of the below-cost price threshold and the scope of the Galland Act were successively amended by the Dutreil Acts I & II (2003, 2005), the Chatel Act (2008) and the Law of Modernization of the Economy (2008), which has also modified the planning system.

[3] For instance, Biscourp, Boutin, and Vergé [2008] and Bonnet and Dubois [2010] have empirically confirmed the inflationary effect of the Galland Act, while the mechanism through which this regulation has relaxed intrabrand competition has been widely documented in the theoretical IO literature (see, e.g., Allain and Chambolle [2011]).

the expected result. Instead of restoring a fierce level of competition, the elimination of "hidden margins" only slightly decreased retail prices (approximately 1%). This result gives support to the idea that consumers suffer noticeably from the low level of competition in the downstream market.

However, professional and academic circles have only recently begun to question the level of competition in the downstream market. For a long time, the concentrated structure of the French grocery retail sector did not raise concerns over retail prices.[4] For example, in its first notice on the sector, the French Competition Authority (henceforth, CA) argued that: "*The concentration of the retail food industry has little effect on the downstream market because competition is fierce among retail chains*" (Competition Authority [1997, p.28]). During the last few years, the accumulation of evidence reflecting the retailers' abilities to raise their prices more intensively in highly-concentrated markets softened this point of view.[5] Far from having fostered a fierce and fair competition, the control exerted by the regional zoning boards at the market entrance appears to have substantially strengthened the retailers' positions locally (Bertrand and Kramarz [2002];Competition Authority [2007]). Hence, the presumption is strong that retailers benefit from weak local competitive conditions and exert significant market power that distorts price competition in local markets.

Although essential in the motivations and the designs of government regulations in this sector, few studies have evaluated the market power of retailers in France or in OECD countries.[6] Dubois and Jódar-Rosell [2010] provide estimates of the margins of retail chains with a household panel survey for France and reveal a moderate level of profitability (around 8-12%). Interestingly, they demonstrate that retailers use private labels to relax price competition, confirming a number of theoretical predictions. Their empirical model builds on a structural approach that borrows from Smith [2004], who was interested in isolating and quantifying the effect of owning multiple stores in the exercise of retailers' market power for the UK market. However, both studies make some simplifying assumptions that may lead to understate the retailers' local monopoly power as well as their abilities to distort price competition. First, the authors assume that retail chains set uniform prices within large regional markets and ignore the practice of "price flexing".[7] This assumption leads the authors to specify the price decision at the retail chain level, which implies identical profitability levels among the stores of a retail chain within

[4] The aggregate market share of the five largest retailers increased from 61.8% to 75.6% between 1995 and 2009, placing France second among European countries.

[5] A survey conducted by Nielsen shows that food retail prices may vary, on average, up to 9% among the 126 largest French cities (see *Libre-Service Actualités (LSA)*, March 6, 2008). In the same vein, the consumer association *UFC-Que Choisir* reveals that prices may vary up to 20% between two hypermarkets of the same retailer depending on the competition that they face (see *UFC-Que Choisir*, December 26, 2007).

[6] In contrast, several papers have conducted appraisals of retailers' market power for a single product in the context of vertically related markets (see, e.g., Villas-Boas [2007] or Cohen and Cotterill [2011]).

[7] The UK Competition Commission called "price flexing" a pricing strategy that consists of charging different prices for the same product in different stores. As shown in its inquiry of 2000, this strategy is rarely used in England, unlike France. When this strategy is practiced, it only concerns a small number of products, and the price variations are low.

large market areas. However, as cited above, the empirical evidence reported by consumers associations shows the existence of spatial price competition in France at a very detailed level. In addition, the coexistence of several ownership structures (integrated, franchised and independent stores) as well as the potential for cooperative behaviors provide conditions that favor the emergence of a price-setting scheme distinct from that set at the retail chain level. Second, the authors simplify the multi-product nature of the retailers' offer by assuming that consumers buy a *representative* shopping basket. However, previous empirical research has found that the price responsiveness of consumers may differ significantly depending on the size of the shopping basket (Bell and Lattin [1998]), which therefore also affects the appraisal of a store's market power.

To adequately measure the local monopoly power enjoyed by large grocery stores, I develop and estimate a structural model of demand and supply. I model consumer store choice by using a two-step process: (i) a consumer decides whether to purchase a particular type of food product in a large grocery store and (ii) conditional on the shopping basket, the consumer determines which store to visit. The probabilities of purchase incidence derived from the first stage are computed by means of a multivariate probit model. Afterwards, these probabilities are plugged into a flexible discrete choice model of demand among spatially differentiated grocery stores that accounts for consumers' preferences over store characteristics and geographic proximity. To this end, I follow recent contributions that extend the methodology proposed by Berry [1994] and Berry, Levinsohn, and Pakes [1995] to account for geographical dimensions when analyzing retail markets (Thomadsen [2005]; Davis [2006]; McManus [2007]; Chiou [2009]; Manuszak [2010]).[8] Using the estimated demand parameters, I recover the stores' price-cost margin under different pricing scenarios and determine the preferred one by applying a non-nested testing procedure.

I conduct my analysis thanks to a unique database that corresponds to a cross-section survey of 1,654 households living in a metropolitan area in southern France during the year 2000. The survey conveys detailed information on the stores visited at the product category level, which allows me to characterize both the shopping basket of households and their primary shopping destination. I supplement this database with information on store characteristics obtained from a national survey of French outlets.

The results suggest that the substantial market power enjoyed by retailers in some local areas does not arise from anti-competitive practices but mainly results from a high degree of concentration. Moreover, I show that a significant level of heterogeneity exists among stores' price-cost margin both for stores under competing retail chains and, more surprisingly, for stores under the same fascia. Pushing the analysis further, I disentangle the sources of stores' market power and reveal that the stores located in a weakly competitive environment are more likely to dis-

[8] Pinkse, Slade, and Brett [2002] propose a different modeling approach that addresses the magnitude of spatial price competition through the estimates of price reaction functions. This original approach differs from the traditional framework of a random utility model used here.

tort their offer to benefit from greater local monopoly power. Finally, the model estimates are used to perform counterfactual experiments and quantify the effects of an additional store on retail prices and consumer welfare. In doing so, I provide valuable insights into the competitive landscape of the grocery retail sector.

The remainder of the paper is organized as follows. First, I briefly depict the market structure of the French supermarket industry and compute some concentration indicators to illustrate the observed local market concentration (Section 2). Section 3 provides an overview of the data used for the estimation. Section 4 describes the demand model and the pricing equations that allow me to back out the stores' margin. In Section 5, I present the estimation procedure and discuss the assumptions required to identify the estimates of the demand parameters. Section 6 presents the estimates of the demand parameters, displays the stores' margin according to the preferred pricing model and reports the results of the performed robustness tests. I discuss the impact of some counterfactual policy simulations on retail prices and consumer welfare in Section 7. Section 8 concludes.

# 2 The French supermarket industry

In 2009, the French grocery retail sector had sales revenues of €181 billion and represented approximately 580,000 jobs. Since the 1960s, the boom in large grocery stores has completely reshaped the retail landscape and consumer habits in many OECD countries. Over the years, the grocery retail sector has become the preferred distribution channel of the French and accounts today for 70% of food sales and 20% of non-food items, although representing barely 3% of commercial equipments (i.e., 21,000 retail outlets). According to the usual categorization, the grocery retail sector in France operates essentially through four store formats: hypermarkets, supermarkets, convenience stores and hard discount stores.[9] For historical reasons, hypermarkets are very active in France and comprise one-third of the total food sales (see Carluer-Lossouarn [2007] for a historical analysis of the French grocery retail sector).

The largest retail groups operate under two organizational forms that differ principally by the integration of their wholesale activities and their capital structures. *Integrated* groups, such as Carrefour, Auchan or Casino, operate either directly (i.e., through a company-owned store) or through affiliated entities (i.e., by franchising). In contrast, a network of *cooperative* groups,

[9] Professionals and the national institute of statistics (INSEE) identify four store formats according to selling area and product-mix offered. *Hypermarkets* are large grocery stores with a selling area over 2,500 $m^2$ whose sales arise from one-third of foodstuffs (i.e., this format is a combination of a supermarket and a department store). They are generally located out of town and operate as an anchor store for other retail facilities that are located closer to consumers. *Supermarkets* have selling areas that range from 400 to 2,500 $m^2$. They differentiate themselves from hypermarkets by offering a lower assortment of products and by being located in city centers or in the suburbs of large cities. *Convenience stores* are proximity outlets that almost exclusively sell grocery items. These stores usually operate with a selling area below 400 $m^2$ and are mostly integrated into the largest retail groups. Lastly, *hard discount stores* correspond to small supermarkets that carry a limited assortment of low- and medium-range foodstuffs. These stores operate under a dedicated fascia and are integrated most of the time into national retail groups.

Table 1: Market structure for the 500 largest French cities

| Population distribution of cities | Retail chain | | | Retail group | | |
|---|---|---|---|---|---|---|
| | Nb | Mkt Sh. 1 (in %) | HHI>2000 (in %) | Nb | Mkt Sh. 1 (in %) | HHI>2000 (in %) |
| $[0, Q1]$ | 13.97 | 29.11 | 32.80 | 8.42 | 35.54 | 65.60 |
| $[Q1, Q2]$ | 15.63 | 27.17 | 23.20 | 8.80 | 34.75 | 53.60 |
| $[Q2, Q3]$ | 15.54 | 26.92 | 19.20 | 8.89 | 36.19 | 66.40 |
| $[Q3, Q4]$ | 16.65 | 26.21 | 15.20 | 9.22 | 35.82 | 62.40 |
| | Nb | Mkt Sh. 1 | $\overline{HHI}$ | Nb | Mkt Sh. 1 | $\overline{HHI}$ |
| Total | 15.45 | 27.35 | 1590.42 | 8.83 | 35.57 | 2310.94 |

Notes: Descriptive statistics are reported for the first quarter of 2000 and based on the 1999 census population. The database surveyed all of the hypermarkets, supermarkets, hard discount stores and convenience stores (with selling areas over 400 $m^2$). In total, I count 30 retail chains that belong to 12 retail groups. The average of the statistics is reported. For the first quarter of 2000, the computation of the HHI based on national market shares is equal to 747 (retail chain level) and 1214 (retail group level). Source: Author's calculations based on *Panorama TradeDimensions* database.

such as Leclerc, Intermarché and Système U, is composed of independent shopkeepers who source to their central purchasing unit, which is managed by the cooperative's member shareholders. Even if most of the decisions concerning the retail offer are made by the head of the group (e.g., advertising campaigns, terms and conditions of negotiations with suppliers, products listings, and private label assortments), retail prices are set locally. Hence, depending on the local market conditions, each store is free to set its prices above the retail prices implicitly imposed by manufacturers.[10]

One of the most striking features of the success of large grocery stores in European countries is the low number of players who share this success. The French retail market does not depart from the rule, as it is dominated by six retail groups that, taken together, held 84% of the market share in 2009. The six retail groups include Carrefour (24%), Leclerc (17%), Intermarché (13%), Auchan (11%), Casino (10%) and Système U (9%). If this concentration level does not raise anti-competitive concerns at the national level, the situation is mixed at the local level. This difference is explained by the importance of travel costs, which limit the size of market areas and thereby induce only a subset of retailers to operate in a local market. According to previous surveys on shopping patterns, consumers travel, on average, from 10 to 20 minutes (drive-time), depending on the store format, to reach a large grocery store. Based on these figures, the European Commission (EC) and the UK Competition Commission usually delimit the boundaries of a store's catchment area to a distance corresponding to a 20-minute driving time (see, e.g., decisions in cases No. IV\M.1085 Promodès/Catteau or No. COMP/M.1221 Rewe/Meinl), whereas the CA retains two relevant markets (one accounts for all of the stores located within a 15-minutes driving time, whereas the other focuses exclusively on hypermarkets located within a 30-minutes driving time).

[10] In practice, this is true for "independent stores" (i.e., franchised stores and stores belonging to a cooperative group) but is much less common for corporate stores.

In line with the former definition, I compute some statistics that illustrate to what extent the level of concentration at the local level differs from the level of concentration at the national level. To this end, I use the *Panorama TradeDimensions* database, which tracks detailed information on the market structure of the French grocery retail sector (e.g., entry, exit and rebranding). Because I do not have information other than the ZIP codes of the stores, I assume that the stores are positioned at the centers of their respective cities. I choose a radius of 10 km to delimit the boundaries of a store's catchment area. In addition, I limit the scope of my analysis to the 500 largest French cities in order to concentrate on the most populated markets. Therefore, a relevant market consists of one of these cities and its neighboring cities located within a 10 km radius. For the first quarter of 2000, Table 1 reports the number of retail chains and retail groups per market, the market share of the leader and a measure of concentration by computing the Herfindahl-Hirschman Index (HHI) based on selling areas (again considering retail chain and retail group). In the top panel, I organize the results by the quartile classes of the population distribution of the cities. Together, the figures demonstrate that the level of market concentration is higher at the local scale than at the national scale because, in part, of incomplete geographical coverage of the retail groups (8.83 retail groups per market are identified, on average, from a total of 12 groups). More troublesome, however, is that a significant number of trading areas are highly concentrated according to a standard interpretation.[11] For instance, the computations at the retail chain level stress that 32.80% of the cities in the first quartile have an HHI above 2000. As demonstrated by empirical studies on the food retail sector, this local concentration is not costless for consumers as a clear positive relationship was emphasized between market concentration and retail prices (see, e.g., Barros, Brito, and de Lucena [2006]; Biscourp, Boutin, and Vergé [2008]).

Finally, it is worth noting that the concerns raised by the level of concentration observed locally cannot be dissipated by pro-competitive industry dynamics. Since the passing of the Raffarin Act, the market structure of local trading areas has hardly evolved because of important barriers to entry. In contrast to other retail sectors, few stores have opened. More store openings could have strengthened the competitive pressure on incumbents located in highly concentrated markets.[12]

[11] Formally, the HHI is defined as the sum of squares of all of the market shares in the market. According to the 2004 EU Merger Guidelines, a HHI over 2000 indicates a highly-concentrated market.

[12] On average, over the last fifteen years, the hypermarket and supermarket formats have entry rates of approximately 0.3% and 0.7% per year, respectively.

# 3 Data

## 3.1 Presentation and descriptive statistics

I use an original database that surveys the store choices of households living in a metropolitan area in southern France for several food and non-food product categories. The area of study is the French administrative *aire urbaine* of Montpellier (henceforth, Montpellier AU), which covers a total number of 459,916 people. Its market structure closely reflects the local-level observations reported in Table 1. In other words, the area covers an urban center where the retail group leader has roughly 30% of the market share and peripheral areas where half of the large grocery stores compete, at most, with 4 rivals within 5 km. Compared with the figures presented in Table 1, the concentration level of the city of Montpellier is somewhat below the average of the concentration levels of the largest French cities (HHI=1934.50 vs. $\overline{HHI}$=2310.94). However, the city of Montpellier still has a concentrated market.

The survey was conducted jointly by Montpellier's chamber of commerce and the department of economics at the University Montpellier I during the year 2000. The survey follows the quota sampling methodology to create a sample representative of the geographical and socio-economic group (including age) composition of the population. The data were collected at the household level, and the store choices of the households were recorded on a yearly basis. A total of 1,654 households were asked to list the stores visited according to 49 product categories. For a given product category, each household reported the stores visited during the year and ordered them according to their purchase frequency (in steps of 25%). Therefore the survey allows me to determine the primary shopping destination of the households and their top-up stores for each product category, regardless of the distribution channel visited (e.g., large grocery store, specialized store and market place). In addition, the survey gathers information about some household characteristics, such as the age and the socio-economic status of the household head, the number of persons per household and their location of residence. In the following, I focus on food products and restrict the analysis to the 8 most frequently purchased categories (from among the twelve recorded) because of the computational burden of the model.

I supplement this database with information on store characteristics, which I obtained from the *Atlas de la distribution*, a national survey of French outlets. Missing data were filled in by *in situ* surveys. I obtain information about store characteristics, such as fascia, location, store size and the number of employees. To determine the distances traveled by consumers to visit the stores, I geocoded the addresses of the stores, the cities that belong to Montpellier AU and the *IRIS* of Montpellier city (French geographical unit similar to a block-group) in a geographical information system. Assuming that the households reside at the centroid of their geographical unit, I am able to compute the Euclidean distances traveled by the households to visit each store belonging to their respective choice set.[13]

[13]By specifying a *single*-address model, I assume that a households' residence corresponds to the unique departure point of the shopping trip, which greatly simplifies the model specification. The need to specify a *multi*-address

Table 2: Summary statistics

**Panel A: Store data**

| Variable | Units | Mean | S.D. | Min | Max | Obs. |
|---|---|---|---|---|---|---|
| Hypermarket | Binary | 0.19 | 0.40 | 0 | 1 | 62 |
| Supermarket | Binary | 0.40 | 0.50 | 0 | 1 | 62 |
| Hard discount | Binary | 0.32 | 0.47 | 0 | 1 | 62 |
| Convenience store | Binary | 0.08 | 0.27 | 0 | 1 | 62 |
| Gas station | Binary | 0.50 | 0.50 | 0 | 1 | 62 |
| Mall | Nb of retail facilities | 8.52 | 20.42 | 0 | 120 | 62 |
| Surface | $m^2$ | 2167.14 | 2503.56 | 450.00 | 11799.94 | 62 |
| # cash registers | Nb of cash reg./100$m^2$ | 0.66 | 0.20 | 0.33 | 1.25 | 62 |
| # employees | Nb of employees | 81.71 | 119.72 | 4 | 550 | 62 |
| # parking slots | Nb of parking slots/$m^2$ | 0.14 | 0.09 | 0 | 0.55 | 62 |

**Panel B: Household data**

| Variable | Units | Mean | S.D. | Min | Max | Obs. |
|---|---|---|---|---|---|---|
| Age group 1 | Binary | 0.1192 | 0.3241 | 0 | 1 | 1611 |
| Age group 2 | Binary | 0.1136 | 0.3174 | 0 | 1 | 1611 |
| Age group 3 | Binary | 0.1930 | 0.3948 | 0 | 1 | 1611 |
| Age group 4 | Binary | 0.2272 | 0.4191 | 0 | 1 | 1611 |
| Age group 5 | Binary | 0.1540 | 0.3610 | 0 | 1 | 1611 |
| Age group 6 | Binary | 0.1930 | 0.3948 | 0 | 1 | 1611 |
| Credit card holder | Binary | 0.8231 | 0.3817 | 0 | 1 | 1611 |
| Living in a house | Binary | 0.5587 | 0.4967 | 0 | 1 | 1611 |
| Montpellier city | Binary | 0.5444 | 0.4982 | 0 | 1 | 1611 |
| Rural town | Binary | 0.2806 | 0.4494 | 0 | 1 | 1611 |
| Single household | Binary | 0.2259 | 0.4183 | 0 | 1 | 1611 |
| Work | Binary | 0.5307 | 0.4992 | 0 | 1 | 1611 |
| # cars | Nb of cars/household | 1.3563 | 0.8472 | 0 | 5 | 1611 |

Notes: S.D. corresponds to standard deviation. There are 6 age groups (20 to 24 years, 25 to 29 years, 30 to 39 years, 40 to 49 years, 50 to 59 years, 60 years and over). Source: Author's calculations.

In what follows, I am interested in predicting the large grocery store visited most frequently by a household.[14] Thus, I remove from the data those stores that were attended only for top-up shopping. As a result, the choice set of a household may include one of the 62 large grocery stores (hypermarkets, supermarkets, hard discount stores and large convenience stores with selling areas over 400 $m^2$) in the area and one outside option that gathers alternative distribution channels (e.g., specialized stores or market places). Additionally, I follow the methodology applied by the EC in previous investigations and restrict each household's choice set to the stores located within a 20 km radius of its residence.[15] Note that one limitation of the definition of a household's choice set is that the purchases realized in outlets far from the household's residence are excluded *de facto*. This situation may arise for consumers who live in a small peripheral rural town but who work in the metropolitan city. I find 43 households that meet

model becomes more obvious if one studies those sectors in which purchases are motivated by impulsive behaviors or immediate needs (see, for instance, the study of Houde [2011] for the gasoline market).

[14]Details on how to determine the primary shopping destination of households are provided in the Online Appendix available at author's webpage (`http://www.rennes.inra.fr/smart_eng/media/pages_individuelles/stephane_turolla`).

[15]Depending on the average driving speed, a 20 km radius corresponds to approximately 20 to 30 minutes.

these criteria. Hopefully, this situation concerns only a small part of the sample (i.e., around 2.60%) that has to be excluded from the study. After eliminating these households, the database used to conduct the analysis corresponds to a cross-section survey of 1611 households.
I present some summary statistics of the store and household characteristics in Table 2.

## 3.2 The price index

Unlike homescan data, this survey does not record information on the characteristics of the products purchased by households such as price or packaging. However, to account for the pecuniary incentives that impact the store choice decision, I construct a price measure for each product category in all of the stores.
To that purpose, I run a price survey on a subset of stores for a sample of items. I collect the prices of 91 national brand products and first-price products from a total of 27 selected stores such that the sample is representative of the formats, retail chains and locations of the stores operating in the area of study.[16] The prices were reported over three days to avoid seasonal variations, especially for the fruits and vegetables category. Collecting the prices of the national brand products allows me to record the prices of *strictly* homogeneous products offered across stores (except hard discount stores). Therefore, I limit any aggregation bias that might arise when computing a price index from products with different characteristics (e.g., brand, packaging or quality). Additionally, I track the prices of first-price products to account for items available in all store formats. Then, for a store $j$ and a category $c$ composed of $K$ items, $(k = 1, \ldots, K)$, the price index is computed according to the following expression:

$$\widehat{p}_{jc} = \frac{\sum_{k=1}^{K} p_{kjc}}{K} \tag{1}$$

where the store with the largest sales area in Montpellier AU is chosen as base 100 of the price indices (i.e., *Carrefour 2*).[17]
Thereafter, I estimate the price indices for the non-surveyed stores by requiring a hedonic price regression. As I suspect, the elements that impact the price decision (e.g., quality or logistics costs) may be correlated across product categories. Thus, I need to account for the potential price correlation across categories. Therefore, I specify a seemingly unrelated regression (SUR) equations model by assuming that unobserved heterogeneity is distributed according to a multivariate normal distribution. The log of the price of the selected basket of items is regressed on a set of retail chain fixed-effects and variables describing both the competitive and demand environments. Table 3 reports the estimates.
Finally, it should be noted that the computation of the price indices for hard discount stores follows a separate procedure. The price reporting reveals that the hard discounters do not adopt

[16] A list of the selected products is available upon request.
[17] Because of a confidentiality agreement, I do not provide the precise names of the stores.

Table 3: Hedonic regression of Log-price

| Dependent variable: (log) price of the product category | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | SUR Model | | | | | | | |
| Variable | Fruits & vegetables | Meat | Cooked meat | Cheese | Other dairy product | Grocery item | Alcoholic drink | Soft drink |
| Constant | 4.9645*** | 4.7724*** | 4.9513*** | 4.9023*** | 4.8043*** | 4.8987*** | 4.7633*** | 4.9432*** |
| | (0.0672) | (0.0430) | (0.0414) | (0.0361) | (0.0332) | (0.0350) | (0.0310) | (0.0458) |
| Hypermarket | -0.2691*** | -0.3146*** | -0.3427*** | -0.3261*** | -0.2309*** | -0.2831*** | -0.2017*** | -0.4503*** |
| | (0.0803) | (0.0419) | (0.0391) | (0.0284) | (0.0214) | (0.0259) | (0.0145) | (0.0469) |
| Supermarket | -0.1789** | -0.2790*** | -0.3794*** | -0.2841*** | -0.2092*** | -0.2644*** | -0.1818*** | -0.3387*** |
| | (0.0756) | (0.0394) | (0.0368) | (0.0266) | (0.0200) | (0.0243) | (0.0134) | (0.0441) |
| log(Household income) | 0.0241** | 0.0241** | 0.0241** | 0.0241** | 0.0241** | 0.0241** | 0.0241** | 0.0241** |
| | (0.0099) | (0.0099) | (0.0099) | (0.0099) | (0.0099) | (0.0099) | (0.0099) | (0.0099) |
| # stores $\leq$ 5 km | -0.0006* | -0.0006* | -0.0006* | -0.0006* | -0.0006* | -0.0006* | -0.0006* | -0.0006* |
| | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) |
| $R^2$ | 0.6983 | 0.8285 | 0.8890 | 0.9070 | 0.9477 | 0.8999 | 0.9366 | 0.8380 |
| Observations | 216 | | | | | | | |

Notes: Standard errors in parentheses.*, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All regressions include retail chain fixed-effects. The omitted retail chain for the hypermarkets is Hyper U and for the supermarkets is Super U . The variable *# stores $\leq$ 5 km* counts, for a given store, the number of rivals within 5 km. The variable *log(Household income)* corresponds to the log of the median income per household at the city or *IRIS* level, depending on the location of the store. The specification adopted implies that these last two variables are equal across the product categories. The Breusch-Pagan LM test for error independence supports the resort to a SUR specification, $\chi^2\,(28) = 139.70$ and $p$-value=0.0000. Source: Author's calculations.

Table 4: Descriptive statistics on the price indices

| Retail chain | Fruits & vegetables | Meat | Cooked meat | Cheese | Other dairy product | Grocery item | Alcoholic drink | Soft drink |
|---|---|---|---|---|---|---|---|---|
| Aldi | 79.67 | 69.19 | 110.18 | 103.21 | 82.08 | 67.48 | 77.77 | 88.41 |
| Atac | 119.14 | 103.45 | 93.19 | 107.16 | 95.07 | 104.42 | 101.41 | 114.34 |
| Auchan | 114.28 | 98.84 | 101.59 | 99.95 | 91.86 | 104.10 | 99.97 | 105.26 |
| Carrefour | 99.73 | 97.00 | 96.65 | 101.02 | 97.34 | 104.10 | 100.99 | 99.22 |
| Casino | 119.25 | 104.88 | 104.88 | 107.82 | 105.79 | 109.80 | 103.78 | 105.92 |
| Cdm | 97.91 | 72.86 | 85.61 | 101.07 | 88.32 | 68.93 | 90.38 | 99.57 |
| Champion | 114.44 | 91.90 | 105.88 | 102.76 | 104.33 | 104.89 | 103.64 | 97.56 |
| Ed | 84.00 | 70.00 | 85.00 | 103.00 | 81.00 | 75.00 | 78.00 | 82.00 |
| Géant Casino | 107.86 | 102.57 | 102.56 | 104.54 | 97.48 | 103.52 | 100.48 | 101.59 |
| Hyper U | 98.23 | 93.40 | 103.95 | 103.14 | 106.38 | 106.82 | 101.03 | 94.28 |
| Inno | 135.95 | 109.46 | 115.65 | 109.06 | 113.09 | 112.23 | 103.12 | 116.11 |
| Intermarché | 112.14 | 96.79 | 102.85 | 106.06 | 97.03 | 107.66 | 103.65 | 108.05 |
| Leader Price | 99.77 | 77.17 | 104.56 | 103.12 | 95.73 | 69.65 | 77.88 | 93.56 |
| Leclerc | 103.32 | 93.13 | 98.62 | 103.54 | 98.45 | 102.32 | 100.35 | 98.98 |
| Lidl | 85.07 | 68.99 | 85.61 | 103.21 | 80.33 | 96.08 | 77.77 | 90.13 |
| Marché U | 155.32 | 127.54 | 152.52 | 145.12 | 131.51 | 144.57 | 126.19 | 151.40 |
| Monoprix | 125.27 | 113.49 | 115.21 | 108.26 | 109.72 | 110.68 | 102.77 | 109.98 |
| Norma | 90.00 | 72.00 | 84.00 | 103.00 | 83.00 | 81.00 | 77.00 | 97.00 |
| Shopi | 146.44 | 119.93 | 126.06 | 128.13 | 124.56 | 120.58 | 112.74 | 120.90 |
| Stoc | 115.38 | 103.52 | 100.60 | 101.38 | 107.70 | 102.90 | 104.01 | 99.43 |
| Super U | 127.46 | 93.84 | 105.98 | 110.99 | 104.18 | 113.67 | 105.97 | 103.91 |

Notes: Price indices reported correspond to the average of the indices by product category and retail chain. Source: Author's calculations.

a price flexing strategy in the area of study. Therefore, it is no longer necessary to estimate the missing price indices. Rather, I depart from the price reporting and compute an average price index by category for each hard discounter according to Eq. (1. Because the hard discounters mainly offer first-price products, I remove the national brand products from the shopping basket of the base outlet to compare the price indices based on similar products. In Table 4, I present the average price indices by category and retail chain.

Note that by estimating the missing prices, I inevitably introduce a measurement error in the price variable. Although the estimated price indices represent only 38% of all prices, the results may be biased by this noise. Traditionally, this issue is addressed by applying instrumental variables techniques, but the modeling framework adopted afterwards prevents me from following this approach (see footnote 24). Instead, I draw $N$ values in the 95% prediction interval for each price estimate and re-estimate the demand model to evaluate the sensitivity of the results to this noise. I discuss the procedure and the results in more detail in subsection 6.3, but the important thing to note is that one observes little difference in the estimated demand parameters over the $N$ replications.

# 4 The empirical framework

In this section, I first specify the formulation of the demand model and then I derive the pricing equations from alternative pricing games that are likely to be played by French retailers.

A large majority of the empirical IO literature devoted to the study of supermarket competition and built on discrete-choice demand models assumes that consumers buy the same basket of goods (Smith [2004, 2006]; Dubois and Jódar-Rosell [2010]).[18] Therefore, this simplifying assumption eliminates the composition of the shopping basket as a determinant of consumer demand. However, several empirical studies have emphasized that consumers may react differently to store characteristics depending on whether the consumers are large or small basket shoppers (see, e.g., Bell and Lattin [1998]).[19] In particular, the price responsiveness of consumers is assumed to vary with the expected number of product categories purchased. This assumption implies that competition among stores may be influenced by this component and, in turn, that stores can adopt different pricing strategies according to the type of shoppers located around them (Ellickson and Misra [2008]). As a result, this simplifying assumption may lead to understated or overstated levels of competition.

By recording households' store choices at the category level, the database allows me to specify a demand model that accounts for heterogeneous shopping baskets and thereby to estimate more accurately the competitive forces among large grocery stores. As observed in Table 5, the basket size is an important element of differentiation among households. For the different product categories, the table displays the frequency of purchasing in the supermarket channel and the standard deviation. These statistics indicate that households attend large grocery stores to fulfill needs that vary in their nature and their number. For instance, fewer than half of the households buy some fruits and vegetables in the supermarket channel, but important heterogeneity around this mean is observed. Similarly, for the other product categories, one finds that households express various preferences for the supermarket channel. This finding suggests that depending on the household visiting the stores, large grocery stores do not necessarily compete over the entire set of products offered. Thus, the size of the shopping list is an element that may influence a consumer's store choice decision.

To account for the heterogeneity across households in terms of the product categories purchased in the supermarket distribution channel, I propose a two-step model in which a household chooses its primary shopping destination conditional on a household-specific bundle of product categories (i.e., a shopping basket). In the first step, I estimate the probabilities that a household will purchase the product categories in a large grocery store. I then use these probabilities to weight the price index of the corresponding category such that a household pay greater attention to the prices of the product categories that it usually buys in large grocery stores. The formulation of a weighted average price index of the shopping basket is close to the one adopted by Briesch, Chintagunta, and Fox [2009]. Nonetheless, I differ from these authors with respect to at least two points. First, in my case, the household's choice relies on the decision to go shopping in the supermarket distribution channel, whereas the purchasing occurrence encompasses all types of retail channels in Briesch, Chintagunta, and Fox [2009]. Second, unlike Briesch,

[18]Richards and Hamilton [2006] are one of the few exceptions. They model the consumer sequential choices of which products to buy and where the resulting shopping basket is purchased through a nested CES discrete choice model.

[19]Using the terminology of Bell and Lattin [1998], I refer to a *large* basket shopper to define a shopper who has a relatively high probability of purchase for any given category.

Table 5: Choice of a large grocery store by category (in %)

| Category | Mean | (S.D.) | Category | Mean | (S.D.) |
|---|---|---|---|---|---|
| Fruits & vegetables | 0.4730 | (0.4994) | Other dairy product | 0.8827 | (0.3219) |
| Meat | 0.6145 | (0.4869) | Grocery item | 0.8492 | (0.3580) |
| Cooked meat | 0.5760 | (0.4943) | Alcoholic drink | 0.7033 | (0.4569) |
| Cheese | 0.7939 | (0.4046) | Soft drink | 0.8908 | (0.3120) |

Notes: The number of observations is 1611. Source: Author's calculations.

Chintagunta, and Fox [2009], I adopt a modeling framework that addresses the central issue of cross-category effects on households' store choice decisions. I present in more detail the two steps in the following subsections.

## 4.1 Demand model: Household's retail channel choice

For each household $h$ $(h = 1, \ldots, H)$ one observes its decision to buy a product in category $c$ $(c_h = 1, \ldots, C_h)$ in a large grocery store among a set of $C$ categories. Following the discrete choice literature, I represent the purchase incidence of this product category with a vector $\mathbf{i_h} = \langle i_{h1}, i_{h2}, \ldots, i_{hC} \rangle$ of binary dependent variables. I estimate the probability $\Pr(i_{hc})$ of a single decision through a system of simultaneous probit equations. Let $i^*_{hc}$ denote the underlying latent variable associated with the $c$-th category. The link between the purchase incidence and the latent variable is expressed as follows:

$$i_{hc} = \begin{cases} 1 & \text{if } i^*_{hc} > 0 \\ 0 & \text{otherwise} \end{cases}$$

These latent variables are defined by a linear combination of a set of explanatory variables and an error term. Using the matrix notation, I can write the system as follows:

$$\mathbf{I}^* = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

where $\mathbf{X} = \langle x_1, \ldots, x_p \rangle$ is a $C \times p$ vector of $p$ explanatory variables, $\beta = \langle \beta_1, \ldots, \beta_p \rangle$ is a corresponding vector of parameters of the same dimension and $\varepsilon$ is a $C \times 1$ vector of the error terms that accounts for the unobservable heterogeneity. The choice of a large grocery store for a given product category is supposed to be explained by household characteristics (i.e., age group, house, card, work) and the attractiveness of its surrounding retail environment (Montpellier city, # hypermarket $\leq$ 10 km). Furthermore, the household's decision of whether to buy the $c$-th category in the supermarket distribution channel is presumed to be related to its decisions with respect to the other categories. In other words, a household's choices may be linked not only because of cross-effects across the categories that reflect their complementarities but also because of unobservable factors at the retail channel level that encourage the concentration of purchases (e.g., quality of products and shopping costs). This assumption is corroborated by the calculation of the pairwise tetrachoric correlation coefficients (see Table 6). As we observe,

Table 6: Tetrachoric correlation coefficients of the product categories

| | Fruits & vegetables | Meat | Cooked meat | Cheese | Other dairy product | Grocery item | Alcoholic drink | Soft drink |
|---|---|---|---|---|---|---|---|---|
| Fruits & vegetables | 1.0000 | | | | | | | |
| Meat | 0.7448* | 1.0000 | | | | | | |
| Cooked meat | 0.6473* | 0.9252* | 1.0000 | | | | | |
| Cheese | 0.6447* | 0.7524* | 0.7332* | 1.0000 | | | | |
| Other dairy product | 0.6732* | 0.7751* | 0.7607* | 0.9239* | 1.0000 | | | |
| Grocery item | 0.5582* | 0.6490* | 0.6445* | 0.7414* | 0.8875* | 1.0000 | | |
| Alcoholic drink | 0.3229* | 0.4164* | 0.4623* | 0.4676* | 0.6023* | 0.5830* | 1.0000 | |
| Soft drink | 0.5033* | 0.5799* | 0.5666* | 0.7644* | 0.8795* | 0.8229* | 0.7014* | 1.0000 |

Note: * significance at the 1% level. Source: Author's calculations.

all of the estimated correlations are significant and positive. In addition, the magnitude of the estimates for some pairs of product categories (e.g., meat and cooked meat or cheese and other dairy products) reveals important cross-effects. This finding suggests that, notwithstanding the heterogeneity in the households' shopping habits, the complementarity effects across categories may foster the concentration of purchases in a distribution channel. This concentration should induce extreme behaviors from the households with respect to the supermarket channel, such as either grouping a large part of their purchases together (large basket shopper) or considering this channel principally for top-up shopping (small basket shopper). Therefore, to control for any possible correlations arising from unobservable factors, I assume that the error terms of the latent equations are distributed according to a multivariate normal distribution, $\varepsilon \sim N(0, \Sigma)$, where $\Sigma = \{\rho_{jk}\}$ is the correlation matrix obtained by considering the Cholesky decomposition of the covariance matrix of the errors. That is, $\Sigma = Lee'L'$, where $e$ are independent standard normal random variables, and $L$ is the lower triangular matrix with diagonal elements equal to unity:

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1C} \\ \rho_{21} & 1 & \cdots & \rho_{2C} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{C1} & \rho_{C2} & \cdots & 1 \end{bmatrix}$$

As a result, the outcome for the $C$ different choices for household $h$ is now specified through a Multivariate Probit model (Chib and Greenberg [1998]; MVP hereafter). The probability of the corresponding combination of choices (conditional on parameters $\beta$ and $\Sigma$) is given by:

$$\Pr(\mathbf{I}_h = i_h|\beta, \Sigma) = \Phi_C\left(x\beta_1, \ldots, x\beta_C\right)$$

where $\Phi_C(\cdot)$ denotes the $C$-variate standard normal distribution. The estimated parameters of the MVP model are reported and discussed in Appendix A.

## 4.2 Demand model: Household's store choice

The second part of the demand model is more familiar with respect to the literature on structural models of demand (Berry [1994], Berry, Levinsohn, and Pakes [1995], Nevo [2001]). Given the discrete nature of a household's decision, I follow the standard random utility approach and specify a discrete choice model to assess the determinants of a household's store choice. The households' preferences are assumed to differ because of the heterogeneity in their locations and their tastes over their chosen store characteristics. To account for this flexibility, I define a random coefficients logit model (or mixed logit model), which allows me to estimate more realistic substitution patterns than a simple "logit-type" model. Concretely, the mixed logit model yields flexible estimates of own- and cross-price elasticities by avoiding the problematic independence of irrelevant alternatives (IIA) assumption involved in discrete choice models where heterogeneity is solely captured through the idiosyncratic term.[20]

Let us assume that a household $h$ chooses its primary shopping destination (conditional on a shopping basket) based on the highest utility rule, which is derived from visiting one of the stores $j$ $(j = 1, \ldots, J)$ that are included in its choice set $J_h$, or by choosing an outside option $j = 0$. Recall that a household's choice set contains all stores located within 20 km of its residence. Thus, according to the typical notation for discrete choice models of demand, the indirect utility that a household $h$ that resides in location $L_h$ derives from visiting store $j \in J_h$ located in $L_j$ is:

$$\begin{aligned} U_{hj} = \alpha_0 \widetilde{p}_{hj} + \sum_{g=2}^{6} \alpha_g \widetilde{p}_{hj} d_{hg}^{age} + \lambda_h DIST\left(L_h, L_j\right) + \sum_{m} \gamma_m DIST\left(L_h, L_j\right) z_{hm} \\ + \phi CASH_j + \kappa_h SURF_j + \sum_{s} \sum_{q} \varphi_{sq} v_q d_s^{format} + \xi_f + \varepsilon_{hj} \end{aligned} \tag{3}$$

where $\widetilde{p}_{hj}$ is the household-specific price of the shopping basket of store $j$, $d_{hg}^{age}$ is a dummy variable equal to one if the head of household $h$ belongs to age group $g$, $DIST\left(L_h, L_j\right)$ is the Euclidean distance between a household's residence $(L_h)$ and store location $(L_j)$ (the disutility to travel is assumed to vary among households), $z_{hm}$ are $M$ observed household characteristics, $CASH_j$ is the number of cash registers per one hundred square meters in store $j$ and $SURF_j$ is the store size of store $j$. Similar to the distance parameter, the parameter $\kappa_h$ of store size is assumed to vary by household. I also include a dummy variable $d_s^{format}$ for one of the $S$ store formats ($s$=hypermarket, supermarket, hard discount store, large convenience store) and interact this dummy variable with $Q$ variables that represent a mix of household and store characteristics denoted by $v$. Finally, $\xi_f$ is an index of unobserved (to the econometrician) retail chain attributes and $\varepsilon_{hj}$ is the idiosyncratic term that is assumed to be i.i.d. according to a type I extreme value distribution.

[20]See Train [2003] for further insights.

Price sensitivity varies among the six age classes of the household head (with the youngest taken as the reference). Thus, the coefficient $\alpha_0$ corresponds to the marginal utility of price of a "representative" household, and a deviation from this mean depends on the coefficient of the interaction of the price with the age class of the household head. Note that following the first step of the model, the price variable is defined as the sum of the price indices, $\widehat{p}_{jc}$, multiplied by the category purchase probabilities, $\Pr(I_{hc})$, and the share of the mean expenditure in category $c$, $\varpi_c$, used to determine the primary shopping destination:

$$\widetilde{p}_{hj} = \sum_{c=1}^{C} \Pr(I_{hc})\, \widehat{p}_{jc} \varpi_c \tag{4}$$

Similar to the price coefficient, the distance coefficient $\lambda_h$ is supposed to vary by household. Rather than introducing the heterogeneity among households through interaction terms, I specify a random coefficient on distance, which is more appropriate for accounting for the diversity of the households' location. $\omega$ denotes the unobserved household characteristics that influence their traveling habits. Precisely, I assume that the distance coefficient is normally distributed and independent of the idiosyncratic term $\varepsilon$. Again, the distance variable is interacted with the observed household characteristics (e.g., the number of cars, the type of residence and whether a household's residence is in a rural town).

I also introduce a random coefficient for the store size variable ($SURF$), which is distributed according to a normal distribution. Additionally, I enrich the specification of the indirect utility function by interacting the store's format with the number of rivals (by store type) within 1 km and 2.5 km radii to account for the competitiveness of the store's environment. Unobserved store characteristics (e.g., shelf display or assortment) are captured by the fixed-effects $\xi_f$. I argue that these unobserved characteristics reflect essentially national strategies enacted by the retailers for their retail chains. Thus, these common shocks are captured by fixed-effects defined at the retail chain level. As usual, I assume that the households value these unobserved characteristics in an identical manner.

Similar to the "outside good" in classical demand models, the households may decide to visit retailing channels other than large grocery stores (e.g., small convenience stores, specialized stores, and market places) or not to purchase the food categories, which is resumed through the outside option $j = 0$. Without additional information on the characteristics of this alternative, I decide to normalize the characteristics of the outside option to zero.

According to the highest utility rule, household $h$ visits store $j$ with the following probability:

$$P_{hj} = \int_{A_{hj}} dF(\varepsilon_h) dF(\omega_h)$$

where $A_{hj} = \{(\varepsilon_h, \omega_h) \,|U_{hj} > U_{hl}; l \neq j\}$, and $F(\cdot)$ denotes the distribution function.

## 4.3 Supply side: Alternative pricing models and selection tests

I describe the different pricing rules that the retailers may apply. I assume that the stores compete in prices and set their prices simultaneously, conditionally on their characteristics supposed chosen prior to this decision (e.g., location, store size or quality). To simplify, I assume that the store managers seek to maintain the price competitiveness of their store across all product categories. That is, the store managers think in terms of the price positioning of the shopping basket and do not adopt category management. Thus, the prices that result from this behavior are a Nash equilibrium of the game. By deriving the pricing equations from the first-order conditions of the profit maximization problem, I recover the stores' marginal cost and consequently compute their price-cost margin.

Rather than assuming an arbitrary pricing behavior, I estimate the stores' marginal cost under alternative pricing models that reflect different degrees of potential coordination across the stores and determine which model fits the data best. I depart from the most competitive case, where the prices are set at the store level (model 1), meaning without ownership consideration. Next, I consider the case in which the integrated stores set their prices at the retail chain level, while the stores belonging to cooperative groups fix their prices individually (model 2). Then, I regard the scenarios in which the pricing decision occurs at the retail chain and retail group levels (models 3 and 4, respectively). I also examine the possibility of spatial collusion by assuming that the stores located nearby one another maximize their joint profits. Precisely, the definition of spatial collusion covers three cases corresponding to a store's cooperative behavior with its nearest, two nearest or three nearest rivals (models 5 to 7). Finally, I consider the extreme case in which the stores in the area of study collectively behave as a monopolist (model 8).[21] Each pricing model is solved as a function of the demand parameters and leads to a specific estimation of the stores' marginal cost. In the following, I derive the expression of the stores' marginal cost for the general case of multi-store retailers. The stores' marginal cost formula for other pricing models are derived by simply adopting a different definition of the ownership matrix $T$ defined below.

Consider the problem of a retailer $R$ that sets its prices in its stores $j = 1, \ldots, \mathcal{J}_R$. The profits of the retailer $R$ are:

$$\Pi_R = \sum_{j \in \mathcal{J}_R} (p_j - c_j) M_j s_j (\mathbf{p}) - C_j \tag{5}$$

where $c_j$ denotes the constant marginal cost of selling a unit of a shopping basket for store $j$, $M_j$ is the size of the market for store $j$, $s_j (\mathbf{p})$ is the market share of $j$ and $C_j$ is a fixed-cost.

If a pure-strategy Nash equilibrium in prices exists, the first-order condition for a typical store

[21] One important feature of the considered market is that the hard discount chains use uniform pricing. Consequently, for all of the pricing models, I impose that the prices of the hard discount stores belonging to the same retail chain are set so as to maximize the joint profits of its members. As a result, the definitions of the pricing models presented here do not apply to hard discount stores.

$j$ belonging to $\mathcal{J}_R$ is:

$$s_j(\mathbf{p}) + \sum_{l \in \mathcal{J}_R} T(p_l - c_l) \frac{\partial s_l(\mathbf{p})}{\partial p_j} = 0,\ j = 1, \ldots, \mathcal{J}_R \tag{6}$$

where $T$ corresponds to the ownership matrix, with general element $T(j,l)$ equals to one if both stores $j$ and $l$ belong to the same retail group and zero otherwise. This gives a system of $\mathcal{J}_R$ equations. After defining $\Delta$ as the retailer's response matrix with element $(j,l) = \partial s_l(\mathbf{p}) / \partial p_j$, I can express the stores' price-cost margin of retailer $R$ in matrix notation by stacking up the first-order conditions and rearranging the terms:

$$(\mathbf{p} - \mathbf{c}) = -\left[\mathbf{T} \otimes \mathbf{\Delta}(\mathbf{p})\right]^{-1} \mathbf{s}(\mathbf{p}) \tag{7}$$

where $\otimes$ corresponds to the Kronecker product. It follows that the estimated stores' marginal cost depends exclusively on the parameters of the demand system and the market conduct assumption:

$$\hat{\mathbf{c}} = \mathbf{p} + \left[\mathbf{T} \otimes \mathbf{\Delta}(\mathbf{p})\right]^{-1} \mathbf{s}(\mathbf{p}) \tag{8}$$

At this point in the estimation procedure, I obtain different sets of the stores' estimated marginal cost (i.e., one for each pricing model). The next step consists of determining which pricing model explains the data best. For this purpose, I adopt the approach followed by Bonnet and Dubois [2010] and conduct pairwise non-nested tests (Rivers and Vuong [2002]). The details of the procedure are provided in Appendix B.

It is worth noting that manufacturers are absent from this scenario. This absence implies that the stores' marginal cost estimated from Eq. (8) are derived under the assumption that manufacturers do not influence the first-order conditions of retailers. This assumption is relevant if manufacturers offer two-part tariff contracts by setting their prices equal to their marginal costs such as retailers act as residual claimants. In this case, the issue of double marginalization ensuing from this vertical relationship may vanish (see Rey and Vergé [2010]), and the store's estimated marginal cost coincides with its true value (i.e. retail cost + cost of goods). Otherwise, alternative vertical pricing models (e.g., linear tariffs and two-part tariffs without resale price maintenance) induce a distortion between a store's estimated and its true marginal cost because the determination of the manufacturers' margin influences the price-setting of the retailers. As the model is defined at the shopping basket level and does not refer explicitly to a set of manufacturers, I do not specify a vertical relationship and thus assume that manufacturers are neutral to the retailers' price-setting behaviors. Thus, we must keep in mind when discussing the results that depending on the supply contracts adopted by the parties, a deviation might exist between the stores' estimated margin and their true margin.[22]

[22] Although resale price maintenance (RPM) is *per se* illegal in France, it is well-documented that the adoption of the Galland Act has indirectly promoted this practice. Bonnet and Dubois [2010] have shown that manufacturers use nonlinear pricing contracts with RPM for the specific retail market of bottled water. Contrary to my

# 5 Identification and estimation strategy

The demand parameters expressed in Eq. (2) and Eq. (3) are estimated with simulated maximum likelihood (SML). Let $\theta^{MVP} = \{\beta, \rho\}$ and $\theta^{MXL} = \{\alpha, \lambda, \gamma, \phi, \kappa, \varphi\}$ denote the set of demand parameters corresponding to the multivariate probit model and the mixed logit model, respectively. Conditional on $\theta^{MVP}$, the log-likelihood function of the combination of the category purchase incidences may be written as:

$$\mathcal{L}\left(I; \theta^{MVP}\right) = \sum_{h} \mathbb{1}\left[\log \Phi_{C_h}\left(X; \theta^{MVP}\right)\right] \tag{9}$$

where $\mathbb{1}\left(\cdot\right)$ is the indicator function. The multivariate normal probability $\Phi_{C_h}\left(\cdot\right)$ does not have a closed-form formula because of the problem of high order multivariate normal integrals. A standard approach consists in approximating the probability by simulations. To that end, I employ the so-called Geweke-Hajivassilou-Keane (GHK) simulator and draw $R$ values from an upper-truncated standard normal distribution.
The log-likelihood of store choice, conditional on $\theta^{MXL}$ and $I$, is given by:

$$\mathcal{L}\left(Y; \theta^{MXL}, I\right) = \sum_{h} \sum_{j=0}^{J_h} \mathbb{1}\left[\log\left(s_{hj}\left(\theta^{MXL}, I\right)\right)\right] \tag{10}$$

where $Y$ is the vector of store choices, and $s_{hj}$ is the probability that household $h$ chooses store $j$ as its primary shopping destination. Because I specify a mixed logit model, the latter is defined as:

$$s_{hj}\left(\theta_h^{MXL}, I\right) = \frac{e^{V_{hj}\left(\theta_h^{MXL}, I\right)}}{\sum_{j=0}^{J_h}\left(e^{V_{hj}\left(\theta_h^{MXL}, I\right)}\right)} \tag{11}$$

Unfortunately, this closed-form expression is conditional on $\theta_h^{MXL}$. Because I do not know the true value of $\theta_h^{MXL}$, I need to integrate Eq. (11) over all off the possible values of $\theta_h^{MXL}$. The unconditional store choice probability is then approximated by numerical simulation:

$$\widetilde{s}_{hj} = \frac{1}{R} \sum_{r=1}^{R} s_{hj}\left(\theta_h^{MXL,r}, I\right) \tag{12}$$

I utilize simulation to evaluate accurately the probability terms in both parts of the model. By proprieties, these simulated probabilities are unbiased, and their variance diminishes as the number of draws increases. Rather than using random draws, I follow recent advances in simulation methods and generate 100 Halton draws. Note that I keep the same set of draws for each iteration.[23]

postulate, the authors find that manufacturers act as residual claimants. Nonetheless, the question remains whether manufacturers do not sell at cost, in exchange of positive fixed fees, in the other sectors of the French agro-food industry.

[23]A consensus exists in the literature regarding the superiority of Halton draws over random draws (see Bhat

At this point, two estimation strategies are conceivable: a full information maximum likelihood (FIML) procedure or a two-step method. The decision to adopt one of these strategies reflects the trade-off between a gain in efficiency and the burden of computation required to achieve it. The joint estimation of the log-likelihood functions (see equations 9 and 10) gives the true standard errors of the estimates, whereas a sequential estimation introduces a measurement error for the estimates of the second model. Hence, adopting the two-step approach would bias the variance-covariance matrix of the estimates of the mixed logit model, as the price of the shopping basket is computed from the estimated probabilities of the category purchase incidences. Nonetheless, the aim of the empirical model is to provide an accurate estimation of the substitution effects, which results from the demand parameters. Efficient estimates then appear to be a second-order concern. In addition, the computation time needed to estimate the multivariate probit model for eight categories is sizeable. As a result, I decide to adopt a two-step approach and adjust the standard errors of the mixed logit model with the correction method proposed by Murphy and Topel [1985] (see Appendix C for further details).

Another (potential) source of bias must be controlled to obtain valid estimates. As stressed many times in the empirical IO literature, the estimation of demand parameters may suffer from an endogeneity bias depending on whether managers choose some product characteristics (store characteristics, in our case) based on attributes that are unobserved by the researcher (Berry [1994], Berry, Levinsohn, and Pakes [1995]). In this case, the variables of concern are correlated with the error term, and we have a traditional endogeneity issue. Past research has emphasized the strong presumption of price endogeneity because of the difficulty of controlling for the product characteristics (especially the characteristics describing product quality) that influence the pricing decision.

One way to control for unobserved store quality is to introduce retail chain fixed-effects $\xi_f$. In the present context, I am not concerned by time-varying unobserved store quality. The occurrence of an endogeneity problem depends on the existence of inter-temporal unobserved store characteristics that influence the store manager's price-setting decision. By introducing retail chain fixed-effects, I assume that the unobserved quality of stores (e.g., advertising or shelf display) depends on business strategies implemented at the retail chain level. Thus, I control for endogeneity bias as long as the unobserved store attributes do not deviate from the mean of their respective retail chain. However, it is likely that other unobserved store-specific attributes influence the price positioning of the stores. Therefore, to soften the previous assumption, I incorporate additional store-specific variables, denoted by $v_q$ in the household indirect utility function, to control for the demand and the competitive environments. By introducing these variables, I expect to control for the main unobserved drivers of the price decision and there-

[2001], Train [2003], Chiou and Walker [2007], among others). For a given number of draws, Halton draws achieve greater efficiency and coverage than random draws because the observations of a Halton sequence are negatively correlated.

fore, the endogeneity issue.[24]

The demand parameters are identified through several sources of variation. First, each choice occasion differs from the others because of the heterogeneity observed in household characteristics. This heterogeneity permits to identify the parameters $\{\alpha, \gamma\}$. Further, for a given choice occasion, a household faces a set of stores whose characteristics differ from one another. Hence, the average valuation of store characteristics identifies $\{\phi, \kappa_h\}$ and, for the same reason, the unobserved characteristics of each retail chain $\xi_f$. Note that by specifying the fixed-effects at the retail chain level rather than at the store level, I avoid the identification problem that may arise for the parameters of store characteristics, as the store's dummy variable should be strongly correlated with the observed store characteristics.

The main source of variation across choice occasions is derived from the heterogeneity in the spatial distribution of the households' residence and the stores' location, which induces different choice sets among households. As a result, one observes different distributions of distance among households located in different cities (or *IRIS*). This finding enables me to identify the parameter $\lambda_h$. Finally, because the price of the shopping basket is specific to a household and varies across the alternatives for a given choice set, there is sufficient variation to identify the parameters associated with $\widetilde{p}_{hj}$.

# 6 Results

## 6.1 Mixed logit demand model

This subsection presents the estimation results for the household's store choice model defined in Eq. (3). The simulated maximum likelihood estimates are reported in Table 7. The estimated parameters must be interpreted with respect to the outside option. Almost all of the coefficients are both statistically and economically significant. Overall, we note that shopping patterns differ significantly across the households, store formats and areas of living. As expected, the households express a disutility of price and distance. More precisely, we observe that the households whose heads are between 30 to 39 years old (age group 3) appear less sensitive to price than the youngest age class. This heterogeneity among household behavior is also notable in the expression of the disutility of traveling. As defined in the indirect utility specification (see Eq.

[24]The data prevent me from adopting correction methods that address more directly the price endogeneity bias. One of these methods, known as the "fixed-effects" approach (Berry, Levinsohn, and Pakes [1995], Nevo [2001]), addresses this issue by introducing an alternative-specific fixed-effect that separates the market-specific valuation of unobserved attributes from the mean valuation of the alternative, with the aim of applying traditional instrumental methods. However, to identify the parameters, I need sufficient variations in the data (i.e., a cross-section of markets or a longitudinal data set). As a result, I rule out this correction method because I can observe only one market for a given period of time in the data. Alternatively, I could refer to the "control function" approach developed by Petrin and Train [2010]. Its principle consists of regressing in a first step the price variable on all exogenous factors. The residuals are then plugged into the indirect utility function along with the price variable. Because the price variable already includes an error term (as it corresponds to an estimated variable), I cannot use this two-stage error correction method. Therefore, I reject this approach in favor of the retained specification.

(3)), the marginal valuation of distance may vary by the observable and unobservable household characteristics. This possibility allows me to reveal that the high disutility of traveling (the estimated mean of the distance coefficient distribution is -2.1542) intensifies for people living in a house and in a rural town. Conversely, the higher the number of cars owned by a household, the lower the household's sensitivity to distance is. Nonetheless, the statistical significance of the estimated standard deviation of the random coefficient on distance suggests that the household's willingness to travel must be explained by other individual characteristics that are not observed. Similar to the distance coefficient, the store size parameter is allowed to vary by household. On average, the households positively value the log of a store's selling area (mean=1.3376), although important heterogeneity around this mean is observed (S.D.=2.4698). Moreover, we note that households seem to pay great attention to the number of cash registers per one hundred square meters, as suggested by the estimated parameter of this variable. I interpret this variable as a proxy for the inverse of the waiting time at the checkout lines.

I introduce several interaction terms with store formats to capture more accurately the variety in shopping patterns. As shown in Table 7, we note that the households living in Montpellier are more likely to choose the outside option at the expense of a large grocery store. This result suggests that these households are sensitive to the number of alternative options located nearby and aggregated into the outside option (e.g., specialized stores). In addition, for a given store, I count the number of rivals by format and by distance bands of 1 km and 2.5 km to interact these variables with the store's format. By doing so, I investigate the nature of the competition among store formats with respect to the distance that separates them and also control for the endogeneity bias discussed above by accounting for the factors that may influence the store manager's price-setting behavior.

Interestingly, it appears that the supermarkets and the hard discount stores compete fiercely when they are close together (within 1 km), but this competitive pressure vanishes for the supermarkets as soon as the radius is extended to 2.5 km. One also notes that the hard discount stores (hypermarkets) take advantage of the commercial attractiveness generated by the hypermarkets (supermarkets). This relationship indicates that a unilateral complementarity effect may exist between these formats depending on the distance between them.

One advantage that a mixed logit model has over a simple logit model is that it provides accurate estimates of substitution patterns, as cross-price elasticities vary by the competing alternatives. I determine the elasticity of the market share of a given store $j$ according to the following expression:

$$\frac{\partial s_j}{\partial p_l}\frac{p_l}{s_j} = \begin{cases} \frac{p_j}{s_j}\int \alpha_h s_{hj}\left(1 - s_{hj}\right) dF(\omega) & \text{if } l = j \\ -\frac{p_j}{s_j}\int \alpha_h s_{hj} s_{hl} dF(\omega) & \text{otherwise} \end{cases} \tag{13}$$

The second and third columns of Table 8 report the average own-price elasticity and the mean of the cross-price elasticities by retail chain. Overall, the mean of the distribution of own-price elasticities across stores is -8.02, with a standard deviation of 2.68. Upon closer examination, we observe that the hypermarkets appear less sensitive to a change in their price (except for

Table 7: Results from the mixed logit model

| Variable | | Interactions with store formats: Hypermarket | Supermarket | Hard discount | Convenience store |
|---|---|---|---|---|---|
| Price | -1.7102*** | | | | |
| | (0.5452) | | | | |
| Price × age group 2 | 0.2655 | | | | |
| | (0.2599) | | | | |
| Price × age group 3 | 0.5621** | | | | |
| | (0.2273) | | | | |
| Price × age group 4 | 0.2400 | | | | |
| | (0.2273) | | | | |
| Price × age group 5 | 0.0943 | | | | |
| | (0.2546) | | | | |
| Price × age group 6 | 0.1998 | | | | |
| | (0.2674) | | | | |
| Ln(distance) | | | | | |
| Mean | -2.1542*** | | | | |
| | (0.1209) | | | | |
| S.D. | 0.8406*** | | | | |
| | (0.0880) | | | | |
| Ln(distance) × # cars | 0.1189* | | | | |
| | (0.0630) | | | | |
| Ln(distance) × house | -0.2420** | | | | |
| | (0.1173) | | | | |
| Ln(distance) × rural town | -0.7544*** | | | | |
| | (0.1677) | | | | |
| Cash registers | 1.6956*** | | | | |
| | (0.4502) | | | | |
| Ln(surface) | | | | | |
| Mean | 1.3376*** | | | | |
| | (0.3478) | | | | |
| S.D. | 2.4698*** | | | | |
| | (0.2708) | | | | |
| Constant | | 39.9938*** | 38.7530*** | 35.4570*** | 39.6400*** |
| | | (6.7247) | (6.4965) | (5.6602) | (7.6948) |
| # hypermarket ≤ 1 km | | – | – | 0.9724* | – |
| | | | | (0.4996) | |
| # supermarket ≤ 1 km | | – | – | -1.0546** | – |
| | | | | (0.4212) | |
| # hard discount ≤ 1 km | | 0.4421 | -0.4741** | – | – |
| | | (0.3588) | (0.1976) | | |
| # hypermarket ≤ 2.5 km | | – | -0.3254 | 0.5851*** | – |
| | | | (0.3283) | (0.2214) | |
| # supermarket ≤ 2.5 km | | 0.5104*** | – | -0.5444*** | – |
| | | (0.1240) | | (0.1644) | |
| # hard discount ≤ 2.5 km | | -0.2441*** | -0.0417 | – | – |
| | | (0.0696) | (0.1020) | | |
| Single household | | -1.8787 | -2.0037 | -2.3837 | -0.6656 |
| | | (1.8408) | (1.7047) | (1.6193) | (1.7586) |
| Montpellier city | | -7.3948*** | -5.5955*** | -4.4682** | -3.8296** |
| | | (2.0413) | (1.9012) | (1.8004) | (1.9492) |
| Rural town | | 0.0368 | 1.4774 | 2.9466* | – |
| | | (1.7009) | (1.5973) | (1.6168) | |
| Observations | 1,611 | | | | |
| Log-likelihood | -3,418.26 | | | | |

| **Choice set radius** | Sensitivity of estimated coefficients to households' choice set definition: Price | Ln(distance) – Mean coef. |
|---|---|---|
| 16 km | -1.9064*** (0.5713) | -2.1479*** (0.1253) |
| 18 km | -1.7735*** (0.5549) | -2.1366*** (0.1206) |
| 22 km | -1.9637*** (0.5135) | -2.1771*** (0.1194) |
| 24 km | -2.0272*** (0.5049) | -2.2142*** (0.1186) |

Notes: Corrected standard errors in parentheses.*, **, *** indicate significance at the 10%, 5%, 1% level, respectively. Results with 100 Halton draws and retail chain fixed-effects. The price variable is divided by 10. The omitted retail chain of hypermarkets is Leclerc. The omitted retail chain of supermarkets is Casino. The omitted retail chain of hard discounts is Norma. The omitted retail chain of convenience stores is Marché U. Source: Author's calculations.

Table 8: Retail chains' price elasticities and (gross) margin

| Retail chain | Elasticities | | PCM (in %) | | | Obs. |
|---|---|---|---|---|---|---|
| | Own-price | Cross-price | Avg. | S.D. | Median | |
| **Hypermarket** | | | | | | |
| Auchan | -9.84 | 0.19 | 10.16 | – | 10.16 | 1 |
| Carrefour | -6.07 | 0.56 | 22.57 | 14.93 | 15.84 | 4 |
| Géant Casino | -7.80 | 0.20 | 12.91 | 1.04 | 12.91 | 2 |
| Hyper U | -11.22 | 0.06 | 8.91 | – | 8.91 | 1 |
| Inno | -5.69 | 0.20 | 17.58 | – | 17.58 | 1 |
| Intermarché | -8.17 | 0.24 | 14.53 | 5.77 | 14.53 | 2 |
| Leclerc | -6.92 | 0.27 | 14.46 | – | 14.46 | 1 |
| **Supermarket** | | | | | | |
| Atac | -10.90 | 0.17 | 9.18 | – | 9.18 | 1 |
| Casino | -10.40 | 0.13 | 9.64 | 0.50 | 9.63 | 3 |
| Champion | -10.14 | 0.05 | 9.95 | 0.90 | 9.95 | 2 |
| Intermarché | -7.47 | 0.13 | 16.45 | 9.46 | 12.55 | 11 |
| Leclerc | -8.37 | 0.24 | 11.94 | – | 11.94 | 1 |
| Monoprix | -9.40 | 0.12 | 10.64 | – | 10.64 | 1 |
| Stoc | -8.70 | 0.18 | 13.28 | 5.93 | 10.24 | 4 |
| Super U | -5.77 | 0.09 | 20.87 | 8.60 | 20.87 | 2 |
| **Hard discount** | | | | | | |
| Aldi | -8.07 | 0.11 | 13.26 | 0.60 | 13.26 | 2 |
| Cdm | -8.92 | 0.02 | 11.24 | 0.47 | 11.24 | 2 |
| Ed | -7.12 | 0.07 | 17.06 | 7.66 | 13.21 | 4 |
| Leader Price | -7.88 | 0.09 | 14.21 | 3.92 | 12.28 | 4 |
| Lidl | -8.51 | 0.06 | 12.26 | 1.32 | 11.76 | 5 |
| Norma | -7.47 | 0.11 | 13.62 | 1.78 | 13.62 | 2 |
| **Convenience store** | | | | | | |
| Marché U | -9.09 | 0.03 | 13.24 | 5.44 | 13.24 | 2 |
| Shopi | -9.58 | 0.04 | 13.16 | 7.03 | 8.39 | 3 |

Notes: Price elasticities and price-cost margins (PCM) are averages by retail chain. Standard deviations and medians of the PCM are also displayed. The PCM reported correspond to the preferred pricing model and are computed following the expression $(p-c)/p$. Source: Author's calculations.

Auchan and Hyper U) than the other store formats, although important heterogeneity within formats and retail chains has yet to be considered.[25] Likewise, a given price decrease in a hypermarket is more profitable for its rivals than a price decrease in a convenience store, as suggested by the cross-price elasticities.

A sample of the estimates of own- and cross-price elasticities is reported in the Online Appendix. Note that some cross-price elasticities are null because certain pairs of stores cannot belong to the same choice set.

### 6.2 Price-cost margins

Following the previously described menu approach, I select the preferred pricing model based on the results of the Rivers and Vuong tests. Table 9 displays the test statistics of the pairwise comparisons among the alternative models. A column (row) model is rejected in favor of a row (column) model if the test statistic value is below -1.64 (above 1.64) at the significance level of 5%. Using this decision rule, I find that the store-level pricing model (model 1) is never rejected in favor of the other models. Therefore, model 1 is the preferred model.

Although quite unexpected given the network structure of French retail groups, this result suggests that stores do not internalize the cannibalization effect of rival stores operating under the same retail chain. This finding is explained by several factors that lead stores to adopt a pricing policy imperfectly coordinated with the other stores in their chain. First, the existence of independent shopkeepers and affiliated stores (alongside company-owned stores) naturally implies that some store manager set prices to maximize their own profits. Even if the "independent" stores of a retail chain source through their central purchasing unit which provides them a recommended resale price, they are free to choose their own retail price.[26] Thus, the managers of independent stores are not compelled to cooperate with the other stores belonging to the same retail chain. Second, all stores do not source *exclusively* through their central purchasing unit. A non-negligible part of their assortment is provided by local producers (especially for perishable products). These "local" products have the advantage of offering greater flexibility in the negotiations and even greater wiggle room for fixing resale prices. Third, a retail chain may internalize the substitution effects among its stores by utilizing vectors of its business strategy other than price (and not accounted for in the model). For instance, advertising, loyalty programs or sales may be more relevant levers for softening the competition among the stores of the same chain. Finally, it is important to stress that according to recent retailers' statements, this result seems consistent with industry practice.[27] Thus, one may argue that assuming a tacit coordination among the stores of a retail chain is not necessarily an appropriate hypothesis if one focuses on local competition.

Moreover, the results of the Rivers and Vuong tests yields valuable information on the determinants of a store's market power. As demonstrated by the non-nested tests, the pricing models built upon the assumption of spatial collusion (models 5, 6, 7) are rejected in favor of model 1. This finding implies that tacit collusion among stores located nearby one another may be excluded from the potential sources of a retailer's market power and therefore, as a consequence

[25]The variation in a store's market share resulting from a change in its price depends both on the store characteristics and its rivals' locations. As the stores of the same retail chain do not face similar competitive environments, their price variations affect their market shares in different ways (for identical characteristics).

[26]The recommended retail price is determined on the basis of the wholesale price and a product-specific margin and adjusted according to the degree of local competition.

[27]For instance, Michel-Edouard Leclerc (the chief executive of the eponym French retail chain) states in the press: "*Nantes is one of the cheapest cities in France, because there are 5 Leclerc franchisees, 10 Système U franchisees and as many Intermarché franchisees who are at each other's throats!*" (*Libre-Service Actualités (LSA)*, March 6, 2008).

Table 9: Rivers and Vuong test statistics for pricing model selection

| $H_1/H_2$ | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
|---|---|---|---|---|---|---|---|
| Model 1 | **-2.50** | **-2.76** | **-4.30** | **-3.15** | **-2.97** | **-3.03** | **-5.08** |
| Model 2 | | -2.29 | -3.92 | -1.46 | -2.85 | -2.94 | -5.08 |
| Model 3 | | | -3.41 | 1.77 | -1.96 | -2.25 | -5.08 |
| Model 4 | | | | 3.43 | -1.38 | -1.81 | -5.07 |
| Model 5 | | | | | -2.76 | -2.87 | -5.08 |
| Model 6 | | | | | | -0.75 | -5.07 |
| Model 7 | | | | | | | -5.08 |

Notes: The test statistic is $T_n = \frac{\sqrt{n}}{\sigma_n}\left(Q_n^h\left(\widehat{\theta}_n^h\right) - Q_n^{h'}\left(\widehat{\theta}_n^{h'}\right)\right) \rightarrow N(0,1)$. A test statistic value below (above) the critical value of -1.64 (1.64) means that the row model $(h)$ is better (worse) than the column model $(h')$ at the significance level of 5% (test statistics distributed standard normal) ; otherwise we cannot reject the null hypothesis that the competing models are asymptotically equivalent. Source: Author's calculations.

of a low degree of competition.

Given the determination of the preferred model, I only comment on the implied marginal costs under this scenario. The right side of Table 8 reports the price-cost margins recovered by retail chain when price fixing occurs at the store level (model 1). It is worth noting that the estimated marginal cost of a supermarket located at the limit of our area of study gives a negative outcome (and thus, a margin higher than 100%), which led me to exclude it from the rest of the analysis.[28] Not surprisingly, we first observe that the outcome of the Bertrand-Nash competition among the stores is far from corresponding to a perfect competitive environment, as the margins are substantially higher than zero. Depending on the retail chain, the store's price-cost margin represents, on average, between 8.91% and 22.57% of the marginal cost (i.e., the costs of production and distribution).[29] The mechanism by which some large grocery stores exert significant market power is easily comprehensible. According to oligopoly theory, a firm's markup is inversely proportional to its price elasticity. If we refer to the estimated elasticities, we note that the retail chains less sensitive to price variations are more likely to set higher prices relative to their marginal cost. Thus, the *good* offered by those retail chains is more differentiated than the *good* offered by their rivals because of factors such as product range, advertising, store amenities and/or location. Considering that, I find that the retail chains Carrefour and Super U offer the most profitable product-mix (i.e., store characteristics, price strategy, and location).
More generally, what conclusion can be drawn from these estimates? Principally, I conclude that the average level of profitability in Montpellier AU is not too excessive, especially when

[28]This outlier results from the definition of the geographical boundaries of the survey. This definition truncates the choice set of households living in the same city as this store. Hopefully, this misspecification in the households' choice set appears only for this city.

[29]The relevance of these figures must be assessed by comparing them with the accounting margins published by the retail groups. For instance, Carrefour (Carrefour, Stoc, Marché Plus and Ed, for the most important retail chains) and Promodès (Continent, Champion and Shopi) have reported an average gross margin of 21.1% and 19.3% for their group in 1999, respectively. These values do not account for the distribution costs and can therefore be considered as an upper bound of the true estimated margins (see Nevo [2001]). However, they are calculated at the national level and encompass all of the store formats. Consequently, they may be used only as a proxy of the upper bound of the true estimated margins. Under these limits, these figures inform about the relevance of my estimates.

compared with the profitability levels of other retail sectors (e.g., department stores (36.4%), homewares (38.6%), clothing and shoes (43.5%), from *INSEE's retail census*). Looking abroad for a sectorial comparison, I turn to the UK market, which presents a similar market structure. Following the gross margins reported by Smith [2004, see Tab.3], I note that the levels of profitability are close in these two countries, albeit slightly in favor of the French retailers (around 3 percentage points above). Nonetheless, the average profitability masks important disparities among the stores of the same format and of the same retail chain, as shown in Table 8. For example, the distribution of the margins for the Intermarché supermarket chain (i.e., the most represented retail chain in this area) lies between 9.19% and 42.62%. Beyond the differences in the store characteristics and marginal costs, which are small within a chain, one of the main reasons for this heterogeneity is the inclusion of market geography in the model. Depending on their location, stores face different demand and competitive environments that influence both their "market share effect" as well as their "competitive effect" in a Hotelling [1929] sense. As a result, the local monopoly power enjoyed by large grocery stores varies substantially depending on their location.
Table 10 further investigates this relationship by regressing a store's price-cost margin on various measures of proximity among the stores while controlling for the unobserved retail chain characteristics and store formats. The estimated coefficients confirm that the spatial distribution of rivals noticeably impacts a store's market power, regardless of the chosen measure. For instance, the store's price-cost margin in models (ii)-(v) is regressed on the number of competitors (defined as stores or retail chains) counted by a distance band of 5 and 10 km. The coefficients show that the greater the number of rival stores (or retail chains) located within 10 km, the lower the store's price-cost margin is. Using model (ii), I observe that each store entry between 1 and 5 km reduces a store's price-cost margin by 1.81% on average. The decrease becomes more significant (2.96%) when a store competes with a new retail chain, as shown by model (iv). Because there are 1.18 stores per retail chain on average, these figures indicate that a store suffers more when competing with a store belonging to a rival retail chain. The importance of the location of rivals is again supported if one explains the market power of a store by the distance that separates it from its closest competitor (model vi) or by the cumulated distance between the store and its three nearest competitors (model vii). The same conclusion may be drawn if, instead of the number of rivals between 1 and 5 km, one accounts for the sum of the selling areas (model viii). More broadly, this analysis revisits the relationship between market concentration and prices in the grocery retail industry. This new empirical evidence reinforces the assertion that local concentration positively affects prices (see, e.g., Asplund and Friberg [2002] or Barros, Brito, and de Lucena [2006] for similar conclusions).

In sum, the findings indicate that competition in the French grocery retail industry is highly localized. The competitive pressure exerted by large grocery stores is limited to a few kilometers. By analyzing thoroughly the intensity of the competition in this sector for a given metropolitan

Table 10: Impact of rivals' locations on a store's price-cost margin

| Dependent variable: (log) price-cost margin (%) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | (i) | (ii) | (iii) | (iv) | (v) | (vi) | (vii) | (viii) |
| Constant | -2.1648*** | -1.9631*** | -1.8624*** | -1.9382*** | -1.8785*** | -2.3612*** | -2.3832*** | -1.9122*** |
| | (0.2082) | (0.2074) | (0.2114) | (0.2089) | (0.2232) | (0.2126) | (0.1987) | (0.2149) |
| Fixed-effects $\widehat{\xi}_f$ | -0.0091 | -0.0080 | -0.0167 | -0.0084 | -0.0192 | -0.0117 | -0.0059 | -0.0068 |
| | (0.0386) | (0.0363) | (0.0349) | (0.0361) | (0.0351) | (0.0368) | (0.0351) | (0.0364) |
| # stores $(1, 5)$ | | -0.0181*** | | | | | | |
| | | (0.0062) | | | | | | |
| # stores $(1, 10)$ | | | -0.0131*** | | | | | |
| | | | (0.0034) | | | | | |
| # stores $(10, 15)$ | | | -0.0024 | | | | | |
| | | | (0.0079) | | | | | |
| # retail chains $(1, 5)$ | | | | -0.0296*** | | | | |
| | | | | (0.0099) | | | | |
| # retail chains $(1, 10)$ | | | | | -0.0281*** | | | |
| | | | | | (0.0075) | | | |
| # retail chains $(10, 15)$ | | | | | 0.0115 | | | |
| | | | | | (0.0199) | | | |
| Dist. to nearest store | | | | | | 0.0820** | | |
| | | | | | | (0.0318) | | |
| Cum. dist. to 3 nearest stores | | | | | | | 0.0287*** | |
| | | | | | | | (0.0080) | |
| Sum of selling areas ('000) $(1, 5)$ | | | | | | | | -0.0104*** |
| | | | | | | | | (0.0036) |
| Format fixed-effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.0202 | 0.1513 | 0.2392 | 0.1579 | 0.2306 | 0.1256 | 0.2059 | 0.1473 |

Notes: Regressions run with 61 observations. *, **, *** indicate significance at the 10%, 5%, 1% level, respectively. Source: Author's calculations.

area, I show that consumers may suffer locally from strong positions held by retailers. Thus, a quarter of the stores of the survey compete, at most, with three rivals within 5 km. This limited level of competition allows stores to inflate their margin by six percentage points on average, compared to the other stores. This inflation results in worse retail offer, which means higher prices and (potentially) lower assortments, lower store quality (e.g., freshness of products and cleanlineless) and fewer services offered.

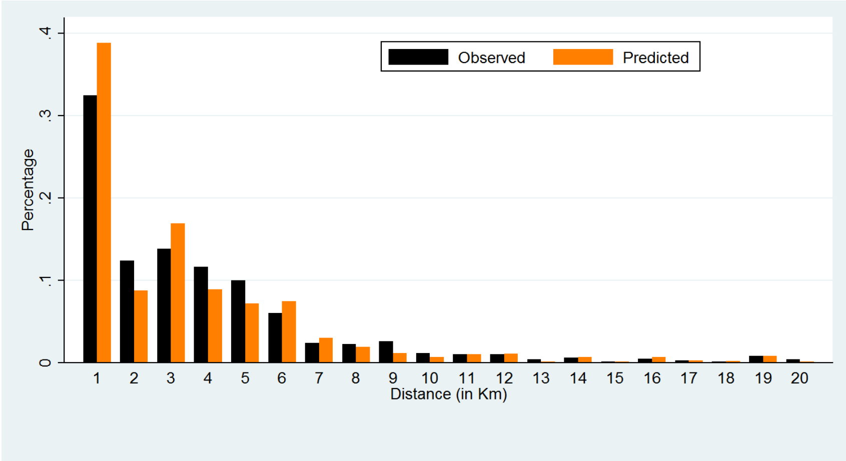
## 6.3 Robustness check

In this subsection, I confirm the robustness of the results by assessing the goodness of fit of the model and the sensibility of the results to several hypotheses. First, I note that the predicted market shares do not differ by more than 0.6 percentage points, on average, from the observed market shares (with the highest deviation equal to 4.8 percentage points). Furthermore, I ensure that the model realistically allocates consumers to stores. To that end, I plot in Figure 1 both the observed and predicted travel distances for each household. We observe that the shapes of the distributions are similar, which confirms the predictive power of the model. I also check that the model does not suffer from identification problems that may arise because of a poor approximation of the random coefficients distributions (Chiou and Walker [2007]). Thus, I re-estimate the model with different numbers of draws (200 and 1,000 Halton draws) and starting values to ensure that the coefficients are identified by checking their stability.
The margins reported in the previous section depend critically on the estimate of the price coefficient. In the following paragraphs, I examine how this coefficient varies depending on the different assumptions made.

**Measurement error in prices** The price variable suffers from a measurement error because some of the price indices correspond to estimated values. Because the price sample is clustered by retail chain and is limited in size, I cannot resort to a bootstrap method to replicate the predictions of the SUR model and fully address this issue. Instead, I merely examine the sensitivity of the results by drawing 100 observations from the 95% prediction interval of each predicted price index. For each replicated sample, I then re-estimate the mixed logit model and re-compute the stores' margin under the different pricing models. Overall, both the demand estimates and the estimated margins vary slightly. The standard deviation of the price coefficient is 0.1402 over the 100 replications. The stores' price-cost margin for pricing model 1 varies, on average, by 8.98% with regard to the values displayed in Table 8. Table 11 displays the original average margins by retail chain and the average margins calculated from the 100 replications.

**Household-specific price of the shopping basket** By incorporating between-household heterogeneity into the computation of the shopping basket's price, I recover a price coefficient that may differ from the one I would obtain with a less flexible model. To evaluate the effect of the first stage of the model on household price sensitivity, I re-estimate the model without this

Figure 1: Distance to chosen store (in km)




stage (i.e., by setting $Pr\left(I_{hc}\right) = 1$ in Eq. (4)). The results of this alternative specification are reported in the Online Appendix. One observes little difference in the demand estimates, except for a slight increase in the price coefficient that must be interpreted with caution given the automatic rise of the shopping basket's price (because now $Pr\left(I_{hc}\right) = 1$). Using these estimates, I derive higher price elasticities and lower margins while obtaining conclusions identical to those of the preferred pricing model. The introduction of a household-specific price to the shopping basket decreases the price disutility of the households. By weighting the prices of the product categories according to their probabilities of being purchased in a large grocery store by a given household, I simply attempt to reproduce more closely the household's shopping list. As a result, the households are more captive to the product categories that comprise their shopping basket in the two-step model rather than in the alternative specification. Consequently, ignoring the between-household heterogeneity in the composition of the shopping basket may underestimate a store's market power.

**Boundaries of choice sets** The manner in which I define the households' choice sets directly impacts the distribution of the prices faced by the households and can therefore significantly affect the price coefficient. To test this link, I re-estimate the model under alternative definitions of the households' choice set. The corresponding estimated price and distance coefficients are reported at the bottom of Table 7. Overall, because the price parameter varies between -1.7735 and -2.0272, there are small variations in the margins, whereas the mean of the distribution of the distance parameter varies between -2.1366 and -2.2142.

Table 11: Original and replicated margins for pricing model 1

| Retail chain | PCM (in %) | | Retail chain | PCM (in %) | |
|---|---|---|---|---|---|
| | Original | Replicated | | Original | Replicated |
| **Hypermarket** | | | **Hard discount** | | |
| Auchan | 10.16 | 11.07 | Aldi | 13.26 | 14.43 |
| Carrefour | 22.57 | 24.57 | Cdm | 11.24 | 12.24 |
| Géant Casino | 12.91 | 14.06 | Ed | 17.06 | 18.59 |
| Hyper U | 8.91 | 9.71 | Leader Price | 14.21 | 15.46 |
| Inno | 17.58 | 19.14 | Lidl | 12.26 | 13.36 |
| Intermarché | 14.53 | 15.87 | Norma | 13.62 | 14.85 |
| Leclerc | 14.46 | 15.74 | | | |
| **Supermarket** | | | **Convenience store** | | |
| Atac | 9.18 | 10.01 | Marché U | 13.24 | 14.46 |
| Casino | 9.64 | 10.54 | Shopi | 13.16 | 14.46 |
| Champion | 9.95 | 10.84 | | | |
| Intermarché | 16.45 | 17.91 | | | |
| Leclerc | 11.94 | 13.04 | | | |
| Monoprix | 10.64 | 11.57 | | | |
| Stoc | 13.28 | 14.46 | | | |
| Super U | 20.87 | 22.66 | | | |

Notes: The figures reported for the "replicated" column correspond to the average margins over the 100 replications.
Source: Author's calculations.

# 7 Counterfactual experiments

Many European countries are interested in enhancing competition in the grocery retail sector. Among the possible levers, both the UK and the French competition authorities have proposed amending the planning regimes of their respective countries to lower barriers to entry encountered locally and thereby promote the opening of new stores. Using the estimates obtained previously and the same Nash-Bertrand equilibrium assumption, I assess the benefits that one can expect from opening a hypermarket by quantifying the effects of a hypermarket's presence on the stores' prices and welfare. Assuming that the stores may react to these new market structures by changing their prices, I perform several counterfactual experiments by alternatively removing each hypermarket and computing the new market equilibrium for each scenario.
To measure the competitive effect of an existing hypermarket, I proceed as follows. I first eliminate the relevant alternative from the households' choice set. Then, I use the demand estimates and the marginal costs recovered under the preferred pricing model to numerically compute the new market equilibrium. Given equation 8, the predicted equilibrium prices $\mathbf{p}^*$ solve the system of $\mathcal{J}-1$ equations:

$$\mathbf{p}^* = \hat{\mathbf{c}} - \left[\mathbf{T}^* \otimes \boldsymbol{\Delta}\left(\mathbf{p}^*\right)\right]^{-1} \mathbf{s}\left(\mathbf{p}^*\right) \tag{14}$$

where $\hat{c}$ are the estimated marginal costs recovered from the store-level pricing model (model 1) and $T^*$ is the updated ownership matrix.[30] This system of $\mathcal{J}-1$ equations has a unique

[30]The prices of the outside option and the aforementioned outlier are assumed to remain constant. Moreover, I do not restrain hard discounters from setting uniform prices in their chain.

Table 12: Equilibrium outcomes following the removal of a hypermarket

| Store removing | Initial Mkt. Sh. | $\%\Delta wp$ | $\%\Delta p_3$ | $\%\Delta$ profit | Decomposition and total change of $\%\Delta$ CV | | |
|---|---|---|---|---|---|---|---|
| | | | | | Removal effect | Price effect | Total change |
| Auchan 1 | 3.72 | 0.17 | 0.63 | 1.43 | -0.18 | -0.17 | -0.35 |
| Carrefour 1 | 4.59 | -0.14 | 3.58 | -5.89 | -0.55 | 0.12 | -0.43 |
| Carrefour 2 | 10.86 | 0.27 | 0.90 | 1.31 | -0.67 | -0.23 | -0.90 |
| Carrefour 3 | 12.29 | -1.83 | -7.01 | -10.81 | -0.96 | 1.10 | 0.14 |
| Carrefour 4 | 5.28 | 0.10 | 0.88 | 0.83 | -0.19 | -0.05 | -0.24 |
| Geant Casino 1 | 3.29 | -0.04 | 0.30 | 0.23 | -0.15 | 0.00 | -0.15 |
| Geant Casino 2 | 2.86 | 0.07 | 0.06 | 0.48 | -0.27 | -0.04 | -0.31 |
| Hyper U 1 | 0.56 | 0.02 | 0.11 | 0.08 | -0.02 | -0.02 | -0.04 |
| Inno 1 | 2.55 | 0.05 | 3.01 | 0.15 | -0.20 | -0.08 | -0.28 |
| Intermarche 8 | 1.55 | 0.04 | 1.90 | -0.85 | -0.13 | -0.01 | -0.14 |
| Intermarche 13 | 0.37 | 0.03 | 0.09 | 0.13 | -0.02 | -0.02 | -0.04 |
| Leclerc 2 | 3.72 | 0.08 | 0.21 | 0.52 | -0.20 | -0.05 | -0.25 |

Note: The variations reported are computed with respect to the pre-removal equilibrium. ($\%wp$) corresponds to the store size-weighted average price change, $\%\Delta p_3$ corresponds to the average price change for the 3 nearest stores of the removed alternative and $\%\Delta$ profit is the total change in the sector's profit. The removal effect captures the "gross" effect on the consumer surplus from eliminating the alternative, whereas the price effect quantifies the impacts of the price variations on the consumer surplus. The last column ($\%\Delta$ CV) reports the total change in consumer surplus. Source: Author's calculations.

solution if the price variable $p^*$ corresponds to the price of the shopping basket and not to the prices of the product categories. It follows from this restriction that I cannot compute a household-specific price of the shopping basket, as I did previously. However, to approximate correctly the gain associated with the presence of an alternative, I drop the first stage of the demand model, which entails identical prices across the households. Based on the discussion in the previous section, this decision might soften the price changes in the simulations because the stores' margin are slightly underestimated. I then compare the consumer surplus, the sector's total profit and the equilibrium prices of the baseline situation with the results obtained from the removal experiments. In what follows, I retain the compensating variation as a measure of the change in the consumer welfare. Hence, assuming that the marginal utility of income for each household remains constant after the removal of an alternative, Small and Rosen [1981] have shown that the amount of money a household would need to be compensated can be derived from:

$$CV_h = \frac{\ln\left[\sum_{j=0}^{J_h} \exp\left(V_{hj}^{post}\right)\right] - \ln\left[\sum_{j=0}^{J_h} \exp\left(V_{hj}^{pre}\right)\right]}{|\alpha_h|} \tag{15}$$

where $V_{hj}^{post}$ and $V_{hj}^{pre}$ are defined by Eq. (3) using the pre- and post-removal predicted prices. Thus, the total change in consumer surplus is obtained by adding this measure over the households. These calculations assume that stores do not respond to the shock by changing their characteristics (observed and unobserved) and that the utility of the outside option remains unchanged.

Table 12 displays the outcomes of the new equilibrium for each hypermarket removal. Columns (5) and (8) give the total change in the sector's profit ($\%\Delta profit$) and the total change in consumer surplus ($\%\Delta CV$). As expected, for almost all cases, the simulations reveal a negative impact on the consumer surplus due to the elimination of a hypermarket. Depending on the hypermarket removed, the total change in consumer surplus varies between -0.90 and 0.14. In

other words, the total amount of money the households would need to be compensated at each period of purchase for the closing of a hypermarket equals the variation multiplied by the total consumer surplus derived from the baseline situation and expressed in monetary units. The magnitude of the changes in the total consumer welfare indicates that the consumers noticeably value the presence of an additional hypermarket. Moreover, one notes that the largest hypermarkets (in terms of market share) do not necessarily contribute the most to the consumer welfare. Along with the valuation of the store characteristics, the consumer surplus also accounts for the competitive pressure exerted by the hypermarket removed on the other stores.
Because the total change in consumer welfare is likely to mask some opposite effects, I decompose the total variation into two components. One component captures the "gross" effect of eliminating the alternative (i.e., without price adjustment (see column 6)), while the other component quantifies the price effect on the consumer welfare (column 7). This decomposition indicates that the magnitude of the effects differs significantly across the simulations for these two components and points out that the expected effect of an additional hypermarket on consumer welfare depends greatly on its characteristics (e.g., location, fascia, and store size). In addition, by decomposing the effect, one can explain why the households benefit from the removal of a hypermarket in one simulation because, in this particular case, one notes that the price effect more than offsets the removal effect. This finding suggests that this simulation generates an important price decrease. Finally, one observes in column (5) that eliminating an alternative benefits the industry more often, as the industry's profits rise significantly.
As indicated by the values of the price effect on consumer welfare, the changes in retail prices are moderated for a large portion of the simulations. Column (3) gives the store size-weighted average price change ($\%wp$). At most, I denote a price increase of 0.27% which is quite significant following a removal of a single hypermarket. However, I more frequently obtain an increase below $0.10\%$. These low variations in price may be explained by the traditional trade-off between the market share effect and the price competition effect faced by stores. To illustrate this trade-off, I report in column (4) the average price change for the 3 stores nearest to the closing hypermarket. The fact that the nearest rivals increase their prices more intensively than the mean of the set of stores indicates that the competitive effect of an additional hypermarket falls rapidly with distance. Hence, the stores located closer to the removed alternative can achieve greater profit by raising their prices than by cutting them to attract more customers. Conversely, the stores located farther away adopt the opposite strategy.

The general message from these experiments is that reducing the market concentration level by promoting the entry of a new competitor is almost always beneficial to consumers. However, to ensure a significant price decrease, it appears necessary to favor the entry of several stores and to pay attention to the product-mix that they offer. More broadly, these simulations suggest that French policymakers should more seriously consider softening the Raffarin Act to decrease the prices of foodstuffs in France.

# 8 Conclusion

This paper addresses the issue of the French grocery retail sector's competitive intensity. I develop and estimate a structural model of demand in which households have preferences over both store characteristics and geographic proximity. The methodology combines previous contributions of the literature on discrete choice models of demand among spatially differentiated firms and an original approach to determine the household-specific shopping basket. In addition, the paper extends the existing literature devoted to appraising retailers' market power by accounting for the store's ability to set its prices according to the local market structure. As shown by recent inquiries, "price flexing" is a key feature of the business strategy applied by French retail chains and consequently, an important dimension of a store's market power. Using the estimated parameters of the demand model, I recover the stores' price-cost margin under alternative pricing strategies, which differ by the degree of cooperation across stores while fitting into this framework. I then select the preferred pricing model by applying a non-nested testing procedure.

The model is estimated for a French metropolitan area using a cross-sectional household survey containing detailed information on the stores visited for the main food product categories. The results show that the stores set prices according to the most competitive scenario. This finding rules out any collusive behavior as a cause of local monopoly power. The average estimated level of profitability exhibits no signs of low degree competition in this market. However, a closer look at the results shows important differences among the stores. These differences indicate that a significant proportion of large grocery stores exert excessive market power. In practice, these stores take advantage of a weakly competitive environment to distort their offer and increase their margin. For instance, I find that the stores that compete with, at most, three rivals within 5 km set their margin six percentage points higher on average than other stores. Although the survey is based on data covering a single market area, I have good confidence that the findings can be extended to France because of its lower degree of concentration compared with other market areas (see Table 1).

Together, the results contribute to the debate on the level of competition in the French grocery retailing sector. I provide new empirical evidence on the existence of market areas with a low degree of competition based on the low density of the stores combined with a high disutility to travel expressed by the consumers in the area. The counterfactual experiments show that by promoting the entry of a new competitor, one significantly improves the consumer welfare and almost always decreases the price levels of foodstuffs. More generally, the simulations speak in favor of a relaxation of the regulation at the market entrance.

# References

ALLAIN, M.-L., AND C. CHAMBOLLE (2011): “Anti-Competitive Effects of resale-Below-Cost Laws,” *International Journal of Industrial Organization*, 29(4), 373–385.

ALLAIN, M.-L., C. CHAMBOLLE, AND T. VERGÉ (2008): *La loi Galland sur les relations commerciales. Jusqu'où la réformer ?* Opuscule du CEPREMAP, ed. de la rue d'ulm edn.

ASPLUND, M., AND R. FRIBERG (2002): “Food Prices and Market Structure in Sweden,” *Scandinavian Journal of Economics*, 104(4), 547–566.

BARROS, P. P., D. BRITO, AND D. DE LUCENA (2006): “Mergers in the food retailing sector: An empirical investigation,” *European Economic Review*, 50(2), 447–468.

BELL, D. R., AND J. M. LATTIN (1998): “Shopping Behavior and Consumer Preference for Store Price Format: Why "Large Basket" Shoppers Prefer EDLP,” *Marketing Science*, 17(1), 66–88.

BERRY, S. (1994): “Estimating Discrete-Choice Models of Product Differentiation,” *RAND Journal of Economics*, 25(2), 242–262.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63(4), 841–890.

BERTRAND, M., AND F. KRAMARZ (2002): “Does Entry Regulation Hinder Job Creation? Evidence From The French Retail Industry,” *Quarterly Journal of Economics*, 117(4), 1369–1413.

BHAT, C. R. (2001): “Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model,” *Transportation Research Part B: Methodological*, 35(7), 677–693.

BISCOURP, P., X. BOUTIN, AND T. VERGÉ (2008): “The Effects of Retail Regulations on Prices: Evidence from the Loi Galland,” mimeo 2008-02, Institut National de la Statistique et des Etudes Economiques, D3E.

BONNET, C., AND P. DUBOIS (2010): “Inference on Vertical Contracts Between Manufacturers and Retailers Allowing for Non Linear Pricing and Resale Price Maintenance,” *RAND Journal of Economics*, 41(1), 139–164.

BRIESCH, R. A., P. K. CHINTAGUNTA, AND E. J. FOX (2009): “How Does Assortment Affect Grocery Store Choice?,” *Journal of Marketing Research*, 46(2), 176–189.

CARLUER-LOSSOUARN, F. (2007): *L'Aventure des Premiers Supermarchés*. Linéaires - Editions du Boisbaudry.

CHIB, S., AND E. GREENBERG (1998): "Analysis of Multivariate Probit Models," *Biometrika*, 85(2), 347–361.

CHIOU, L. (2009): "Empirical Analysis of Competition Between Wal-Mart and Other Retail Channels," *Journal of Economics & Management Strategy*, 18(2), 285–322.

CHIOU, L., AND J. L. WALKER (2007): "Masking identification of discrete choice models under simulation methods," *Journal of Econometrics*, 141(2), 683–703.

COHEN, M. A., AND R. W. COTTERILL (2011): "Assessing the Impact of Retailer Store Brand Presence on Manufacturer Brands in an Equilibrium Framework," *Journal of Industrial Economics*, LIX(3), 372–395.

COMMISSION HAGELSTEEN (2008): *La négociabilité des tarifs et des conditions générales de vente*. La documentation Française, Paris.

COMPETITION AUTHORITY (1997): "Avis relatif à diverses questions portant sur la concentration de la distribution," Discussion Paper 97-A-04, Paris.

——— (2007): "Avis relatif à la législation relative à l'équipement commercial," Discussion Paper 07-A-12, Paris.

DAVIS, P. J. (2006): "Spatial Competition in Retail Markets: Movie Theaters," *RAND Journal of Economics*, 37(4), 964–982.

DUBOIS, P., AND S. JÓDAR-ROSELL (2010): "Price and Brand Competition Between Differentiated Retailers: A Structural Econometric Model," mimeo, Toulouse School of Economics.

ELLICKSON, P. B., AND S. MISRA (2008): "Supermarket Princing Strategies," *Marketing Science*, 27(5), 811–828.

HOTELLING, H. (1929): "Stability in Competition," *Economic Journal*, 39, 41–57.

HOUDE, J.-F. (2011): "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, forthcoming.

MANUSZAK, M. D. (2010): "Predicting the impact of upstream mergers on downstream markets with an application to the retail gasoline industry," *International Journal of Industrial Organization*, 28(1), 99–111.

MCMANUS, B. (2007): "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee," *RAND Journal of Economics*, 38(2), 512–532.

MURPHY, K. M., AND R. H. TOPEL (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 3(4), 88–97.

NEVO, A. (2001): "Measuring Market Power in the Ready-To-Eat Cereal Industry," *Econometrica*, 69(2), 307–342.

PETRIN, A., AND K. TRAIN (2010): "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47(1), 3–13.

PINKSE, J., M. E. SLADE, AND C. BRETT (2002): "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70(3), 1111–1153.

REY, P., AND T. VERGÉ (2010): "Resale Price Maintenance And Interlocking Relationships," *Journal of Industrial Economics*, 58(4), 928–961.

RICHARDS, T. J., AND S. F. HAMILTON (2006): "Rivalry in Price and Variety Among Supermarket Retailers," *American Journal of Agricultural Economics*, 88(3), 710–726.

RIVERS, D., AND Q. H. VUONG (2002): "Model selection tests for nonlinear dynamic models," *Econometrics Journal*, 5(1), 1–39.

SMALL, K. A., AND H. S. ROSEN (1981): "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 49(1), 105–130.

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

——— (2006): "Store Characteristics in Retail Oligopoly," *RAND Journal of Economics*, 37(2), 416–430.

SMITH, R. J. (1992): "Non-Nested Tests for Competing Models Estimated by Generalized Method of Moments," *Econometrica*, 60(4), 973–980.

THOMADSEN, R. (2005): "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," *RAND Journal of Economics*, 36(4), 908–929.

TRAIN, K. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press.

VILLAS-BOAS, S. (2007): "Vertical Relationships Between Manufacturers and Retailers: Inference with Limited Data," *Review of Economics Studies*, 74(2), 625–652.

VUONG, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57(2), 307–333.

# Appendix

## A Results from the MVP model

This section presents the results from the MVP model. Table 14 reports the estimates. The reference age group for the system of probit equations is age group 1. We note that the binary variables, Age group, and the variables describing the retail environment (i.e., Montpellier city and # hypermarket $\leq$ 10 km), determine a significant part of each household's retail channel choice. The households in which the household head's age corresponds to age groups 1 or 3 are more likely to visit a large grocery store than the other households. Additionally, living in a retail environment with numerous hypermarkets positively influences the choice to patronize a large grocery store. Conversely, the households living in Montpellier are more likely to perform their errands in other retail channels (e.g., specialized stores). This difference shows that people living in urban areas demonstrate different shopping patterns.

Overall, we note that the model performs well ($\chi^2_{(80)} = 1504$, *p*-value=0.0000). On average, the predicted probabilities of visiting a large grocery store conditional on a category do not differ by more than 0.82 percentage points from the actual probabilities. I perform a LR test for the joint restriction of all correlation coefficients equal to zero. The null hypothesis is rejected ($\chi^2_{(28)} = 16543$, *p*-value=0.0000), which supports the adoption of a MVP model vs. a system of simultaneous probit equations. Many of the correlation coefficients are significant and positive (see Table 13). This finding implies that factors other than the observed covariates are at work to explain multiple category purchases in the supermarket channel.

Table 13: Correlation coefficients from the MVP model

| | Fruits & vegetables | Meat | Cooked meat | Cheese | Other dairy product | Grocery item | Alcoholic drink | Soft drink |
|---|---|---|---|---|---|---|---|---|
| Fruits & vegetables | 1.0000 | | | | | | | |
| Meat | 0.7375*** | 1.0000 | | | | | | |
| Cooked meat | 0.6405*** | 0.9209*** | 1.0000 | | | | | |
| Cheese | 0.6388*** | 0.7462*** | 0.7278** | 1.0000 | | | | |
| Other dairy product | 0.6498*** | 0.7406*** | 0.7284 | 0.9081*** | 1.0000 | | | |
| Grocery item | 0.5548*** | 0.6409*** | 0.6516*** | 0.7175*** | 0.8577*** | 1.0000 | | |
| Alcoholic drink | 0.3216*** | 0.4138*** | 0.4572*** | 0.4624*** | 0.5899*** | 0.5698 | 1.0000 | |
| Soft drink | 0.4815*** | 0.5593*** | 0.5570 | 0.7316*** | 0.8553*** | 0.8044** | 0.6704*** | 1.0000 |

Notes: Standard errors in parentheses.***, **, *: significance at the 1%, 5%, 10% level respectively. Source: Author's calculations.

## B Selection tests of supply models

Given that the pricing models defined in subsection 4.3 are non-nested, I infer the model that fits the data best by applying a non-nested testing procedure. I follow Rivers and Vuong [2002]'s procedure, which proceeds via pairwise comparisons of the competing models.

For each competing model (denoted $h$), there is a corresponding pricing equation derived from the profit maximization problem and expressed as a function of the implied price-cost margin

Table 14: Results from the MVP model

| Variable | Fruits & vegetables | Meat | Cooked meat | Cheese | Other dairy product | Grocery item | Alcoholic drink | Soft drink |
|---|---|---|---|---|---|---|---|---|
| Constant | -0.0271 | 0.2780* | 0.2116 | 0.4170*** | 0.7426*** | 0.5684*** | -0.0232 | 0.7215*** |
| | (0.1466) | (0.1484) | (0.1478) | (0.1583) | (0.1712) | (0.1663) | (0.1523) | (0.1762) |
| Age group 2 | 0.0232 | -0.0998 | -0.0419 | 0.2378 | 0.2458 | 0.1757 | 0.2402* | 0.4428*** |
| | (0.1328) | (0.1354) | (0.1348) | (0.1500) | (0.1671) | (0.1563) | (0.1396) | (0.1765) |
| Age group 3 | -0.1948 | -0.1170 | -0.1126 | 0.2680** | 0.4562*** | 0.3429*** | 0.3187*** | 0.5260*** |
| | (0.1210) | (0.1228) | (0.1224) | (0.1342) | (0.1526) | (0.1425) | (0.1279) | (0.1562) |
| Age group 4 | -0.1188 | -0.1095 | -0.0647 | 0.2630** | 0.2104 | 0.1966 | 0.2481* | 0.3054** |
| | (0.1204) | (0.1231) | (0.1224) | (0.1342) | (0.1473) | (0.1392) | (0.1268) | (0.1494) |
| Age group 5 | -0.2191* | -0.3052*** | -0.2856** | -0.0197 | 0.2371 | 0.2453* | 0.2148 | 0.1930 |
| | (0.1260) | (0.1277) | (0.1272) | (0.1365) | (0.1535) | (0.1479) | (0.1320) | (0.1537) |
| Age group 6 | -0.4203*** | -0.4397*** | -0.4867*** | -0.1150 | -0.1218 | -0.0103 | -0.0935 | -0.0738 |
| | (0.1233) | (0.1245) | (0.1237) | (0.1329) | (0.1444) | (0.1407) | (0.1257) | (0.1458) |
| Montpellier city | -0.4159*** | -0.6468*** | -0.6546*** | -0.2422* | -0.0783 | -0.3731*** | -0.9305*** | -0.3785** |
| | (0.1200) | (0.1256) | (0.1262) | (0.1378) | (0.1519) | (0.1522) | (0.1440) | (0.1720) |
| # hypermarket $\leq$ 10 km | 0.0752*** | 0.1027*** | 0.0968*** | 0.0518*** | 0.0272 | 0.0612*** | 0.0998*** | 0.0547*** |
| | (0.0160) | (0.0166) | (0.0166) | (0.0180) | (0.0198) | (0.0199) | (0.0188) | (0.0225) |
| Card | -0.00573 | -0.0533 | -0.0638 | 0.1297 | 0.2110** | 0.1799* | 0.3112*** | 0.1838* |
| | (0.0855) | (0.0862) | (0.0861) | (0.0921) | (0.0983) | (0.0975) | (0.0875) | (0.1040) |
| House | -0.0086 | 0.0070 | 0.0226 | 0.0358 | 0.0010 | 0.0241 | 0.0182 | 0.0987 |
| | (0.0778) | (0.0783) | (0.0785) | (0.0863) | (0.0947) | (0.0928) | (0.0834) | (0.1009) |
| Work | -0.1282* | -0.1091 | -0.1096 | -0.0759 | -0.0919 | -0.1049 | 0.0750 | -0.1060 |
| | (0.0730) | (0.0739) | (0.0738) | (0.0823) | (0.0921) | (0.0880) | (0.0785) | (0.0971) |
| | | | | | | | | |
| Observations | 1611 | | | | | | | |
| Log-likelihood | -4,890.62 | | | | | | | |
| $\chi^2_{(80)}$ | 1504.5 | | | | | | | |

Notes: Standard errors in parentheses.*, **, *** indicate significance at the 10%, 5%, 1% level, respectively. The variable *# hypermarket $\leq$ 10 km* corresponds to the number of hypermarkets within 10 km of the household's residence, and *Work* is a binary variable that takes a value of 1 if the person who performs the errands in the household works. Source: Author's calculations.

and costs shifters:

$$p_j^h = \mu_j^h + \underbrace{\omega_j^h + W_j^h \lambda_h}_{\text{cost shifters}} + \eta_j^h$$

where $\mu_j^h$ corresponds to the price-cost margin of store $j$ under model $h$, while $\omega_j^h$ and $W_j^h$ are the unobservable and observable costs shifters, respectively. By subtracting the stores' price-cost margin from both sides of the equation, I can express the estimated marginal cost as follows:

$$c_j^h = \omega_j^h + W_j^h \lambda_h + \eta_j^h$$

The test principle consists in comparing the explanatory power (or, more precisely, the lack-of-fit criterion $Q_n^h$, see definition below) of the cost shifters on the marginal costs recovered from two competing models, $h$ and $h'$. Hence, the implied marginal costs of the preferred model are best explained by exogenous cost shifters. Identifying the cost parameters requires sufficient variation among the stores' marginal cost, even though part of this variation may be independent from a rival's behavior.
In what follows, I assume that the marginal costs may be expressed as an exponential function of the cost shifters:

$$c_j^h = \exp\left(\omega_r^h + W_j^h \lambda_h\right) + \eta_j^h$$

where the unobserved cost variables $\omega_r^h$ are supposedly identical among the stores of the same retail group.
The parameters of the cost equations are estimated using a nonlinear least squares (NLLS) estimator. Thus, the lack-of-fit criterion, $Q_n^h$, is defined as the objective function of the NLLS (i.e., the residual sum of squared).
The Rivers-Vuong test statistic, defined as $T_n = \frac{\sqrt{n}}{\widehat{\sigma}_n}\left(Q_n^h\left(\widehat{\theta}_n^h\right) - Q_n^{h'}\left(\widehat{\theta}_n^{h'}\right)\right)$, is expressed as a function of the difference in the lack-of-fit criteria between the two competing models and an estimate of the sampling variance between the objectives $(\sigma^2)$. The Rivers-Vuong test statistic is asymptotically distributed according to a standard normal distribution.
Based on Rivers and Vuong [2002], the null hypothesis is that the two competing models, $h$ and $h'$, are asymptotically equivalent when

$$H_0 : \lim_{n\to\infty}\left\{\overline{Q}_n^h\left(\overline{\theta}^h\right) - \overline{Q}_n^{h'}\left(\overline{\theta}^{h'}\right)\right\} = 0$$

where $\overline{Q}_n^h\left(\overline{\theta}^h\right)$ is the expectation of the lack-of-fit criterion $Q_n^h\left(\theta^h\right)$ evaluated at the pseudo-true values of the parameters of model $\overline{\theta}^h$ (similar to model $h'$). The first alternative hypothesis is that $h$ is asymptotically better than $h'$ when

$$H_1 : \lim_{n\to\infty}\left\{\overline{Q}_n^h\left(\overline{\theta}^h\right) - \overline{Q}_n^{h'}\left(\overline{\theta}^{h'}\right)\right\} < 0.$$

Table 15: NLLS estimates of the cost equations

| Dependent variable: Estimated marginal costs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Pricing model | | | | | | | |
| Cost variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Constant | 1.98*** | 2.01*** | 1.90*** | 1.78*** | 1.94*** | 1.67*** | 1.63*** | 3.26*** |
| | (0.07) | (0.07) | (0.09) | (0.14) | (0.08) | (0.22) | (0.26) | (0.65) |
| Gas station | 0.06 | 0.06 | -0.05 | 0.01 | 0.05 | 0.18 | 0.17 | 0.52 |
| | (0.05) | (0.05) | (0.07) | (0.09) | (0.06) | (0.14) | (0.17) | (0.48) |
| Mall | 0.30** | 0.33** | 0.32 | 0.22 | 0.26* | 0.23 | 0.24 | 0.07 |
| | (0.14) | (0.15) | (0.20) | (0.26) | (0.16) | (0.24) | (0.28) | (0.63) |
| # employees | -0.07*** | -0.12*** | -0.10*** | -0.09** | -0.06** | -0.08* | -0.07 | 0.02 |
| | (0.03) | (0.03) | (0.04) | (0.04) | (0.03) | (0.04) | (0.04) | (0.11) |
| # parking slots | -0.67** | -0.69** | -0.66 | -0.45 | -0.62* | -0.29 | -0.28 | 1.22 |
| | (0.33) | (0.33) | (0.44) | (0.59) | (0.36) | (0.59) | (0.68) | (1.81) |
| LR-test $\left\{\omega_r^h = 0\right\}$ | 56.95 | 28.29 | 28.22 | 13.16 | 9.94 | 5.70 | 4.85 | 5.24 |
| *p-value* | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |

Notes: Robust standard errors in parentheses.***, **, *: significance at the 1%, 5%, 10% level respectively. The number of retail facilities (Mall variable) and the number of employees (# employees variable) are divided by 100. Retail group fixed-effects $\omega_r^h$ are not reported. The omitted retail group is Leclerc. Source: Author's calculations.

Similarly, the second alternative hypothesis is that $h'$ is asymptotically better than $h$ when

$$H_2 : \lim_{n\to\infty} \left\{\overline{Q}_n^h\left(\overline{\theta}^h\right) - \overline{Q}_n^{h'}\left(\overline{\theta}^{h'}\right)\right\} > 0.$$

The Rivers and Vuong [2002] test is a generalization of the Vuong [1989] test to a broad class of estimation methods (e.g., NLLS or GMM). Compared with other non-nested model selection procedures, such as the Cox-type statistic developed by Smith [1992], one advantage of this test is that it addresses a mis-specification of the competing models. However, by definition, this test is non-transitive and therefore may fail to determine a unique preferred model.

NLLS estimates of the cost equation derived from the different pricing models are reported in Table 15. The results show that the stores' marginal cost rises significantly with the number of facilities in the shopping mall. By contrast, both the number of employees and the number of parking slots per square meter negatively influence the marginal cost of the stores. This finding indicates that these variables participate in the local monopoly power of these stores. Moreover, the unobserved cost shifters captured through the retail group fixed-effects appear to be highly significant, as suggested by the likelihood ratio test. This suggests that the buyer power of the purchasing central unit greatly influences the stores' marginal cost.

# C Adjusting the variance of the mixed logit estimator

Because I choose to estimate the demand model in two stages and because the second stage model (i.e., the mixed logit model) contains explanatory variables constructed from the probabilities estimated in the first stage, I have to adjust the variance-covariance matrix of the second stage model. More specifically, I have to account for the noise of the estimates from the first stage model while computing the standard errors of the parameters in the second stage model. From the formula given by Murphy and Topel [1985], I derive the correct asymptotic variance-covariance matrix. Let $\widehat{\theta}^{MVP}$ and $\widehat{V}^{MVP}$ be the estimated parameters and the estimated variance-covariance matrix of the multivariate probit model, respectively. Assume that $\widehat{\theta}^{MXL}$ and $\widehat{V}^{MXL}$ are the estimated parameters and the estimated variance-covariance matrix of the mixed logit model, respectively. According to Murphy and Topel [1985], the correct asymptotic variance-covariance matrix of the mixed logit model is given by:

$$\widehat{V}^{MXL} + \widehat{V}^{MXL}\left(\widehat{C}\widehat{V}^{MVP}\widehat{C}' - \widehat{R}\widehat{V}^{MVP}\widehat{C}' - \widehat{C}\widehat{V}^{MVP}\widehat{R}'\right)\widehat{V}^{MXL}$$

where

$$\widehat{C} = \left\{\sum_h \left(\frac{\delta \ln L_h^{MXL}}{\delta \widehat{\theta}^{MXL}}\right)\left(\frac{\delta \ln L_h^{MXL}}{\delta \widehat{\theta}^{MVP}}\right)\right\}$$

$$\widehat{R} = \left\{\sum_h \left(\frac{\delta \ln L_h^{MXL}}{\delta \widehat{\theta}^{MXL}}\right)\left(\frac{\delta \ln L_h^{MVP}}{\delta \widehat{\theta}^{MVP}}\right)\right\}$$

with $L_h^{MVP}$ and $L_h^{MXL}$ are household $h$'s contributions to the likelihood function of the MVP model and the MXL model, respectively.

Les **Working Papers SMART – LERECO** sont produits par l'UMR SMART et l'UR LERECO

- **UMR SMART**

L'Unité Mixte de Recherche (UMR 1302) *Structures et Marchés Agricoles, Ressources et Territoires* comprend l'unité de recherche d'Economie et Sociologie Rurales de l'INRA de Rennes et le département d'Economie Rurale et Gestion d'Agrocampus Ouest.

Adresse :
UMR SMART - INRA, 4 allée Bobierre, CS 61103, 35011 Rennes cedex
UMR SMART - Agrocampus, 65 rue de Saint Brieuc, CS 84215, 35042 Rennes cedex

- **LERECO**

Unité de Recherche *Laboratoire d'Etudes et de Recherches en Economie*
Adresse :
LERECO, INRA, Rue de la Géraudière, BP 71627 44316 Nantes Cedex 03

Site internet commun : http://www.rennes.inra.fr/smart/

**Liste complète des Working Papers SMART – LERECO :**
http://www.rennes.inra.fr/smart/Working-Papers-Smart-Lereco
http://econpapers.repec.org/paper/raewpaper/

The **Working Papers SMART – LERECO** are produced by UMR SMART and UR LERECO

- **UMR SMART**

The « Mixed Unit of Research » (UMR1302) *Structures and Markets in Agriculture, Resources and Territories*, is composed of the research unit of Rural Economics and Sociology of INRA Rennes and of the Department of Rural Economics and Management of Agrocampus Ouest.

Address:
UMR SMART - INRA, 4 allée Bobierre, CS 61103, 35011 Rennes cedex, France
UMR SMART - Agrocampus, 65 rue de Saint Brieuc, CS 84215, 35042 Rennes cedex, France

- **LERECO**

Research Unit *Economic Studies and Research Lab*
Address:
LERECO, INRA, Rue de la Géraudière, BP 71627 44316 Nantes Cedex 03, France

Common website: http://www.rennes.inra.fr/smart_eng

**Full list of the Working Papers SMART – LERECO:**
http://www.rennes.inra.fr/smart_eng/Working-Papers-Smart-Lereco
http://econpapers.repec.org/paper/raewpaper/

**Contact**

**Working Papers SMART – LERECO**
INRA, UMR SMART
4 allée Adolphe Bobierre, CS 61103
35011 Rennes cedex, France
**Email :** smart_lereco_wp@rennes.inra.fr

**2012**

**Working Papers SMART – LERECO**

UMR INRA-Agrocampus Ouest **SMART** (Structures et Marchés Agricoles, Ressources et Territoires)

UR INRA **LERECO** (Laboratoires d'Etudes et de Recherches en Economie)

Rennes, France