Competition between French grocery stores: Evidence from a price comparison website

Working paper

Etienne Chamayou*

CREST-LEI and Department of Economics, Ecole Polytechnique

December 11, 2017

Abstract:

The French grocery store chain Leclerc operates a price comparison website which allows to compare each of its stores with a selection of competitors, and performs chain comparisons at the national level. A large price cross section collected from the website, including over 4 million price records from nearly 2,300 supermarkets, is used to investigate price levels and dispersion across local markets in an unprecedented way. Prices are found to depend strongly on store chain affiliation, and are poorly explained by observable market characteristics, in particular one-size-fits-all measures of market power. One chain is observed to virtually operate a uniform pricing policy. Store price comparisons yield more volatile results when stores are separated by a higher distance, which supports a relation between price dispersion and consumer search costs. At the market level, the lowest and the highest prices of a given product are separated by an average 17% difference. Controlling for store fixed effects only reduces this gap to 10%. Price dispersion is loosely related to observable market characteristics, but increases significantly with an index of market prices. These findings are consistent with stores generally using noisier price strategies to extract higher profits when consumer loyalty or search costs allow them to do so.

Keywords:

JEL Classification Numbers: D83, L81, M31

*e-mail: *etienne.chamayou@ensae.fr*

1 Introduction

Since the development of supermarket chains in France, a regulation has been developed with a view to protect small retailers and suppliers from increasingly large retail chains. Empirical investigations performed by Bertrand and Kramarz (2002) and Biscourp et al. (2013) suggest that passed laws have had significant unexpected consequences, stressing the need to develop a better understanding of competition in the branch. Following Stigler (1961), many papers have documented the existence of price dispersion, contradicting the "law of one price", and stressed the impact of imperfect consumer information on competition. In that respect, the large number of available products within grocery stores leads to question to which extent consumers are actually able to compare prices and find the lowest price. To date, very few papers have investigated price dispersion in grocery markets, and all investigations have remained largely constrained by data limitations.

In this paper, the richness of the data allows to study competition in food retailing across local markets in an unprecedented way.[1]. The main data source is a price comparison website, www.quiestlemoinscher.com, which is operated since 2006 by one of the largest chains in France. The website was scraped once in March 2015 to extract all available price records, thus yielding a cross section of millions of product prices posted by nearly 2,300 supermarkets across France.

The first section provides an overview of the literature on price dispersion in the grocery market and gives some context about the French market. The second section describes the data collected from the price comparison website www.quiestlemoinscher.com and the methodology it uses to compare stores and national chains. The third section investigates how store chain affiliation, local competition and socio-demographic variables account for prices. Finally, the last section tests whether price dispersion appears to be driven by imperfect information and investigates the relation between dispersion and competition.

Prices are found to depend strongly on store chain affiliation, and are poorly explained by observable market characteristics, in particular one-size-fits-all measures of market power. One chain is observed to virtually operate a uniform pricing policy. Store price comparisons yield more volatile results when stores are separated by a higher distance, which supports the hypothesis that the presence of consumer search costs relaxes competition, allowing price dispersion to emerge. At the market level, the lowest and the highest prices of a given product are separated by an average 17% difference. Controlling for store fixed effects only reduces this gap to 10%. Price dispersion is loosely related to observable market characteristics, but increases significantly with an index of market prices. These findings are consistent with stores generally using noisier price strategies to extract higher profits when consumer loyalty or search costs allow them to do so.

[1] Lach (2002) and Zhao (2006) are strongly limited both in terms of store number and product sample. In Dubois and Perrone (2015), the size and structure of the consumer panel, meant to be representative of France as a whole, do not allow to measure dispersion within local markets.

2 Literature and context

2.1 Price dispersion in the grocery market

Since the seminal paper of Stigler (1961), a large literature has investigated the link between "consumer ignorance" and price dispersion, namely the persistence over time of different prices for a homogeneous good in a given market. Following Varian (1980), a rich theoretical paradigm has emerged in which price dispersion results from price randomization by sellers in equilibrium. Empirical research, on the other hand, has long been hampered by a scarcity of adequate data.

Lach (2002) studies price dispersion with data originally collected by the Israeli Central Bureau of Statistics to measure inflation. The necessity to have a significant number of price records of well identified products leads the paper to focus on four grocery store products sold in Israel over four years. Dynamic price dispersion is documented in the form of sellers frequently changing quartiles in the price distribution over months at the national level. Data do not allow to observe price dispersion within local markets

Zhao (2006) investigates the relation between price dispersion, measured through the coefficient of variation, with consumer search costs, competition intensity, and consumer heterogeneity. A positive correlation with dispersion is found for each of these elements. Data consist in a scanner panel covering 23 product categories of 6 supermarkets within a suburban area of Chicago from June 1991 to Junz 1993. The analysis focuses on the largest 10 brands in terms of market share within each product category. Product categories are narrow enough (e.g. Butter, Coffee) for products within categories to be considered as alternatives for a given consumer. Price dispersion related to product conditioning and brand diversity can then be quantified. The unit price of a given product (as defined by brand and quality) is found to generally decrease with size, while significant heterogeneity across intra-brand price dispersion is measured, meaning that brands exhibit different levels of differentiation. The paper uses the frequency of store visits and the frequency of product category or brand purchases as proxies for search costs. Intensity of competition varies due to a store entry in the market over the studied period. Higher industry concentration is expected and found to lead to lower price dispersion. Consumer heterogeneity is measured through the coefficients of variation in various consumer demographic variables. Lower variability is expected and empirically observed to involve a lower role of price discrimination hence lower dispersion.

Dubois and Perrone (2015) analyse price dispersion in the French supermarket industry with four product categories: beer, cola, coffe and whisky. Data come from a panel of households which were asked to register all their food purchases using a scanner between 1999 and 2001. They find that stores frequently move across quartiles of the product price distributions that they observe over time, and estimate a structural model which accommodates sequential search, vertical product differentiation and heterogeneous consumer tastes. They find that search cost are high and that the majority of consumers is thus poorly informed about prices in equilibrium. Price elasticities

differ significantly from the perfect information model.

2.2 The French Market

The regulation of the French grocery store market has motivated two papers which provide interesting insights regarding its evolution.

Bertrand and Kramarz (2002) analyse the impact of a restriction on large store openings introduced in 1974 to protect small retail stores. They find that a stronger deterrence of entry, decided by boards at the regional level, is associated with increased retailer concentration and weaker employment growth.

Biscourp et al. (2013) study the effects of the Loi Galland, passed in 1997, which modified existing below-cost pricing regulations with a view to protect small retailers and producers from larger retail chains. Existing regulation, dating back to 1963, had indeed proved to be inefficient given its loose definition of cost. The new law was thus meant to clarify the rules by defining the threshold as the invoice price. This forbade to take ex-post rebates into account in the final price. The paper documents a weakening of the relation between concentration and retail prices which is likely to reflect a reduction in intra-brand competition resulting from the Loi Galland. Indeed, the new regulation gives suppliers of branded products the possibility to impose industry-wide price floors (minimum RPM), while negotiating the actual wholesale price with retailers through rebates. As expected, the measured effect is stronger for branded product than for store brand products, which are less likely to have been by the change in the regulation. Price dispersion is found to be reduced for branded products and the price gap between most expensive and less expensive stores is reduced. Price dispersion is yet observed to remain significant.

In 2015, the French food retailing industry was dominated by six firms, which accounted for over 80% of total sales. Carrefour and Leclerc were the two largest groups with respective 22% and 20% market shares, followed by Intermarche (14%), Casino (12%), Auchan (11%) and Systeme U (10%)[2]. A remarkable difference between chains lies in their ownership structure. While Carrefour, Casino and Auchan generally own the large stores operated under their brands, Leclerc, Intermarche and Systeme U are cooperatives. The creation of the comparison website quiestlemoinscher.com (thereafter "Qlmc") is part of a long term strategy of the chain Leclerc to prove the competitiveness of its prices. Soon after the launch in May 2006, Carrefour filed a complaint about the lack of transparency and potential biases in comparisons. The website was forced to close by a court decision. An updated version of the website was released on November 2006 and has since then remained in operation. Legal proceedings continued until the rejection by the court of cassation of Carrefour's claims in January 2010. In 2015, Leclerc was still using the website on a frequent basis for advertising.

[2]Source: Kantar Worldpanel 2015.

3 Data and descriptive statistics

The following section provides an overview of the data collected from the website, and replicates the comparisons it performs while discussing the employed methodology. In March 2015, a script was used to extract all price records made available on the website. The collect was achieved by looping through all comparisons made between Leclerc stores and listed local competitors. This implies that the obtained database is a subset of the price cross section collected by the website, and that the price comparisons between national chains cannot be exactly replicated. Data collected from the website include the following variables: product name, section and family, store name (including chain, city and additional information if necessary to disambiguate stores), unit price and date of price record. These data were merged with a database of store characteristics including store gps coordinates, size, and municipality code. Store location and size were used to estimate local Herfindahl-Hirschman Indexes following standard competition authority practices. The municipality code allowed to add socio-demographic data describing local population size and revenue. Some sections of the paper use smaller cross sections of prices collected before and after March 2015 to discuss the robustness of results and add some insights about price dynamics.

3.1 Stores and competition

Until 2013, the website only offered comparisons between Leclerc and competitors at the chain level. For each competing chain, prices were collected at a sample of stores meant to be representative of the store network. Some constraints were thus imposed on store location and size, while exact store choice was claimed to be random. From 2013 on, the development of the "drive" concept in France has allowed the comparison website to cover far more stores, and thus to start displaying store level comparisons. The concept of "drive" implies that consumers are offered the opportunity to shop online from a physical store (at the same prices) and collect their purchases whenever it suits them. As a consequence, the collection of prices can be achieved on the internet, as opposed to costly physical store visits. As of March 2015, Qlmc claimed to cover 60% of the stores of the 10 supermarket chains compared (44% in August 2013).

Regarding store level comparisons, the website states that each Leclerc is compared with a selection of its most relevant competitors within 30 km, based on Leclerc managers' expertise. The website also indicates that stores the surface of which is smaller than 1,000 m^2 are excluded, as are stores belonging to chains which are deemed to be too differentiated such as hard discount chains. Finally, Leclerc stores are not included among potential competitors. A total number of 575 Leclerc stores were found to be listed on the website in March 2015. The comparison of each store with its respective selection of competitors yielded 2,390 pairs of stores, involving 1,815 non Leclerc stores. Data were missing for 14 Leclerc stores and 51 competitors. This implies that among competitors of the 561 Leclerc stores for which price data have been collected, 36 out of 1811 are missing ($\leq 2\%$).

Table 1: Representation of major national chains on Qlmc and in the data

	France	QLMC		Data	
	Nb stores	Nb stores	Coverage	Nb stores	Coverage
Auchan	142	125	88%	112	79%
Carrefour	222	188	85%	171	77%
Carrefour Market	925	421	46%	239	26%
Casino	392	151	39%	76	19%
Cora	58	58	100%	54	93%
Geant	108	108	100%	92	85%
Intermarche	1,770	1,022	58%	530	30%
Leclerc	579	579	100%	561	97%
Simply Market	305	50	16%	49	16%
Systeme U	1,030	632	61%	413	40%
Total	5,531	3,334	60%	2,297	42%

Table 1 provides an overview of stores covered by Qlmc and in the data as of March 2015 for the ten national chains compared on Qlmc. The first "Nb Stores" column indicates the total number of stores by retail chain in France according to LSA. The second one, under "QLMC", gives the number of stores for which Qlmc claims to have price records, and the last one, under "Data" show how many stores are covered in the data that were collected from the website. The "Coverage" columns are simply obtained by dividing the number of stores, respectively on Qlmc and in the data, by the actual total number of stores in France according to LSA. The coverage rates in the data are relatively high and rather close to Qlmc rates for chains which are characterized by large store surfaces: Auchan, Carrefour, Cora, Geant and Leclerc. This can be explained by the fact that Leclerc is present across all regions and operates rather large stores[3]. Regarding chains with smaller store formats, coverage is lower both for Qlmc and in the data with respect to the website (e.g. 19% for Casino in the data vs. 39% on Qlmc). Two natural explanations are the slower development of "drive" within smaller stores[4] and the fact that stores from these chains are less likely to be listed as relevant local competitors for Leclerc stores on Qlmc.

Table 2 provides an overview of competition according to Qlmc comparisons[5]. On average, a Leclerc store is compared with 5 competitors, and separated by 2.4 km or a 6.1 minute drive distance from its closest competitor. from its closest over 50% of all Leclerc supermarkets are compared with a store located within 2 km or a drive distance of 6 minutes (cf. Q50 of "closest" columns). Except for 28 stores, the furthest competitor is located within 30 km (respectively 29 stores with a 35

[3]Only stores which are listed on Qlmc as local competitors of Leclerc stores could be collected in our data.

[4]Collecting prices from an additional store which has a "drive" is virtually costless. Once the program has been written, it works with any store of the same chain.

[5]The website does not claim to be comprehensive.

Table 2: Overview of competition around the 575 Leclerc stores in Qlmc

	Nb competitors	Distance (km) mean	closest	median	furthest	Drive time (minutes) to mean	closest	median	furthest
Mean	5.0	8.8	2.4	8.5	15.9	13.5	6.1	13.4	21.0
Std	1.6	5.1	2.5	6.0	9.7	4.7	3.3	5.4	8.6
Min	1.0	0.8	0.1	0.5	0.9	3.5	0.0	1.8	4.0
Q10	3.0	3.0	0.7	2.5	4.6	8.4	2.8	7.5	11.8
Q25	4.0	4.8	1.1	3.7	8.4	10.2	4.0	9.6	15.1
Q50	5.0	7.8	1.8	6.5	15.3	12.8	5.7	12.4	19.6
Q75	6.0	12.3	2.7	12.5	21.5	16.0	7.4	16.7	25.6
Q90	7.0	15.7	4.7	18.0	26.3	19.6	9.5	21.1	30.9
Max	12.0	28.6	21.1	28.5	67.0	36.7	30.9	34.9	78.1

Distance (km) as the crow flies. Drive time (minutes) was obtained from Google.

minute maximum drive time). For 14 Leclerc stores, the closest listed store is over 10 km away (respectively 12 stores without competitor within 15 minutes). No store meets these two criteria, hence it does not seem that the lack or omission of nearby competitors led to include stores beyond reasonable distance. For instance, the Leclerc outlet which has the furthest competitor in the data (67 km) is listed with 7 competitors, of which 5 are located within 30 km.

Biscourp et al. (2013) define catchment areas by radiuses (i.e. distances as the crow flies), following Barros et al. (2006) and the method then employed by competition authorities. As they do not have exact store locations, they define catchment areas around city centers. Local market concentration is measured by the Herfindahl-Hirschman Index (HHI). Market shares are approximated by selling areas in the computations. Store turnover is indeed unknown, but expected to be strongly correlated with size (They also argue that a HHI based on size may be a better indicator in their specific case). Main estimations are performed with a 10 km distance. Allain et al. (2016), studying the impact of a large merger in the French market, also compute HHI based on store size and use radiuses of 10 and 20 km respectively for supermarkets and hypermarkets. They note than in the case of the merger they investigate, the French competition authority considered that consumers were willing to drive 15 to 30 minutes to reach a hypermarket, and 10 to 15 minutes to a smaller supermarket or discount store.

3.2 Products and comparison methodology

As of March 2015, only national brand products are covered by the website. Even though products are identified by the bar code on Qlmc to ensure precision of comparisons, in our data products are identified by their section, family and exact product name including format. Product families within each of the seven product section are detailed in Table 3. There are seven food

product sections: meat and fish, vegetables and fruits, bakery, fresh food, frozen food, savoury grocery, sweet grocery, baby food and drinks. Non food products are split in four sections: health and beauty, household, pets and home and textile. The methodology note on Qlmc indicates that for chain comparisons, the number of products covered in each family is determined by the volume of national hypermarket and supermarket sales, with a global objective of 3,000 products. Within each family, products are chosen based on the national hypermarket and supermarket detention rates. Products whose detention rate is below 30% (i.e. products referenced by less than 30% of the stores) are dropped. This led to a total of 2,461 national brand product references covered for March 2015 (2,510 in August 2013). As regards local competitor comparisons, all products for which price records are available at both stores are used.

Price records obtained from the website include all products used in each store level comparison. As a consequence, there are 12,318 product references in the data as of March 2015. Table 4 provides an overview of the relative weights of each section in terms of product number and value. Column "Nb %" is obtained by computing the number of product references within each section by the number of unique products in the data. Column "Value %" accounts for the sum of the average prices of each product reference within the section divided by the sum of the average prices of all product references. The "$\geq$ 500 obs.", respectively "700", columns show how the relative weights of each section vary if we drop product references for which less than 500, respectively 700, price records are available. The 700 observation threshold allows to roughly align the number of product references with the one used by Qlmc in national chain comparisons. We use these restrictions to perform robustness checks when we replicate national comparisons. The five largest sections, regardless of the criterion, are fresh products, health and beauty, savoury grocery, sweet grocery and drinks. Drinks and health and beauty products tend to have larger values than products from other categories, so that they account for a significantly higher share in terms of value than product count.

The comparison of Leclerc with its competitors follows two simple steps. First, the average price of each product is computed for each chain, provided the product is observed within enough stores of the chain. Leclerc is then successively compared to each of its competitors based on all products for which a chain price was computed. The result displayed on the website is the percentage difference between the price of the basket for the competing chain and for Leclerc:

$$\frac{\sum_i P_{iC} - \sum_i P_{iL}}{\sum_i P_{iL}}$$

where i refers to all products in the baskets, P_{iC} and P_{iL} respectively stand for the average price of product i for the competing chain (C) and for Leclerc (L). The comparison between two stores is similar except that it uses store prices instead of average chain prices.

Table 3: Product sections and families

Section	Families
Baby and dietetic food (573)	Baby food (418); Dietetic products (155)
Drinks (1,233)	Beer and Spirits (443); Fizzy drinks and Cola (244); Water (176); Juices and Smoothies (110); Squash and Cordial (101); Wine, Champagne and Cider (159)
Fresh products (2,595)	Butter and Cream (199); Meat (490); Cheese (491); Milk and eggs (150); Fish (98); Delicatessen (660); Yoghurts and Chilled Desserts (507)
Frozen food (368)	Ice cream and Frozen yoghurt (101); Frozen vegetables and fries (91); Frozen pizzas, pies and ready meals (128); Frozen Meat and Fish (48)
Health and Beauty (2,127)	Kitchen Roll and Tissues (86); Oral care (169); Feminine care and Baby changing (138); Drugstore (97); Haircare (558); Face and body skincare (951); Men toiletries (128)
Home and textile (308)	DIY and Car (9); Kitchen and dining room (50); Home Office (171); Batteries, lightbulbs and plugs (54)
Household (679)	Air fresheners and insect killers (118); Laundry (124); Cloths, Gloves and Scourers (45); Cleaning (225); Dishwashing (64); Specialist laundry and Washing machine cleaner (103)
Pets (239)	Cat and dog food (233); Litter (6)
Savoury grocery (2,032)	Snacks (214); Condiments and Spices (609); Canned goods (406); Precooked dishes (205); Pasta, Rice and Flour (328); Soups (270)
Sweet grocery (2,099)	Biscuits (294); Coffee and Tea (368); Chocolates ans sweets (450); Desserts, Sugar and Sweeteners (318); Breakfast (453); Cakes (215)
Vegetables and fruits (65)	Fruits (65)

Number of products within each section or family in parentheses.

Table 4: Number and cumulated value of products by section

	All products		$\geq$ 500 obs		$\geq$ 700 obs	
	Nb %	Value %	Nb %	Value %	Nb %	Value %
Baby and dietetic food	4.7	4.3	3.9	3.0	3.3	2.4
Drinks	10.0	15.3	10.9	20.4	11.1	21.9
Fresh products	21.1	15.5	19.8	16.7	18.4	15.2
Frozen food	3.0	3.1	3.0	3.9	2.4	3.1
Health and beauty	17.3	26.9	11.5	12.8	12.4	13.4
Home and textile	2.5	3.4	0.5	0.7	0.3	0.4
Household	5.5	6.8	5.5	6.8	5.8	7.2
Pets	1.9	2.8	3.0	4.4	3.0	4.5
Savoury grocery	16.5	9.4	19.6	12.5	20.4	12.6
Sweet grocery	17.0	12.3	22.1	18.8	22.8	19.2
Vegetables and fruits	0.5	0.4	0.2	0.2	0.2	0.2
Total	100.0	100.0	100.0	100.0	100.0	100.0
Total Nb or Value (euros)	12,318	43,883	3,467	9,138	2,578	6,682

3.3 Price comparison results

Results of chain level comparisons performed according to the website methology are reported in Table 5. The first two columns under "Nb stores" respectively indicate the number of stores used by Qlmc and the number of stores actually present in our data. The next two columns under "Nb products" similarly display the respective numbers of product references used in the comparisons. The "Qlmc" column under "Comparison vs. Leclerc" displays the comparison result as it was observed on Qlmc, while the "Data" column on its right corresponds to its replication attempt on available data. Finally, the last two columns provide comparison outcomes which are obtained when the comparison is biased against Leclerc, by dropping from the comparison the 10% or 20% products which are the most favorable to Leclerc. Despite the fact that data collected differ from these used by Qlmc, results are very similar and are found to be relatively robust to variations in product basket. Geant Casino is the second cheapest chain as of March 2015, only 1.5% more expensive than Leclerc (1.8% according to Qlmc). Dropping the 20% products which weigh in most favorably for Leclerc reduces the difference to 0.4%.

Results of store level comparisons performed according to the website methology are reported in Table 6. The "Nb pairs" column indicates the total number of comparisons performed between Leclerc stores and competitors from a given chain. The next columns provide a description of the distribution of all comparison outcomes. For instance, there were 99 comparisons involving a Leclerc and a Geant Casino supermarket as of March 2015. On average, the Geant Casino is found to be 1.8% more expensive than its Leclerc competitor. In one comparison, a Geant Casino

Table 5: Comparisons at the chain level

	Nb stores		Nb products		Comparison vs. Leclerc			
	Qlmc	Data	Qlmc	Data	Qlmc	Data	Bias 10%	Bias 20%
Auchan	125	112	1,976	2,382	+7.6%	+6.5%	+5.5%	+5.0%
Carrefour	188	171	1,294	1,284	+7.8%	+8.2%	+7.0%	+6.0%
Carrefour market	421	239	2,032	3,401	+13.5%	+12.4%	+11.6%	+10.2%
Casino	151	76	na	1,650	+16.7%	+16.8%	+15.8%	+15.4%
Cora	58	54	1,326	2,994	+10.2%	+9.4%	+8.3%	+7.3%
Geant Casino	108	92	1,582	1,582	+1.8%	+1.5%	+0.7%	+0.4%
Intermarche	1,022	530	1,971	6,287	+7.0%	+7.1%	+5.8%	+5.0%
Simply market	50	49	na	1,070	+12.9%	+13.4%	+11.6%	+11.2%
Systeme U	632	413	2,386	4,565	+6.7%	+5.8%	+4.8%	+4.7%

Comparisons are based on 561 Leclerc stores (vs. 581 in Qlmc). In the column "Bias 10%", the 10% products which compare most favorably for Leclerc in terms of percent price difference are dropped.

supermarket is found to be 0.6% cheaper than its Leclerc competitor. Except for Geant Casino, all chains have at least one store which is largely more expensive ($\geq 15\%$) than its Leclerc competitor, while Leclerc is never observed to compare too badly ($\geq -5\%$). Though there is heterogeneity across pairs, chain level comparisons appear to provide relatively meaningful information.

Table 6: Comparisons between Leclerc stores and their competitors by chain

	Nb pairs	Comparison of Leclerc stores vs. chain competitors Mean	Std	Min	Q25	Q50	Q75	Max
Auchan	118	+6.5%	3.3%	+1.6%	+4.1%	+5.7%	+8.3%	+19.5%
Carrefour	175	+8.2%	5.2%	-3.5%	+5.8%	+8.1%	+9.4%	+36.2%
Carrefour market	235	+13.8%	3.3%	+1.3%	+11.7%	+13.5%	+15.8%	+24.5%
Casino	57	+17.9%	4.8%	+0.5%	+16.8%	+18.7%	+21.0%	+27.5%
Cora	57	+8.6%	2.4%	+3.6%	+6.7%	+8.4%	+10.3%	+15.6%
Geant Casino	99	+1.8%	1.5%	-0.6%	+0.7%	+1.3%	+2.3%	+5.3%
Intermarche	525	+7.1%	2.8%	+2.0%	+5.4%	+6.6%	+8.2%	+28.4%
Simply market	49	+13.4%	6.2%	+6.5%	+9.8%	+10.6%	+15.4%	+31.8%
Systeme U	355	+6.7%	4.0%	+1.1%	+3.8%	+5.8%	+8.7%	+26.0%

Pairs were kept only when at least 400 products were available for comparison. There are 118 comparisons between a Leclerc store and an Auchan store. On average, an Auchan store is 6.5% more expensive than its Leclerc competitor.

3.4 Comparison dynamics

In March 2015, the comparison website allowed to download files containing price records used to perform comparisons between 2007 and 2012, and in May 2014. These files were aggregated in a database[6] which was used to compute evolutions of chain prices between each period. This allows to understand the variations in price comparison results displayed by the website over time. A price comparison (following the website methodology) is performed with all products of one chain for which an average price can be computed in two successive periods. Variations can then be chained to obtain statistics over longer periods. Indeed, product turnover generally does not allow meaningful direct comparisons between non successive price records. Table 6 provides an overview of the evolution in chain prices between 2007 and 2015. Base 100 is Leclerc in March 2015. Leclerc price indices were computed by comparing Leclerc prices between successive available price records. Competing chain indices were computed by comparison with Leclerc prices within each period.

Leclerc prices between May 2007 and May 2012 have increased by 1.13% (average annual increase of 0.25%). Until May 2011, other chain display similarly low variations. This translates in a relative status quo in chain comparison results. Geant Casino is then the most expensive chain relative to Leclerc (from +6% to +10%), followed by Cora (+5%). Auchan, Carrefour, Geant Casino, Intermarche and Systeme U display rather similar price levels (+3% to +4%). After May 2011, most chains exhibit a progressive loss in competitivess as compared to Leclerc. Geant Casino, however, constitutes a remarkable exception. After a peak in September 2012 (13.8% more expensive than Leclerc), the chain becomes increasingly price competitive from May 2013 on. As of March 2015,

[6] As mentioned in section 2.1, data prior to 2013 are very limited in terms of store and product coverage. As a consequence, they are simply used to give some context and perform robustness checks.

Table 7: Chain price indices from 2007 to 2015 (base 100: Leclerc in March 2015)

Date	Auchan	Carrefour	Cora	Geant	Intermarche	Leclerc	Systeme U
05/2007	113	116	118	117	115	110	115
04/2008	117	119	121	na	117	113	118
04/2009	116	114	118	123	116	112	116
04/2010	116	116	120	122	116	112	116
05/2011	119	117	118	121	115	112	116
06/2012	116	116	122	125	118	111	116
05/2014	111	109	120	110	118	105	114
03/2015	106	108	109	102	107	100	105

Base 100: Leclerc in March 2015. Leclerc price indices were computed by comparing Leclerc prices between successive available price records. Competing chain indices were computed by comparison with Leclerc prices within each period.

Geant Casino is the closest competitor of Leclerc in terms of price level (+1.3% vs. Leclerc), while it was actually the most expensive chain at the beginning of the period, and was still 12.2% more expensive than Leclerc as of March 2013. The history of comparisons also reveals that Carrefour, after a progressive increase in price competitiveness in the second half of 2013 and the first half of 2014 (+2.6% vs. Leclerc in September 2014), catches up abruptly with other comparable chains (Auchan, Intermarche and Systeme U) in March 2015 which are between 6% and 7% more expensive than Leclerc.

Intra-chain comparisons between May 2014 and March 2015 suggest that the relative loss of price competitiveness exhibited by Carrefour actually results from a mild change in prices by Carrefour (-1.4%) constrasting with significant cuts implemented by other chains (e.g. -4.3% for Auchan, -5.1% for Leclerc, -5.2% for Intermarche). Geant Casino achieves its unprecedented level of price competitiveness through an 8.5% decrease.

Overall, the history of comparisons reveals that beyond some stability at both extremities of the price ranking (Cora is persistently relatively expensive while Leclerc is always the cheapest chain), one chain, Geant Casino, radically changes its pricing policy in less than a year, and the ranking between the remaining national chains (Auchan, Carrefour, Intermarche and Systeme U) exhibits significant volatility over time[7].

[7]Our observations are consistent with price indices published by a blog dedicated to food retailing, www.olivierdauvers.fr, in association with a business intelligence firm (A3 Distrib, purchased by Nielsen in 2016) which collects and compares prices from drive websites.

4 Price determinants

Since its creation in 2007, Qlmc prominently displays aggregate comparisons with its major national competitors. On the one hand, such information may be considered relevant by consumers willing to shop based on rules of thumb, comparisons may largely reflect heterogeneity in store and market characteristics. This section investigates potential determinants of supermarket price heterogeneity, among which an approximation of the HHI which is commonly used in food retailing, and discuss the importance of chain pricing strategies.

4.1 Store price determinants

In order to study the relation between store prices and their potential determinants, we start by aggregating price information at the store level. Denoting P_{ij} the price of a product i observed at store j, Product_i a dummy variable which takes value 1 for all price records of product i and Store_j a dummy variable equal to 1 for all prices observed at store i, the coefficients of the following regression are estimated:

$$\log P_{ij} = \alpha_i \text{Product}_i + \beta_j \text{Store}_j + \epsilon_{ij} \tag{1}$$

Residuals ϵ_{ij} can be interpreted as the percentage deviation of a store product price from its expected geometric mean. The average of the residuals for each store (respectively product) is approximately null. The store coefficients β_j are used to compute store prices indexes which can be directly compared to previous indexes computed at the chain level. Formally, $(\beta_j + 1) * 100$ yields a price index for store j with base 100 for the store used as a reference store in the estimation. Distributions of price indexes by chain are reported in Table 8.

These indexes are used to investigate the extent to which the heterogeneity in store price indexes may reflect store characteristics, as well as socio-economic parameters and differences in competition intensity. Denoting $\text{Store price index}_i$ the price index of store i and $\text{Store characteristics}_i$ a vector of variables accounting for store characteristics[8], we estimate the coefficients of the following equation:

$$\text{Store price index}_i = \mu + \beta \text{ Store characteristics}_i + \epsilon_i \tag{2}$$

Store characteristics include store size, dummy variables corresponding to chain affiliation, and proxies to account for the intensity of competition, the potential store demand and the revenue of the local population. Following the literature, we use HHI to account for the intensity of competition.

[8] Accordingly, β is a vector of coefficients.

Table 8: Distribution of store price indexes by chain

	Nb Stores	Store price indexes Avg	Std.	Min	Q25	Q50	Q75	Max
Auchan	112	107	3.7	102	104	106	109	120
Carrefour	165	108	4.7	95	106	109	110	128
Carrefour market	85	113	2.2	108	112	113	114	118
Casino	23	115	5.4	100	115	116	118	123
Cora	54	110	2.0	104	109	110	111	116
Geant Casino	88	101	1.6	100	100	100	100	105
Intermarche	178	107	2.1	102	106	107	108	114
Leclerc	510	100	1.3	93	100	100	101	106
Simply market	9	112	3.0	108	109	111	115	116
Systeme U	203	105	3.6	99	103	104	108	117
All	1 427	105	5.0	93	101	103	109	128

Base 100: Leclerc in Limoges (reference store in the estimation of Equation 1)

The HHI variable is built is by considering each store as the center of a market, and weighting each store surface by the distance to the center of the market. Formally, for a given store i, denoting distance_{ij} the distance from store i to any store j, we weight the surface of store j by $e^{-\text{distance}_{ij}/10}$. This means that the surface of the store at the center of the market is weighted by 1, while distances of 2, 10 and 20 km respectively imply approximate weights of 0.8, 0.4 and 0.1. A similar method is employed to create a variable which aims at capturing the potential demand of each store. Considering successively all municipalities, we attribute a share of their population to each store depending on the store surfaces weighted by their distance to each municipality center. For each store, we then sum the population that can be met in each municipality to build an index of potential store demand. Robustness checks are performed with non weighted measures of HHI and demand, based on simple radiuses of 10 km for supermarkets and 25 km for hypermarkets. Population revenue is the median household revenue taken at the municipality level. Results are reported in Table 9.

Store and market characteristics are found to account for a small share of the variance in store indexes. In particular, Leclerc does not appear to be significantly less price competitive relative to competitors once the size of its stores and their location is taken into account. Chain affiliation appears to be a strong determinant of store price level, which is consistent with the relative stability in store level comparisons previously observed and previous studies on retail chain prices. Hosken et al. (2008) and Chamayou (2016) obtain similar results with gas stations respectively in the US and in France (even though gas station chains do not follow uniform pricing policies). Turolla (2016) investigates the impact of concentration in the region of Montpellier, in Southern France,

Table 9: Regressions of store price indexes

	(0)	(1)	(2)
Intercept	100.26***	101.97***	105.99***
	(0.12)	(0.68)	(1.26)
Auchan	6.70***	6.99***	
	(0.29)	(0.30)	
Carrefour Market	12.75***	12.34***	
	(0.33)	(0.31)	
Carrefour	7.93***	8.13***	
	(0.25)	(0.26)	
Casino	14.94***	14.01***	
	(0.59)	(0.55)	
Cora	9.61***	9.78***	
	(0.40)	(0.40)	
Geant Casino	0.71**	0.90***	
	(0.32)	(0.31)	
Intermarche	6.76***	6.43***	
	(0.24)	(0.23)	
Simply Market	11.26***	10.11***	
	(0.94)	(0.87)	
Systeme U	5.07***	5.38***	
	(0.23)	(0.22)	
HHI		0.00	0.01
		(0.02)	(0.04)
Surface		-0.36***	-0.58***
		(0.06)	(0.11)
Population revenue (th. euros)		0.02	0.03
		(0.02)	(0.04)
Population size (th. inhab)		0.06**	0.14***
		(0.02)	(0.03)
R^2	0.70	0.75	0.03
N	1,426	1,426	1,426

Columns (1) and (2) include dummies to control for regional specificities, the coefficients of which are not reported. Standard errors in parentheses. * p<.1, ** p<.05, ***p<.01.

using prices of 91 products collected from 27 stores. Using a mixed logit demand model and refined measures of competition intensity, the paper finds the market to be generally competitive, but also notes that a significant number of stores enjoy a large market power which allows them to achieve higher relative margins. While our findings suggest that a unique definition of market concentration at the national level is unlikely to constitute an efficient screening tool to detect insufficient competition, they should thus not lead to rule out the impact of local market power. Rather, they suggest that a finer approach is required.

4.2 Chain pricing policies

We further investigate the predictive power of chain affiliation on prices by focusing on product price distributions within chains. Even though large French food retailers generally do not implement uniform national pricing policies, empirical investigations reveal various degrees of uniformity at the chain level. Table 10 details the frequency of the mode (most common price) of each product within each supermarket chain listed on the price comparison website. Geant Casino stands out in terms of product price homogeneity. On average, a product is sold at the very same price in 89% of the chain stores. This implies that a random basket of goods has a relatively high probability to have the very same price in two Geant Casino stores, even if both are located far apart from each each other. The closest followers are Systeme U and Leclerc, for which the mode accounts for 39% and 38% of price observations on average.

Intra-brand price heterogeneity can also investigated from a store prospect. Table 11 accounts for the percentage of products carried by each store the price of which is found to be equal to the mode of the observed chain prices. The average Geant Casino store appears to follow a standard chain price for approximately 80% of its products. The median is 94% while is the min is 6% hence it appears that a limiter number of stores depart significantly from standard prices while price uniformity is the rule for the bulk of the store chains. Leclerc also exhibits a relatively strong concentration at the store level.

From a methodological point a view, it must be noted that the maximum values observed at the store level must be interpreted with caution. Absent standard national product prices, product price modes typically result from a few stores setting the same prices. The analysis can be refined by discarding price modes which are not followed by a large enough proportion of all chain stores. Robustness checks performed with thresholds of 33% and 50% confirm that Geant Casino and Leclerc stand out in terms of price concentration.

This analysis was replicated for each period of available price records. Results are similar across periods except for Geant Casino. In June 2012, the last observed period preceding its sharp increase in price competitiveness, the average product price mode accounts for 32% of observations. This is to be compared with 82% in May 2014. The increase in price competitiveness has thus been

Table 10: Distribution of the frequency of the mode (most common price) per product

	Nb	Mean	Std	Min	Q25	Q50	Q75	Max
Auchan	416	19	11	5	12	16	22	63
Carrefour	319	29	17	7	17	23	36	87
Carrefour Market	777	33	19	11	20	26	42	100
Geant Casino	417	89	10	45	83	91	97	100
Casino	157	37	15	6	29	33	44	86
Cora	364	20	11	6	14	17	23	90
Intermarche	1,326	25	19	5	13	18	29	97
Leclerc	1,788	38	23	3	14	38	59	95
Super U	1,077	39	12	9	32	37	44	91

On average, 38% of all Leclerc stores set the very same price for a given product.

Table 11: Distribution of the frequencies of "standard" prices per store

	Nb	Mean	Std	Min	Q25	Q50	Q75	Max
Auchan	107	14	7	2	9	13	18	37
Carrefour	146	28	15	0	19	28	36	67
Carrefour Market	223	32	16	0	19	32	45	60
Geant Casino	91	81	23	6	71	94	96	98
Casino	74	16	11	2	7	13	27	49
Cora	54	13	8	1	6	14	18	29
Intermarche	513	24	11	0	15	24	32	50
Leclerc	552	44	18	4	31	47	58	80
Super U	409	35	37	0	6	11	83	98

On average, the prices of 44% of the products carried by a Leclerc store are equal to the most common prices observed at Leclerc stores.

accompanied by a large price uniformization. Such a shock, having apparently affected a large number of markets across France in an essentially undifferentiated way, opens interesting research prospects. With quantity data, it would allow an approach similar to Allain et al. (2016) which combines a standard econometric analysis (differences in differences) with a structural approach, contributing to address the criticisms levelled by Angrist and Pischke (2010) against the empirical Industrial Organization literature[9].

Finally, observations on price uniformity are to be analysed in the light of the theoretical literature on uniform pricing. Dobson and Waterson (2008), referring to UK grocery retailing, show that uniform pricing can be used to soften price competition across markets. Allain et al. (2016) investigate the consequences of spatial discrimination and uniform pricing strategies on mergers. They stress how when one retailer implements uniform pricing, the anticompetitive effects of a merger typically affect consumers in markets which are not directly affected by the merger.

5 Price Dispersion

We now turn to the measure and analysis of price dispersion in the French food retailing industry. From a consumer viewpoint, this addresses the broad question of the validity of aggregate comparisons, at the store or chain level. From a research prospect, price dispersion has been noted to reflect imperfect information from consumers about prices, hence an important deviation from perfect competition. We first examine competition between pairs of rival stores to look for evidence of dispersion at the local level and evaluate the role of consumer information as a determinant of dispersion. We then quantify and investigate price dispersion at the national and local market levels.

5.1 Price dispersion and consumer information

We first measure price dispersion between pairs of competitors, following an approach introduced in Chandra and Tappata (2011) which aims at testing the relation between consumer information and price dispersion. Pairs of competitors which are separated by a very low distance are expected to compete fiercely, so that they constitute a population in which the "law of one price" is more likely to hold. On the other hand, a larger distance is expected to be associated with more limited consumer information. Models of search, often inducing mixed strategy equilibria, may then be more adequate to model competition[10]. In the single product case, following Varian (1980), mixed strategy equilibria have been given a dynamic interpretation, corresponding to the changes in ranks

[9] Angrist and Pischke (2010) criticize the overwhelming use of structural approaches as they generally require strong hypotheses. They call for more evidence relying on "simple, transparent empirical methods that trace a shorter route from facts to findings".

[10] Cf. Baye et al. (2006) for a survey.

that can be observed among sellers over time[11]. In the multi-product case, McAfee (1995) have shown that sellers can randomize margins on each product, either simply replicating the single product case of Varian (1980), or in a way that involves a correlation between a seller's various product prices. Chandra and Tappata (2011), with gasoline, measures rank reversals as the number of days during which the generally cheapest gas station is found to be more expensive. In this paper, rank reversals are measured in one period over products. Formally, considering the prices p_{il} and p_{jl} of two supermarkets i and j over $l \in L$ products, the rank reversals statistics between store i and j writes:

$$r_{ij} = \min \left\{ \frac{1}{L} \sum_{t=1}^{L} \mathbb{1}_{p_{il} > p_{jl}}, \frac{1}{L} \sum_{t=1}^{L} \mathbb{1}_{p_{jl} > p_{il}} \right\} \tag{3}$$

If one store is always more expensive than the other, or both always set the same price, rank reversals are equal to 0. Rank reversals can reach a maximum value of 50% when half of the products are strictly cheaper at store i while the other half are strictly cheaper at store j. Importantly, differentiation between stores tends to mechanically decrease rank reversals, hence it must be taken into account when comparing rank reversals across pairs of competitors. Table 12 provides an overview of rank reversals of all comparisons between chains previously found to operate at relatively similar price levels. The Leclerc vs. Geant Casino confrontation is the most stable across competitor pairs, and within pairs across products. Over 215 pairs of competing stores, Geant Casino is +1.4% more expensive on average, and Leclerc is less expensive in 85% of the store confrontations. On average, regardless of the affiliation of the cheapest store in the Leclerc vs. Geant Casino confrontation, the most expensive store is cheaper on 20.4% of the products available at both stores.

In addition to static dispersion, we use a second cross-section of prices collected in May 2014[12] to achieve a measure of dynamic price dispersion. More precisely, we look for evidence of changes in price rankings at the store pair level, namely when a store becomes strictly cheaper than a competitor in March 2015 while it was strictly more expensive in May 2014, and at the product level, that is looking at the share of products for which the price rank has reversed between the two periods. Descriptives statics are reported in Table 13. Among 114 store comparisons involving a Leclerc and a Geant Casino, 4.4% are won by a different store in the two periods. On average, 21.2% of products taken into account in the comparison changed order between the two periods i.e were strictly cheaper at Leclerc in first period and became strictly cheaper at Geant Casino in

[11] The absence of a pure strategy equilibra results from a tension between an incentive to undercut competitors' prices to attract perfectly informed consumers, and the possibility to extract a rent from uninformed consumers who are willing to accept any offer below their reservation price.

[12] Data for May 2014 were obtained after the first version of the paper was written and are less comprehensive than those of March 2015.

Table 12: Static store level comparisons

Chain A	Chain B	Nb pairs	Pairs (%) B vs. A avg	Pairs (%) A cheaper	Product avg Same price	Product avg Rank reversals
Leclerc	Geant Casino	215	+1.4	85	15.8	20.4
Leclerc	Carrefour	555	+9.1	98	6.4	14.7
Geant Casino	Carrefour	89	+7.6	99	4.1	25.1
Carrefour	Auchan	191	-0.3	52	9.4	28.9
Carrefour	Intermarche	365	-1.0	39	3.0	34.0
Carrefour	Systeme U	196	+2.6	61	4.1	27.3
Auchan	Intermarche	212	+0.8	62	3.0	32.9
Auchan	Systeme U	145	+3.1	66	4.3	27.0
Intermarche	Systeme U	490	+1.0	51	7.3	25.3

Among 215 pairs of Leclerc and Geant Casino competitors, Geant Casino is +1.4% more expensive on average, and Leclerc is less expensive in 85% of the pairs. Regardless of whether Leclerc or Geant Casino wins the overall comparison, on average, the loser i.e. most expensive store is strictly cheaper on 20.4% of products.

second period or the reverse.

Importantly, store differentiation leads to mechanically record relatively low rank reversals[13]. This issue is addressed by imposing a restriction on aggregate price differences and by running quantile regressions, as in Chandra and Tappata (2011). In order to test the link between distance, taken as a proxy for consumer search cost, and rank reversals, we denote Nearby_{ij} a dummy which takes value 1 when supermarkets i and j are separated by a relatively short distance and X_{ij} a vector of controls which account for their market characteristics[14]. We then run the following regression:

$$\text{r}_{ij} = \mu + \alpha \text{ Nearby}_{ij} + \beta X_{ij} + \epsilon_{ij} \tag{4}$$

In a first specification, distance as the crow flies is used, with a threshold of 5 km for the dummy variable Nearby. All pairs separated by less than 10 km are included in the regression. The second specification uses distances in minutes computed by Google, including all pairs for which the driving distance is below 20 minutes. The definition of the variable Nearby is based on a 12 minute threshold, which is found to be roughly equivalent to a 5 km distance in the data, namely when running a simple regression of driving distance on distance as the crow flies.

Rank reversals are found to be significantly less frequent for pairs which are separated by a short distance. Being separated by less than 5 km is associated with reductions of 5.31 and 4.71 points in rank reversals respectively in period 0 and 1 according to the OLS regressions. The same conclusion is reached with dynamics rank reversals between period 0 and period 1, with rank

[13] Cf. Wildenbeest (2011) for a model of search and price dispersion with vertical differentiation.
[14] Accordingly, β is a vector of coefficients

Table 13: Dynamic store level comparisons: March 2015 vs. May 2014

Chain A	Chain B	Nb pairs	Rank reversals	
			Pairs	Product avg
Leclerc	Geant Casino	114	4.4	21.2
Leclerc	Carrefour	152	5.9	24.6
Geant Casino	Carrefour	46	71.7	42.5
Carrefour	Auchan	49	42.9	38.0
Carrefour	Intermarche	119	53.8	38.6
Carrefour	Systeme U	102	48.0	37.2
Auchan	Intermarche	86	22.1	32.4
Auchan	Systeme U	101	34.7	29.9
Intermarche	Systeme U	322	32.8	30.5

Among 114 store comparisons involving a Leclerc and a Geant Casino, 4.4% are won by a different store in the two periods (draws can be neglected as they virtually never happen). On average, 21.2% of products taken into account in the comparison changed order between the two periods i.e were strictly cheaper at Leclerc in first period and became strictly cheaper at Geant Casino in second period or the reverse.

Table 14: Regressions of product price dispersion measured at the national level

Rank reversals	Nearby definition	Regression			
		OLS	Q25	Q50	Q75
May 2014	Distance	-5.31***	-8.00***	-4.19***	-4.25***
		(1.03)	(1.99)	(1.44)	(1.22)
	Time	-5.38***	-8.18***	-6.88***	-4.79***
		(1.19)	(1.77)	(1.98)	(1.49)
March 2015	Distance	-4.71***	-7.06***	-5.63***	-1.97
		(1.04)	(1.58)	(1.45)	(1.36)
	Time	-5.85***	-7.94***	-7.11***	-1.61
		(1.07)	(1.54)	(1.38)	(1.42)
Intertemporal	Distance	-4.62***	-6.48***	-4.19***	-2.20*
		(0.90)	(1.08)	(0.98)	(1.16)
	Time	-4.73***	-6.69***	-3.90***	-2.91**
		(0.97)	(1.11)	(1.08)	(1.28)

Standard errors in parentheses. Significance thresholds: * p<.1, ** p<.05, ***p<.01.

reversals being 4.62 and 4.73 point lower respectively with distance in km and time. Estimates for the Nearby dummies tend to be smaller or non significant for the Q75 quartile in the last column, which indicates that distance is less relevant for pairs of competitors which exhibit high rank reversals. This does not contradict the hypothesis of a link between consumer information and price dispersion. From a theory viewpoint, if consumer search cost prevent the existence of pure strategy equilibria, dispersion arises, hence rank reversals, but not with a frequency that depends on consumer information. Results from quantile regressions are thus consistent with the hypothesis that virtually all pairs exhibiting high rank reversals are good candidates for theoretical explanations involving mixed strategy equilibria.

5.2 National price dispersion

Product price dispersion is measured at the national level, both with raw prices and with residuals prices obtained from regression (1). Descriptive statistics are provided in Table 15. The "Nb Prod." columns provides the number of product references for which the price dispersion could be computed within each category. Product references which did not meet a threshold of 100 available price records were dropped. Column "Mean" under "Raw prices" reports the mean product price within each category. Overall, price dispersion could be computed for 6,935 product references, the mean price of which was 3 euros. The third column under "Raw prices" indicates the average percentage price difference between the 50% prices in the middle of the price distribution. For instance, for a baby food product, the average ratio of the third over the first price quartile is 9.1%, and 21.2% for the average ratio of 95th percentile over the 5th percentile. These ratios respectively decrease to 4.5% and 12.9% with price residuals (in the last two columns). Overall, the magnitude of price dispersion, regardless of its proxy, is relatively similar across product sections. Price dispersion is substantially reduced once store fixed effects are controlled for, but remains significant.

Denoting Product dispersion$_i$ a measure of price dispersion for product i, Price$_i$ the average price of product i over all stores for which a price record is available, and Section$_{ij}$ a dummy variable which takes value 1 if product i belongs to section j, we run the following regression:

$$\text{Product dispersion}_i = \mu + \alpha \text{ Price}_i + \beta_j \text{Section}_{ij} + \epsilon_i \tag{5}$$

Different measures of price dispersion are used depending whether the regression is performed with raw prices or price residuals. Results are reported in Table 16. The two first columns, which were obtained with raw prices, emphasize the link between dispersion and product value. Price dispersion measured by standard deviation can indeed largely be explained by product value. The coefficient of variation essentially cancels this effect out, with some overshoot as α becomes significantly negative.

Table 15: National price dispersion by product section

	Nb	Raw prices				Residuals		
	Prod.	Mean	CV	$\frac{Q75}{Q25}-1$	$\frac{Q95}{Q5}-1$	Std	Q75-Q25	Q95-Q5
Baby food	307	2.6 (2.4)	6.8 (2.2)	9.1 (4.9)	21.2 (7.8)	4.4 (1.6)	4.5 (2.6)	12.9 (5.2)
Pets	185	4.7 (3.2)	5.6 (1.7)	7.6 (3.4)	17.1 (5.6)	3.7 (1.1)	3.9 (1.6)	11.0 (3.5)
Drinks	688	5.2 (5.8)	5.9 (2.2)	7.6 (4.6)	17.9 (8.1)	4.4 (1.5)	4.6 (2.3)	12.9 (5.1)
Savoury grocery	1 358	1.9 (1.0)	6.7 (2.3)	8.4 (5.1)	21.5 (8.3)	4.8 (1.7)	5.1 (2.9)	14.0 (5.6)
Sweet grocery	1 380	2.4 (1.2)	7.0 (2.7)	9.3 (6.9)	22.1 (9.4)	5.0 (2.1)	5.5 (4.1)	14.4 (6.2)
Fresh	1 423	2.4 (1.1)	6.5 (2.1)	7.9 (5.1)	20.5 (8.1)	5.2 (1.6)	5.4 (2.9)	15.1 (5.5)
Health/Beauty	993	3.9 (2.4)	7.0 (2.2)	9.1 (4.5)	23.0 (8.8)	5.1 (1.7)	5.5 (2.8)	15.2 (5.9)
Household	403	3.9 (2.7)	6.9 (2.1)	8.8 (5.0)	22.0 (7.5)	5.1 (1.5)	5.4 (2.4)	14.7 (5.1)
Frozen food	198	3.4 (1.6)	6.9 (2.4)	8.8 (5.6)	22.3 (7.9)	5.2 (1.5)	5.6 (2.7)	15.7 (5.3)
All sections	6 935	3.0 (2.6)	6.6 (2.3)	8.5 (5.4)	21.2 (8.6)	4.9 (1.8)	5.2 (3.0)	14.3 (5.7)

Standard error in parentheses.
Column "Mean" under "Raw prices" is the mean price in euros across markets and products. All columns to its right are measures of dispersion to be read as percentages. The coefficient of variation ("CV") was indeed multiplied by 100, as were all variables describing quartile comparisons.

Product section coefficients capture minor differences. A similar result is obtained by considering the relative differences between the third and the first quartiles of the price distribution. The last two columns report results obtained with price residuals. By construction, differences in product prices are cancelled out in regression (1). Results differ slightly regarding product section coefficients but still explain a very small share of heterogeneity across products. Similar results are obtained when estimations are performed with product families. Zhao (2006) finds that price dispersion is correlated with product characteristics. For instance, the purchase frequency of a product is negatively correlates with dispersion, corresponding to the intuition that consumers a likely to be better informed on the prices of the products they buy more frequently. Unfortunately, we lack precise product characteristics to perform such an analysis. We note that product section, family or price have a limited predictive power regarding product market dispersion. This implies that our aggregate measures and analyses of price dispersion are unlikely to be strongly dependent on the sample of products.

Table 16: Regressions of product price dispersion measured at the national level

Prices	Raw	Raw	Raw	Res.	Res.
Dispersion measure	Std	CV	Q75/Q25-1	Std	Q75-Q25
Intercept	0.05***	7.25***	9.80***	4.60***	4.63***
	(0.00)	(0.13)	(0.31)	(0.10)	(0.18)
Price	0.05***	-0.18***	-0.28***	-0.08***	-0.06***
	(0.00)	(0.01)	(0.03)	(0.01)	(0.02)
Section Pets	-0.01*	-0.84***	-0.87*	-0.53***	-0.47*
	(0.01)	(0.21)	(0.50)	(0.16)	(0.28)
Section Drinks	-0.03***	-0.47***	-0.79**	0.19	0.32
	(0.01)	(0.16)	(0.37)	(0.12)	(0.21)
Section Savoury grocery	-0.01*	-0.23	-0.89***	0.36***	0.59***
	(0.01)	(0.14)	(0.34)	(0.11)	(0.19)
Section Sweet grocery	0.01	0.14	0.20	0.54***	1.05***
	(0.01)	(0.14)	(0.34)	(0.11)	(0.19)
Section Fresh food	-0.00	-0.37***	-1.27***	0.76***	0.88***
	(0.01)	(0.14)	(0.34)	(0.11)	(0.19)
Section Health and Beauty	0.05***	0.43***	0.37	0.80***	1.08***
	(0.01)	(0.15)	(0.35)	(0.11)	(0.20)
Section Household	0.03***	0.32*	0.06	0.75***	1.03***
	(0.01)	(0.17)	(0.41)	(0.13)	(0.23)
Section Frozen food	0.02***	0.30	-0.04	0.89***	1.17***
	(0.01)	(0.21)	(0.49)	(0.16)	(0.28)
R^2	0.70	0.06	0.03	0.04	0.02
N	6935	6935	6935	6935	6935

Reference product section is Baby food.
Standard errors in parentheses. Significance thresholds: * p<.1, ** p<.05, ***p<.01.

5.3 Market price dispersion

We now turn to the measure of price dispersion at the market level. Markets are defined according to the comparisons made available on Qlmc, namely around each Leclerc store. All products for which prices are available at all retailers in the market are taken into account in the analysis. We drop markets for which we are not able to compute the dispersion of at least 100 products. Measures of price dispersion are computed both with raw prices and with price residuals obtained from regression (1). Figures obtained with raw prices are likely to overestimate consumer search related price dispersion since price comparison results suggest that persistent price differences are non negligible. The method used to compute price residuals implies that the expected value of a large enough basket should be similar for each store in the market. Descriptive statistics are provided in Table 17 for each product section. The second column indicates the number of observation, where each observation corresponds to the dispersion of one product in a given local market. The average product has a coefficient of variation of 6.4% and a range of 17.1%, which roughly means that the highest price of a product is typically around 17% higher than the lowest price in the market. With residual prices, this gap drops to 10.5%. Measures of price dispersion do not exhibit significant variations across product sections[15]. Finally, the coefficients of variation, with raw prices, and the standard deviations, with price residuals, are hardly smaller than the ones obtained with price distributions at the national level.

We investigate how market price dispersion relates to market characteristics, in particular the intensity of competition. As previous results suggest that competition is imperfectly captured by available proxies, we introduce an index of market price among explanatory variables. Our hypothesis is that the presence of higher dispersion may reflect poorer consumer information, and thus be associated with higher prices.

Denoting Product market dispersion$_{ij}$ a measure of price dispersion for product i in a local market j, Product$_i$ a dummy variable which takes value 1 for all local measures of price dispersion corresponding to product i, and Market characteristics$_{ij}$ a vector which accounts for various market characteristics[16], we run the following regression:

$$\text{Product market dispersion}_{ij} = \mu + \alpha_i \text{ Product}_i + \beta \text{ Market characteristics}_{ij} + \epsilon_{ij} \qquad (6)$$

Results are reported in Table 18. Price dispersion is computed with price residuals, successively by the standard deviation in prices and the range. Estimations in the third and fourth columns differ only through the introduction of the variable meant to account for the market price level. The latter is built by computing the mean of the ratios of each store price index to their average

[15] The same can be observed with product families (available upon author request).
[16] Accordingly, β is a vector of coefficients.

Table 17: Market price dispersion by product section

		Raw prices		Residuals	
	Count	CV	Range	Std	Range
Baby food	2 798	7.2 (4.6)	19.4 (13.2)	4.3 (3.4)	10.3 (8.1)
Pets	2 858	6.1 (3.6)	15.8 (10.3)	3.9 (2.6)	9.5 (6.7)
Drinks	16 759	5.7 (3.9)	15.1 (11.3)	4.1 (2.8)	10.1 (7.0)
Savoury grocery	28 348	6.4 (4.1)	17.0 (11.6)	4.0 (2.8)	9.8 (6.8)
Sweet grocery	29 332	6.8 (4.7)	18.1 (13.4)	4.3 (3.5)	10.5 (8.3)
Fresh	24 889	6.3 (4.3)	16.8 (12.2)	4.5 (3.2)	10.9 (7.8)
Health and Beauty	15 148	6.9 (4.4)	18.3 (12.5)	4.6 (3.0)	11.4 (7.6)
Household	6 840	6.4 (4.3)	16.9 (12.1)	4.6 (3.0)	11.3 (7.6)
Frozen food	2 258	6.7 (4.5)	17.2 (12.0)	4.6 (3.3)	11.0 (7.9)
All sections	129 230	6.4 (4.3)	17.1 (12.3)	4.3 (3.1)	10.5 (7.6)

Only products with 100 observations or more are included.

chain index[17].

While the HHI, the population size and the population revenue are all estimated to be significantly correlated with price dispersion, their impact remains relatively small compared to the relation with market price. For the latter, an increase of one standard deviation, namely 1.63 points, implies an increase in the range of approximately 1.8 points. This is consistent with firms using noisier price strategies to achieve higher prices and margins across markets where relaxed competition and consumer information allow them to do so.

[17] Alternatively, we considered a simple average of store price indexes and obtained similar results.

Table 18: Regressions of market dispersion

	Std	Range	Std	Range
Intercept	3.56***	4.41***	-38.83***	-106.60***
	(0.83)	(2.19)	(3.60)	(9.67)
Market price index			0.43***	1.11***
			(0.03)	(0.08)
HHI	-4.83***	-13.52***	-2.67***	-7.86***
	(1.73)	(4.66)	(1.34)	(3.65)
Population size (th. inhab.)	0.01*	0.04**	0.01*	0.04***
	(0.01)	(0.02)	(0.01)	(0.02)
Population revenue (th. euros)	0.14***	0.35***	0.10**	0.23**
	(0.04)	(0.10)	(0.04)	(0.10)
Nb stores	-0.03	0.74***	0.04	0.92***
	(0.05)	(0.13)	(0.04)	(0.10)
Loc - City center	-0.38**	-1.08**	-0.27**	-0.79**
	(0.16)	(0.41)	(0.13)	(0.34)
Loc - Isolated	-0.61***	-1.51***	-0.40**	-0.97**
	(0.18)	(0.47)	(0.15)	(0.40)
Loc - Rural	-0.16	-0.63	-0.45**	-1.29***
	(0.37)	(0.90)	(0.22)	(0.52)
R^2	0.30	0.29	0.36	0.35
N	47 113	47 113	47 113	47 113

Standard errors (clustered at the product and market level) in parentheses.
Product fixed effects α_i are omitted in results.
Significance thresholds: * p<.1, ** p<.05, ***p<.01.

6 Conclusion

Using a large cross section of French supermarket product prices, we do not find a significant relation between variables accounting for store market power and prices. This suggests that one-size-fits-all measures of competition intensity such as the commonly used HHI are not good screening tools when it comes to detecting markets characterized by insufficient competition. On the other hand, empirical investigations support the hypothesis that consumer search costs soften competition, thereby allowing stores to set higher prices. Comparisons between pairs of store indeed reveal that product price rankings are more volatile when stores are separated by a higher distance, and measures of price dispersion at the market level are strongly correlated with market price levels. Finally, we observe that the chain affiliation largely determines store prices, and that there exists a large heterogeneity in chain pricing policies. These findings call for measures favoring price transparency, and further empirical investigations aimed at achieving a better understanding of competition at the local level. In this regard, the major change of pricing strategy implemented by the chain Geant Casino in 2013 within its supermarkets all over France provides an interesting research opportunity.

References

Allain, M.-L., C. Chambolle, and S. Turolla (2016). Politique tarifaire locale ou nationale : Quel impact pour le contrôle des concentrations dans le secteur de la distribution. *Revue Economique 63*, 53–67.

Allain, M.-L., C. Chambolle, S. Turolla, and S. Villas-Boas (2016). Retail mergers and food prices. *Journal of Industrial Economics (Forthcoming)*.

Angrist, J. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Barros, P., D. Brito, and D. De Lucena (2006). Mergers in the food retailing sector: An empirical investigation. *European Economic Review 50*, 447–468.

Baye, M. R., J. Morgan, and P. Scholten (2006). Information, search, and price dispersion.

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.

Biscourp, P., X. Boutin, and T. Verge (2013). The effects of retail regulations on prices: Evidence from the loi galland. *The Economic Journal 123*(573), 1279–1312.

Chamayou, E. (2016). Price dispersion in the french retail gasoline market. *Working paper*.

Chandra, A. and M. Tappata (2011). Consumer search and dynamic price dispersion: an application to gasoline markets. *RAND Journal of Economics 42*(4), 681–704.

Dobson, P. W. and M. Waterson (2008). Chain-store competition: Customized vs. uniform pricing. *International Journal of Industrial Organization 26*(6), 1425–1436.

Dubois, P. and H. Perrone (2015). Price dispersion and informational frictions: Evidence from supermarket purchases. *Working paper*.

Hosken, D. S., R. S. McMillan, and C. T. Taylor (2008). Retail gasoline pricing: What do we know? *International Journal of Industrial Organization 26*(6), 1425–1436.

Lach, S. (2002). Existence and persistence of price dispersion: An empirical analysis. *The Review of Economics and Statistics 84*(3), 433–444.

McAfee, R. P. (1995). Multiproduct equilibrium price dispersion. *Journal of Economic Theory 67*(1), 83–105.

Stigler, G. (1961). The economics of information. *The Journal of Political Economy 69*(3), 213–25.

Turolla, S. (2016). Spatial competition in the french supermarket industry. *Annals of Economics and Statistics (Forthcoming)*.

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

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

Zhao, Y. (2006). Price dispersion in the grocery market. *Journal of Business 79*(3), 1175–1192.