

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

Working paper

E. Chamayou*
CREST-LEI

October 6, 2016

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 price comparisons at the national level. This paper uses price data collected from the comparison website to investigate static and dynamic price dispersion across French grocery stores. Although chains are found to largely determine store pricing policies, chain level comparisons are often of little informative value given heterogeneity observed at the store level. The comparison appears to exhibit more dispersion when stores are separated by a higher distance, which supports the idea that dispersion is related to consumer information. At the market level, price dispersion is found to increase with price levels and to decrease with HHI.

Keywords:

JEL Classification Numbers: D83, L81, M31

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

1 Introduction

Since the development of supermarket chains in France, several regulations have been implemented with a view to protect smaller retailers and producers from increasingly large retail chains. The academic literature has yet painted an unflattering portrait of passed laws.

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.

Since the seminal paper of Stigler (1961), a large literature has investigated the link between "consumer ignorance" and price dispersion i.e. 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 relevant panel data.

Lach (2002) studies price dispersion with CPI data of 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 are yet not rich enough 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 June 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 size 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) with data spanning 1999-2001. They find that stores frequently move across quartiles of the product price distributions that they observe over time, and estimate a structural model which accomodates 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.

Using rich price data collected from a comparison website, this paper investigates static and dynamic price dispersion in the French supermarket industry. The first part of the paper provides a description of the data and reviews the methodology used by the comparison website from which they are extracted. The second part discusses the relevance of chain comparisons and investigate price dispersion at the store and market level. Chain affiliation is found to be a strong determinant of prices, while market characteristics account for a small share of price variations. Two chains are observed to be significantly cheaper than others as of March 2015. They are also characterized by high degrees of price uniformity compared to other chains. Whenever the average price difference across products between two competing stores is small enough, the result of the price comparison largely depends on product choice. (Todo: finish result summary + take aways)

2 Price comparison on quiestlemoinscher.com

As of 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%)¹. A remarkable difference difference between chains lies in their ownership structure. While Carrefour, Casino and Auchan own most of the stores operated under their brands, Leclerc, Intermarche and Systeme U are essentially franchise networks. The creation of the comparison website

¹Source: Kantar Worldpanel 2015.

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. The following section provides an overview of the methodology of the comparison website, two crucial aspects of which are competitor and product choice. Data collected from the website are used to replicate its comparison results. The website was explored in March 2015, with a view to extract its price data. The only feasible solution appeared to successively crawl comparisons between Leclerc stores and each of their competitors. This implies that obtained data differ from the price database used by Qlmc to establish comparisons performed at the national level with competing supermarket chains. Data collected from Qlmc include the following variables: product name (including format), section, family, store name (including chain, city and additional information if necessary to disambiguate), unit price and date of price record. These data were merged with a database of store characteristics including store surface.

2.1 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. The collection of prices can then be achieved efficiently 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 whose surface is smaller than $1,000\text{ m}^2$ are excluded, as well as stores belonging to chains which are deemed to be too differentiated (e.g. discounters). 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 unique Leclerc competitors (the number of pairs is larger since a store can be listed as a competitor of several

Leclerc stores). Data were missing for 14 Leclerc stores and 51 Leclerc 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

	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%

Data about store chains were provided by LSA, which is the source used by Qlmc.

Table 1 provides an overview of store coverage for the ten national chains compared on Qlmc in March 2015. Coverage is high and rather close to coverage in the full Qlmc sample 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. Regarding chains with smaller store formats, coverage is lower both for Qlmc and in the data with respect to the website (e.g. 22% for Casino vs. 39% claimed by the website). Two natural explanations are the fact that stores from these chains are generally less likely to be considered relevant local competitors for Leclerc stores, and the slower development of "drive" within smaller stores (which make price collection less costly).

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

	Nb competitors	Distance (km) to			
		closest	furthest	mean	median
Mean	5.0	2.4	15.9	8.8	8.5
Std	1.6	2.5	9.7	5.1	6.0
Min	1.0	0.1	0.9	0.8	0.5
Q10	3.0	0.7	4.6	3.0	2.5
Q25	4.0	1.1	8.4	4.8	3.7
Q50	5.0	1.8	15.3	7.8	6.5
Q75	6.0	2.7	21.5	12.3	12.5
Q90	7.0	4.7	26.3	15.7	18.0
Max	12.0	21.1	67.0	28.6	28.5

Table 2 provides an overview of competition according to qlmc comparisons (the website does not claim to be comprehensive). On average, a Leclerc store is compared with 5 competitors, and over 50% over Leclerc stores are compared with a store located within 2 km. The furthest competitor is generally within 30 km (15 km for almost half of them), except for 28 stores. For 14 Leclerc stores, the closest listed store is over 10 km away. No store meets these two criteria, hence it does not seem obvious 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 competitor authorities. As they do not have exact store locations, they center catchment areas on cities. They use the Herfindahl-Hirschman index (HHI) as a measure of local market concentration, and approximate market shares by selling areas. 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). The distance used in displayed estimations is 10km, while robustness checks are performed with 2.5, 5 and 20km.

Allain et al. (2016) also compute HHI based on store size and use a 10km radius for supermarkets and a 20km radius for hypermarkets, which they deem to be relatively high. They note that 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 a discount store. Estimations displayed take into account all hypermarkets within a 20km radius, and all other stores within 10 km. Robustness checks are performed with 30km/15km and 10km/5km distances.

2.2 Products and comparison methodology

As of March 2015, only national brand products are covered by the website. Products are identified at the bar code level. There are seven food product categories: 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 categories: health and beauty, household, pets and home and textile. Products are further classified within product families. Regarding chain comparisons, the number of products covered within 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 rate. 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 store level comparisons, all products found at both stores are used in

comparisons.

Table 3: Overview of product families within sections

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); Pre-cooked dishes (205); Pasta, Rice and Flour (328); Soups (270)
Sweet grocery (2,099)	Biscuits (294); Coffee and Tea (368); Chocolates and sweets (450); Desserts, Sugar and Sweeteners (318); Breakfast (453); Cakes (215)
Vegetables and fruits (65)	Fruits (65)

Price records obtained from the website include all products used in each store level comparison. As a consequence, there are 12,318 unique products 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. The five largest sections, regardless of the criterion, are Fresh products, Health and Beauty, Savoury Grocery, Sweet Grocery and Drinks. Families within each sections are detailed in Table 3. 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.

Table 4: Overview of product section weights

	All products		≥ 500 obs		≥ 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

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.

2.3 Price comparison results

Results for chain level comparisons performed according to the website methodology are reported in table 5. 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 included products. 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 comparison result to 0.4%.

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.

Results for store level comparisons performed according to the website methodology are reported in table 6. The website lists 99 comparisons between a Leclerc and a Geant Casino. On average, the Geant Casino is found to be 1.8% more expensive than its Leclerc competitor. In some cases, Geant Casino is cheaper than Leclerc. A Geant Casino is indeed found to be 0.6% less expensive than its Leclerc competitor. In general, national level comparisons appear to be relatively representative considering the distributions of store level comparisons.

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

	Nb pairs	Comparison of Leclerc stores vs. competitors by chain						
		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.

2.4 Comparison dynamics

Price records can be used to compute variations in chain prices across periods, which allows to gain some understanding regarding the evolution 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 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.

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

Date	Auchan	Carrefour	Cora	Geant	Intermarche	Leclerc	Système 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

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 and Carrefour Market (+5%). Auchan, Carrefour, Geant Casino, Intermarche and Système U display rather similar price levels (+3% to +4%). After May 2011, most chains exhibit a progressive loss in competitiveness 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, 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 Système U) in March 2015 which are between 6% and 7% more expensive than Leclerc.

Intra-chains 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%) contrasting 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 and Carrefour Market are persistently found to be 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.

3 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 adopt a rule of thumb which weighs the cost and time of transportation with prices and products expected to be offered in a store of a given chain. On the other hand, without being deceptive per se, such comparisons may simply reflect differences in store characteristics such as location, size or market competitiveness. In order to address this issue, this section investigates the impact of store and market characteristics on prices, and study intra brand price heterogeneity.

3.1 Store price determinants

In order to study the relation between price level, price dispersion, and market characteristics, we need to build measures of store price levels and store price dispersion. This is achieved by running the following regression:

$$\log P_{ij} = \text{Product}_i + \text{Store}_j + \epsilon_{ij} \quad (1)$$

where Product_i is a product effect and Store_j is a store effect. The coefficient Product_i associated with a dummy variable equal to 1 for products carried by store i provides information regarding the price level of store i compared to a reference store. The residual variation ϵ_{ij} can be interpreted as the percentage deviation of a store product price from its expected geometric mean. The average of the residuals is approximately null for each product and each store. Price dispersion is approximated at the store level by computing the standard deviation of the residuals. As a robustness check, the regression is also run successively for each supermarket chain so that the estimates of product fixed effects are chain specific. This specification is supported by Table 10 and Table 11 as they reveal significant degrees of price homogeneity within chains. These can be seen to be relatively consistent with price dispersion measured through the standard deviation of price residuals.

The store fixed effects estimated with specification 1 can be used to compute national store prices indexes. Formally, $(\text{Store}_j + 1) * 100$ yields a price index for Store j with base 100 for the

store used as a reference store in the estimation. Distributions by chain are reported in Table 8 and can be seen to be consistent with results from comparisons following the website methodology.

Table 8: Distribution of store price indexes by chain

	Nb	Avg	Std.	Min	Q25	Q50	Q75	Max
Auchan	112	107	4	102	104	106	109	120
Carrefour	165	108	5	95	106	109	110	128
Carrefour market	85	113	2	108	112	113	114	118
Casino	24	115	5	100	115	116	118	123
Cora	54	110	2	104	109	110	111	116
Geant Casino	88	101	2	100	100	100	100	105
Intermarche	184	107	2	102	106	107	108	114
Leclerc	513	100	1	93	100	100	101	106
Simply market	9	112	3	108	109	111	115	116
Systeme U	222	105	4	99	103	104	108	117
All	1456	105	5	93	101	103	109	128

Base 100: Leclerc in Limoges, used as reference store in the estimation of Equation 1.

These indexes are used to investigate the extent to which the heterogeneity in (aggregate) store prices may reflect store characteristics, as well as socio-economic and competition contexts. The following regression is estimated for this purpose:

$$\text{Store price index} = \text{Surface} + \text{HHI} + \text{Population revenue} + \text{Population size} + \epsilon \quad (2)$$

where Surface is the store surface, HHI is computed by considering store surface as a proxy for store sales and discounting market shares by the distance to each competitor (each store is considered to be the center of a market for which a HHI is measured), population revenue is the median household revenue taken at the municipality level, and population size a sum of nearby municipality population sizes weighted by distance. Results are reported in Table 9.

Table 9: Distribution of store price indexes by chain

	(0)	(1)	(2)	(3)
Intercept	100.33*** (0.12)	98.95*** (4.50)	108.95*** (8.08)	103.02*** (8.53)
Auchan	6.72*** (0.29)	7.07*** (0.30)		
Carrefour Market				
	(0.33)	(0.31)		
Carrefour	7.94*** (0.25)	8.21*** (0.26)		
Casino	14.95*** (0.59)	14.05*** (0.55)		
Cora	9.72*** (0.40)	9.82*** (0.40)		
Geant Casino	0.72** (0.32)	0.95*** (0.31)		
Intermarche	6.78*** (0.24)	6.50*** (0.23)		
Simply Market	11.28*** (0.94)	10.14*** (0.88)		
Système U	5.08*** (0.23)	5.42*** (0.22)		
HHI		0.01 (0.02)	-0.03 (0.03)	0.03 (0.04)
Surface		-0.17*** (0.03)	-0.11*** (0.04)	-0.15*** (0.04)
Population revenue		0.06 (0.46)	-1.35 (0.84)	-0.32 (0.87)
Population size		0.22** (0.11)	0.89*** (0.14)	0.52** (0.21)
R2	0.70	0.75	0.03	0.09
N	1,426	1,426	1,426	1,426

Columns (1) and (3) include dummies to control for regional specificities, the coefficients of which are not reported. Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$.

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. The affiliation of a store appears to be a strong determinant of its overall price level, which is consistent with the relative strong intra-chain price concentration previously obtained and other studies on retail chain prices. Hosken et al. (2008) and Chamayou (2016) observe similar results with gas station prices respectively in the US and in France (even though gas station chains do not follow uniform pricing policies).

Turolla (2016) estimates a structural model of demand and recover stores' price-cost margins. The paper finds that stores set prices according to the most competitive scenario, but that a

significant portion of large grocery stores take advantage of insufficient competitive pressure to distort offer and increase margins.

3.2 Chain level pricing policies

Even though French supermarket chains generally do not follow 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 other. The closest followers are Systeme U and Leclerc, for which the mode accounts for 39% and 38% of price observations on average.

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.

Intra-brand price heterogeneity can also be 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 the min is 6% hence it appears that a limited 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.

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.

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 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 literature².

In a theoretical paper, Allain et al. (2016) investigate the consequences of spatial discrimination and uniform pricing strategies on mergers. They show that when one retailer implements uniform pricing, the anticompetitive effects of a merger can affect consumers in markets not directly affect by the merger.

²Angrist 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".

4 Price Dispersion

4.1 Price dispersion and consumer information

Price dispersion is investigated at the pair level, based on the idea introduced in Chandra and Tappata (2011) that distance between sellers can be used as a proxy for consumer information. Pairs of competitors which are separated by a very low distance can indeed be expected to compete fiercely, so that they constitute a population in which the "law of one price" is the most likely to hold. On the other hand, a larger distance is likely to be tantamount to poorer consumer information, namely a competition setting which could be described by a model of search. A common feature of such models is the absence of pure strategy equilibria. In the single product case, mixed strategy equilibria have traditionally been given a dynamic interpretation, corresponding to the changes in ranks that can be observed among sellers over time. In the multi-product case, McAfee (1995) have shown that sellers can randomize on each product, either in a way that simply replicates the single product case (i.e. Varian (1980)), or in a way that involves a correlation between a seller's prices). Chandra and Tappata (2011) remarks than rank reversals, namely changes in the ranking of competing retailer prices, are to be expected whenever a significant distance between sellers is likely to generate search costs. On the other hand, when two retailers operate very close from each other, the "law of one price" is more likely to hold i.e. one can expect price alignment to be the norm. Chandra and Tappata (2011) measures rank reversals as the number of days during which the generally cheapest store is found to be more expensive. In this paper, rank reversals is measured over products. Formally, considering the prices p_{il} and p_{jl} of two supermarkets i and j over $l \in L$ days, such that $\sum_{l=1}^L \mathbb{1}_{p_{jl} > p_{il}} \geq \sum_{l=1}^L \mathbb{1}_{p_{il} > p_{jl}}$ is more frequent than the converse, the rank reversals statistic writes:

$$r_{ij} = \frac{1}{L} \sum_{l=1}^L \mathbb{1}_{p_{il} > p_{jl}} \quad (3)$$

Higher rank reversals are expected for pairs which are separated by a higher distance. Importantly, store differentiation leads to spuriously record low rank reversals. This issue is addressed by considering only pairs which are found to exhibit small aggregate price differences, and by running quantile regressions as in Chandra and Tappata (2011).

Table 12: Static store level comparisons (15 km - 100 obs min)

Chain A	Chain B	Nb pairs	B vs A avg comparison	Pairs won by A (%)	Share of products (avg %)			
					A wins	B wins	Draw	Reversed
Leclerc	Geant Casino	215	+1.4%	85.1	61.8	22.4	15.8	20.4
Leclerc	Carrefour	555	+9.1%	98.4	78.5	15.1	6.4	14.7
Geant Casino	Carrefour	89	+7.6%	98.9	70.8	25.1	4.1	25.1
Carrefour	Auchan	191	-0.3%	51.8	46.3	44.3	9.4	28.9
Carrefour	Intermarche	365	-1.0%	38.6	45.8	51.2	3.0	34.0
Carrefour	Systeme U	196	+2.6%	60.7	57.1	38.8	4.1	27.3
Auchan	Intermarche	212	+0.8%	61.8	54.0	43.0	3.0	32.9
Auchan	Systeme U	145	+3.1%	66.2	60.5	35.2	4.3	27.0
Intermarche	Systeme U	490	+1.0%	51.2	51.5	41.3	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 85.1% of the pairs. On average, a Leclerc sells 61.8% of products strictly cheaper than its Geant Casino competitor. 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.

Table 12 provides an overview of rank reversals measured by formula 3. Of all comparisons between chains exhibiting 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 85.1% 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 order to test the link between distance, taken as a proxy for consumer search cost, and rank reversals, the following regression is run:

$$r_{ij} = \beta_0 + \beta_1 \mathbb{1}\text{Short distance}_{ij} + X_{ij} + \epsilon_{ij} \quad (4)$$

where Short distance is a dummy which is defined to have value 1 when stores are relatively close to each other. In a first specification, distance as the crow flies is used, with a threshold of 5 km for the dummy variable. Only pairs separated by less than 10 km are included in the regression. The second specification uses distances in minutes computed by Google. The threshold is 12 minutes and pairs are only kept when the driving distance is below 20 minutes. Though 12 minutes may sound to be a relatively long trip criterion are roughly equivalent on aggregate.

AGG REG RESULTS + COMMENTS

Regression results provide support for a positive relation between separating distance and rank reversals.

Descriptives statics of dynamic price dispersion 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 second period or the reverse.

Table 13: Dynamic store level comparisons (15 km - 100 obs min)

Chain A	Chain B	Nb pairs	Dynamic "Rank reversals"	
			Pairs (%)	Products (%)
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.

4.2 Market price dispersion

Product price dispersion is measured at the national level, with a view to get some measure of potential heterogeneity among products. Given previous observations on chain pricing policies, only products where X observations are available from at least X different supermarket chains are taken into account. Indeed a product for which prices would be available only within chains which are characterized by significant price uniformity would be likely to be spuriously found to have highly concentrated prices. Product mean price, section and families are used to investigate variations in observed product price dispersion.

Table 14: Overview of average product dispersion by section at the national level

	Count	Price	Raw prices			Std	Residuals	
			CV	$\frac{Q_{75}}{Q_{25}} - 1$	$\frac{Q_{95}}{Q_5} - 1$		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	1358	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	1380	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	1423	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)

Standard error in parentheses.

Column "Price" is the mean product price in euros within each section. 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.

Table 14 provides an overview of average product price and dispersion measures by store section, while Table 15 reports the results of the regression:

$$\text{Product dispersion} = \text{Product mean price} + \text{Product store section} + \epsilon \quad (5)$$

Standard deviation and the spread between Q75 and Q25 can be seen to significantly increase with the value of the product. Coefficient of variation is found to slightly overcorrect this increase, as well as the specification used to compute price residuals.

Table 15: Regressions of product price dispersion measured at the national level

Prices Dispersion measure	Raw Std	Raw CV	Raw Q75/Q25-1	Res. Std	Res. Q75-Q25
Intercept	0.05*** (0.00)	7.25*** (0.13)	9.80*** (0.31)	4.60*** (0.10)	4.63*** (0.18)
Price	0.05*** (0.00)	-0.18*** (0.01)	-0.28*** (0.03)	-0.08*** (0.01)	-0.06*** (0.02)
Section Pets	-0.01* (0.01)	-0.84*** (0.21)	-0.87* (0.50)	-0.53*** (0.16)	-0.47* (0.28)
Section Drinks	-0.03*** (0.01)	-0.47*** (0.16)	-0.79** (0.37)	0.19 (0.12)	0.32 (0.21)
Section Savoury grocery	-0.01* (0.01)	-0.23 (0.14)	-0.89*** (0.34)	0.36*** (0.11)	0.59*** (0.19)
Section Sweet grocery	0.01 (0.01)	0.14 (0.14)	0.20 (0.34)	0.54*** (0.11)	1.05*** (0.19)
Section Fresh food	-0.00 (0.01)	-0.37*** (0.14)	-1.27*** (0.34)	0.76*** (0.11)	0.88*** (0.19)
Section Health and Beauty	0.05*** (0.01)	0.43*** (0.15)	0.37 (0.35)	0.80*** (0.11)	1.08*** (0.20)
Section Household	0.03*** (0.01)	0.32* (0.17)	0.06 (0.41)	0.75*** (0.13)	1.03*** (0.23)
Section Frozen food	0.02*** (0.01)	0.30 (0.21)	-0.04 (0.49)	0.89*** (0.16)	1.17*** (0.28)
R2	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: * $p < .1$, ** $p < .05$, *** $p < .01$.

While in the single product case, the main source of complexity attached to measuring market dispersion is the definition of markets, the presence of several products, in particular when their number is large as is the case with supermarkets, raises another major methodological issue. Absent quantity data, product choice is bound to be largely data dependent, and thus exposes results to various biases. Also, differentiation cannot be ignored by focusing on non differentiated pairs such as with competitor pair analysis, hence results are necessarily sensitive to potential misspecifications and must therefore be interpreted with great care.

The first approach followed in the paper consists in adopting markets as defined by the price comparison website. They are consequently all centered around one Leclerc store, and contain not other Leclerc store. Robustness checks are run to evaluate the impact of two potential issues: the absence of some supermarkets in the price data, and the presence of overlapping markets whenever two Leclerc are too close to each other. The first issue is addressed by narrowing the analysis to markets where virtually all active supermarkets are observed, and the second by considering that two Leclerc can be active in the same market, and alternatively by dropping overlapping markets.

The second approach simply consists in considering radiuses around Leclerc stores as is standard in the literature. This typically excludes some competitors listed by Qlmc (and may include others?).

All products for which prices are available at 2/3 or more of observed supermarkets are taken into account in the analysis. Markets for which less than 100 products satisfy this criterion are dropped. Measures of price dispersion are computed both with raw prices and price residuals. Figures obtained with raw prices are likely to overestimate consumer search related price dispersion since price comparison results suggest that persistent price differences (across products and time) 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. As a consequence, price dispersion can be evaluated either by considering the distribution of all price residuals, or by studying separately each product distributions (which can then be aggregated to obtain a measure of price dispersion at the market level).

We investigate how market price dispersion varies with the number of sellers or the intensity of competition captured by HHI. We also investigate the link between market price dispersion and an index of market prices. The market index is built by considering how expensive each store operating in a given market is in comparison to all stores affiliated to the same chain. Alternatively, store surfaces are taken into account to weight prices indexes in the computation.

5 Conclusion

Analysing competition between French supermarkets from different prospects leads to mixed conclusions regarding the competitiveness of the market. On the one hand, the presence of a "maverick" firm and the recent downward change of pricing strategy of a large supermarket chain in France suggest a significant degree of competitive pressure and a trend favorable to consumers. Meanwhile, evidence of price dispersion reveals that consumers generally face high search costs when it comes to comparing prices at supermarkets, and heterogeneity in supermarket price levels is not easily explained by proxies for competition intensity and socio-demographic variables. Observed shocks in pricing strategies in the recent period, if used to analyse consumer purchases (home scanner data) should allow to gain a better understanding of the role of competition.

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.
- Bertrand, M. and F. Kramarz (2002). Price dispersion then and now: Evidence from retail and e-tail markets. *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.
- 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.

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