Price dispersion and consumer search : Evidence from the retail gasoline market and the supermarket industry in France

Etienne Chamayou

NNT : 2017SACLX067

Thèse de Doctorat

de

l'Université Paris-Saclay

préparée à

l'École Polytechnique

ECOLE DOCTORALE n°578
Sciences de l'Homme et de la Société

Spécialité de doctorat: Sciences Economiques

Par

Etienne Chamayou

Price dispersion and consumer information: Evidence from the retail gasoline market and the supermarket industry in France

---

Thèse présentée et soutenue à l'ENSAE, le 21/09/2017
Composition du jury:

Claire Chambolle	Professeur, Ecole Polytechnique (Présidente)
Philippe Gagnepain	Professeur, Université Paris 1 (Rapporteur)
Erwan Gautier	Professeur, Université de Nantes (Rapporteur)
Thibaud Vergé	Professeur, ENSAE - CREST (Examinateur)
Francis Bloch	Professeur, Université Paris 1 (Directeur de thèse)

L'Ecole Polytechnique n'entend donner aucune approbation, ni improbation, aux opinions émises dans les thèses. Ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

Je tiens tout d'abord à remercier Philippe Février et Francis Bloch qui ont encadré cette thèse. Dès ma première année de Master, Philippe a répondu à mon intérêt pour l'étude de la concurrence en me proposant d'étudier le marché des stations services à partir de données de prix encore inexplorées. La perspective théorique et la bienveillante curiosité de Francis m'ont aidé à structurer mon questionnement et prendre du recul par rapport aux données tout au long de ma thèse.

Je remercie les rapporteurs, Erwan Gautier et Emmanuel Combe, pour le temps qu'ils ont consacré à la lecture de mes chapitres et les précieuses critiques qu'ils ont formulées.

Ce travail n'aurait pu voir le jour sans le soutien financier du CREST ainsi que l'environnement extrêmement favorable qu'il propose à ses doctorants. Je tiens en particulier à exprimer ma gratitude à Laurent Linnemer, Thibaut Vergé et Michael Visser pour leurs conseils avisés, leurs relectures et leur soutien quotidien. Mes remerciements vont également à tous les autres membres du Laboratoire d'Economie Industrielle qui ont accompagné mon cheminement: Saïd Souam, Philippe Choné, Julie Poirier, Isis Durrmeyer, Idrissa Sibailly, Sergei Kovbasyuk, Perrin Lefèbvre, Christophe Bellego, Andreea Enache, Manuel Marfan, Robert Somogyi, Laura Lasio, Emmanuel Paroissien, Amelie Mauroux, Arthur Cazaubiel, Alessandro Iaria, Alessandro Ispano, Julien Monardo, Hugo Molina et Jiekai Zhang.

Je remercie Ronan Le Saout avec qui j'ai pris beaucoup de plaisir à encadrer des projets de statistiques appliquées à l'ENSAE, échanger sur nos recherches respectives concernant le marché des carburants, et finalement collaborer à l'écriture d'un article. Notre participation aux Journées de Microéconomie Appliquée à Montpellier restera gravée dans ma mémoire comme l'un des meilleurs moments de ma thèse.

Enfin, je remercie ma femme Ria, mes parents, Nicolas et Béatrice, mes frères, Amaury et Jean, ainsi que tout mes amis pour le soutien moral qu'ils m'ont apporté dans les moments difficiles, m'aidant à garder le cap afin pouvoir les réunir pour la soutenance tant attendue.

Résumé

Cette thèse est un travail empirique sur la dispersion de prix, c'est à dire le fait qu'un bien homogène puisse être acheté à différents prix selon le point de vente visité, en violation de la célèbre loi du prix unique. L'approche développée s'inspire d'un courant de littérature initié par Stigler (1961), qui note que "la dispersion des prix est une manifestation -et, de fait, une mesure- de l'ignorance dans le marché". Cette proposition a une implication intéressante: tandis que le niveau d'information des consommateurs est difficile à caractériser, la dispersion des prix peut, quant à elle, aisément être quantifiée. Ainsi, il est potentiellement suffisant d'observer les prix pour détecter la présence de frictions informationnelles et mesurer le degré de compétitivité d'un marché.

Les deux premiers chapitres de la thèse explorent un large jeu de données de prix décrivant la concurrence entre les stations services françaises. Le premier chapitre étudie l'impact de la création d'une enseigne "discount" sur le marché. Fin 2011, Total S.A., le premier opérateur de stations services en France, a en effet créé une nouvelle chaine, "Total Access", avec un objectif explicite de reconquête de parts de marchés perdues au profit de la grande distribution. En l'espace de deux ans, 600 stations services ont été rénovées pour constituer le nouveau réseau. Pour la moitié d'entre elles, la conversion s'est accompagnée d'une baisse de prix de l'ordre de 10 centimes d'euro par litre. La réaction des stations concurrentes est analysée à l'aide de la méthode des doubles différences. Au niveau agrégé, la réponse enregistrée est une baisse de faible ampleur, inférieure à 1 centime d'euro par litre. Celle-ci masque néanmoins des hausses et des baisses en proportions équivalentes pour les 40% des stations dont la réaction excède 1 centime. Les baisses sont principalement mises en oeuvre par des stations opérées par la grande distribution, tandis que les hausses sont plutôt le fait de stations de groupes pétroliers et indépendants. Ces réactions soulignent l'importante segmentation du marché, au-delà de l'homogénéité apparente du produit.

Le deuxième chapitre interroge la pertinence des modèles, qui, à la suite de Varian (1980) identifient la dispersion des prix à un équilibre en stratégies mixtes. L'analyse dynamique des prix au niveau local soutient l'hypothèse de l'existence d'une relation entre la dispersion des prix et la présence de frictions informationnelles. L'ordre des prix des stations concurrentes tend à fluctuer d'autant plus dans le temps que la distance

qui les sépare est importante. Ainsi la dispersion est croissante d'un paramètre qui exprime un coût de recherche imposé aux consommateurs. De plus, les supermarchés et les stations des enseignes "discount" des groupes pétroliers, qui affichent généralement les prix le plus bas, sont plus susceptibles que les autres de maintenir des prix parfaitement alignés sur ceux de leurs concurrents. L'étude des mouvements de prix permet d'identifier nombre de ces stations comme leaders ou suiveurs dans leur marché. L'enseigne de la station s'avère déterminer fortement la stratégie de prix implémentée. Concernant les stations les plus onéreuses, la dispersion mesurée à travers les marchés locaux s'avère être corrélée négativement avec le coût du diesel, lequel fluctue principalement en raison du prix du pétrole, et positivement avec le nombre de stations présentes sur le marché. Ainsi, les résultats obtenus suggèrent que les groupes pétroliers et les indépendants s'adressent à une demande relativement peu sensible au prix et caractérisée par des coûts de recherche significatifs tandis que les supermarchés et les stations "discount" se disputent des consommateurs bien informés et très sensibles au prix.

Le dernier chapitre s'intéresse à la concurrence dans la grande distribution. Le groupe français Leclerc opère un site internet qui permet de comparer chacun de ses magasins à une sélection de ses concurrents, et affiche des comparaisons agrégées entre enseignes au niveau national. Un large jeu de données de prix a été extrait du site, comprenant plus de 4 millions de prix provenant de près de 2,300 supermarchés. Ces données sont utilisée pour investiguer l'hétérogénéité des niveaux de prix et de la dispersion à travers les marchés locaux. L'analyse révèle que les prix sont largement déterminés par l'enseigne du magasin, et difficiles à expliquer par les caractéristiques observables du marché, en particulier par des mesures standardisées de la concentration. Une enseigne se distingue par la très grande uniformité de ses prix au niveau national. Les comparaisons de prix entre supermarchés présentent des résultats d'autant plus volatiles que la distance qui les sépare est importante, ce qui est cohérent avec l'hypothèse d'un lien entre dispersion des prix et frictions informationnelles. Au sein de chaque marché, le prix le plus bas et le prix le plus élevé présentent en moyenne un écart de 17%. La prise en compte d'effets fixes magasins ramène cet écart à 10%. La dispersion des prix est faiblement expliquée par les caractéristiques observables du marché, mais elle croît fortement avec le niveau de prix. Ces observations sont cohérentes avec l'hypothèse que les magasins sont en capacité d'extraire des rentes plus élevées lorsque la fidélité ou le manque d'information des consommateurs relâchent la concurrence.

Contents

Chapter 1

General Introduction

This dissertation explores the possibility to learn about the intensity of competition within local markets though the study of rich price data. This approach was motivated by a famous paper, "The Economics of Information" (Stigler (1961)), which stressed the necessity to take the role of information into account in our understanding of competition: "A buyer (or seller) who wishes to ascertain the most favorable price must canvass various sellers (or buyers) - a phenomenon I shall term "search". [...] Price dispersion is a manifestation - and, indeed, it is the measure - of ignorance in the market.". The first two chapters of this thesis use daily price data of nearly all gas stations in France to study competition on the French retail gasoline market. The last chapter measures and analyses price dispersion in the French supermarket industry, exploiting a large cross section of prices. Importantly, while we exclusively study prices of perfectly homogeneous products: standard diesel and national brand products available all over France on supermarket shelves, retailer locations and characteristics play a significant part in shaping competition. The study of price dispersion thus requires thorough discussions regarding the definition of local markets and the role of differentiation. The main conclusions of this dissertation are that consumer search costs do soften competition in studied markets, and that even in the absence of an exogenous shock or data on market shares, a detailed observation of price strategies typically reveals a lot of information about local competition, and can thus provide good guidance for a regulator. Both these findings support initiatives in favor of price transparency.

1 Price dispersion and consumer search: a short survey of the literature

Stigler (1961), the seminal paper of the literature, does not model the link between price dispersion and search. Price dispersion is assumed to be "a function of the average amount of search", hence the necessity to understand the first order determinants of

search. The paper lists several intuitions:

"1. The larger the fraction of the buyers expenditures on the commodity, the greater the savings from search and hence the greater the amount of search.

2. The larger the fraction of repetitive (experienced) buyers in the market, the greater the effective amount of search (with positive correlation of successive prices).

3. The larger the fraction of repetitive sellers, the higher the correlation between successive prices, and hence, the larger the amount of accumulated search.

4. The cost of search will be larger, the larger the geographic size of the market."

In order to provide some evidence, Stigler (1961) cites prices observed at competing car dealers for a specific automobile, as well as price quotes for anthracite coal collected from bids for federal government purchases[1].

Ten years later, Diamond (1971) raises a famous paradox, noting that the "law of one price" can hold despite imperfect information. The paper indeed shows that the introduction of positive search costs in a standard Bertrand setting implies a unique equilibrium at the monopoly price. Indeed, in equilibrium, consumers do not search as they expect the monopoly price to be set by all firms, and the ensuing absence of search makes any deviation from the monopoly price unprofitable.

1.1 Varian (1980): the emergence of a paradigm

The foundations of the theory of consumer search and price dispersion were laid during the 1970s, with the study of optimal search rules[2], and the 1980s regarding the modelling of price dispersion with firms and consumers respectively optimizing price and search decisions in equilibrium. Most common assumptions of the literature are the following:

- Firms sell a perfectly homogeneous good
- Consumers incur search costs to discover prices beyond the first price
- Distribution of prices is perfectly anticipated by consumers
- Consumers can follow different search protocols:

[1] Prices for the automobile ranged from $2,350 to $2,515. The average price was $2,346 and the standard deviation was 42. Regarding anthracite coal, price per ton varied between $15.46 and $18.92. The average was $16.90 and the standard deviation $1.15. Stigler (1961) argued that the car represented a higher percentage of the budget of a household budget than anthracite coal did for the government budget, and the higher price dispersion for the former thus seemed to support the first intuition listed above.

[2] Cf. Rothschild (1973) for a survey

- Sequential search: consumers discover one price after another in a random order. Their optimal search rule then boils down to a reserve price (constant if cost of search is linear). Search stops as soon as a price below the reserve price is discovered.
- Fixed sample size search: consumers decide how many prices they want to know and will then buy at the lowest price received. Optimal search basically implies the computation of the optimal size of the sample.
- Clearinghouse model: consumers can either discover all prices or remain uninformed (made popular by its simplicity and later the internet context)

A paradigm has emerged following Varian (1980) which obtains price dispersion through the use of price randomization by firms in equilibrium. The paper considers a clearinghouse setting where consumers are either perfectly informed about prices, or uninformed consumers who have no information. All consumers are assumed to share the same reserve price for the good of interest. In a situation where all sellers would set the reserve price, each seller has an incentive to slightly undercut its competitors in order to attract informed consumers. Bertrand competition thus operates, but at some point in the undercutting process, each seller is better off giving up price competition and moving back to a price equal to reserve value. Varian (1980) shows that there is no equilibrium in pure strategies, but that a symmetric mixed strategy equilibrium can easily be derived. The Diamond paradox and Bertrand competition correspond to the extreme cases where all consumers are uninformed or informed. The paper offers an interesting prediction about the effect of an increase in the number of firms operating in the market. When the number of competitors increases, the chance for any given firm to capture the demand from informed customers decreases. As a result, higher prices tend to gain weight in the equilibrium price distribution. On the other hand, the net effect is still positive for informed customers: expected price paid decreases (more firms compete and shoppers pay the lowest price). For uninformed customers, expected price unambiguously increases. Mixed strategy equilibria have been shown to result from imperfect information with all standard information settings[3]. Regarding the relation between consumer information and price dispersion, the survey Baye et al. (2006) stresses that the literature leads to conclude that "price dispersion is not a monotonic function of consumers' information costs or the fraction of "shoppers" in the market". In relation with retail gasoline prices, Tappata (2009) has established that different levels of price dispersions, as a consequence of cost variations, could justify different consumer search intensities and hence different demand elasticities. This provides a potential explanation for asymmetries in downward and upward adjustment of retail prices to cost fluctuations.

[3]E.g. Burdett and Judd (1983) with identical consumers and non sequential search, Stahl (1989) with sequential search.

1.2 Empirical evidence

For lack of adequate data, empirical evidence on price dispersion has somewhat lagged behind the developments of the theory of price dispersion and consumer search. In the 2000s, the development of the internet has fortunately allowed the emergence of a dedicated stream of literature. Baye et al. (2006) note that the statistics most commonly used to account for price dispersion are:

- the standard deviation (sensitive to inflation)
- the coefficient of variation (standard deviation divided by expectation so that it is homogeneous of degree zero in the level of prices) which allows comparison across time and between products
- the sample range (p_{max}-p_{min}, highly sensitive to extreme values)
- the gap (p_2-p_{min}, the difference between the two lowest prices) proposed by Baye et al. (2004) which, in the absence of quantity, allows to give more weight to low price firms (yet also sensitive to extreme values)

Lach (2002) proposes a pioneering study seeking support for mixed strategy equilibria. The paper uses a dataset of store-level monthly prices of four homogeneous products sold in Israel, and shows that price dispersion persists even after controlling for observed and unobserved product heterogeneity. The paper furthermore exhibits significant intra distribution mobility (changes in firm prices' ranking over time) and argues that "we would need a lot of idiosyncratic "large" shocks arriving every month to destroy intertemporal rank correlation". The author thus concludes that the finding is "consistent with Varian (1980) argument about the need for "sales" (randomized prices) when consumers search rationally for the lowest price".

Several other papers have examined price transition probabilities (change of quartile in the price distribution at each date etc.), in particular with shopping comparison site data (e.g. Baye et al. (2004)). Results also tend to support the existence of randomization of prices by firms. Chandra and Tappata (2011) take advantage of a rich dataset including daily prices of gas stations over one year and a half. Gas stations' localization is known which offers the possibility of a detailed analysis of rank reversals. The authors find that price rankings vary significantly over time but are more stable among stations at the same street intersection (consistent with the idea that a better information should reduce price dispersion). Price dispersion increases with the number of firms in the market, decreases with the production cost and increases with search costs.

2 Data

The two main sources of data used in this dissertation are price comparison websites. The first dataset describes retail gasoline prices in France between 2011 and 2014. Prices were collected on a daily basis via a script. The other dataset is composed of two large cross sections of French supermarket price records.

2.1 The French retail gasoline market

Since 2006, French law makes it mandatory for any gas station having sold over 500m^3 of gasoline the previous year to keep prices posted on a governmental comparison website[4]. This legal obligation was introduced by the French Ministry of Finance with a view to increase transparency in the retail gasoline market.

From September 2011, I have used a script to collect retail gas station prices on a daily basis. Data downloaded from the website include over 8,000 gas station addresses, amenities (presence of a shop, car wash etc.). Geocoding APIs were used to obtain gas station gps coordinates and compute the distance separating gas stations.

2.2 The French supermarket industry

The last chapter of the thesis focuses on price dispersion in the French supermarket industry. Data were collected from www.quiestlemoinscher.com, a comparison website that was created by one of the largest grocery store chains in France. Two large price cross sections, respectively dating back to May 2014 and March 2015, allow to compare prices posted by competing supermarkets on two occasions.

[4] http://www.prix-carburants.economie.gouv.fr

Chapter 2

Rebranding in the French gasoline market: local competitive effects of price decreases

Abstract. Total S.A., the supermajor oil company, operates the largest gas station network in France. End of 2011, the company launched a new chain, "Total Access", with the stated goal of recapturing market shares lost to supermarkets. Within two years, 600 existing gas stations across France were thus rebranded to form the new chain. For half of them, the conversion was accompanied by an approximate 10 euro cent per liter drop in prices. The reaction of competitors is studied using difference in differences regressions with daily price data obtained from a comparison website. The measured aggregate response is a slight decrease, of less than one euro cent per liter. It yet conceals increases and decreases in equivalent proportions for the 40% of the competitors which are found to change their pricing policy by one cent or more. Decreases are mainly implemented by supermarkets, whereas gas stations operated by oil groups and independent networks account for most of the increases. These reactions lead to highlight the role of market segmentation, beyond apparent product homogeneity.

Key Words: *Retail gasoline, pricing, event study*

1 Introduction

During the 2012 French presidential election campaign, historically high retail gasoline prices stirred a controversy regarding the competitiveness of the market. Following the election, a report was thus requested by the newly formed government (Bellec et al. (2012)), and claims of insufficient competition in the retail market were found to be without merit. Cited evidence included the development of supermarket networks, the withdrawal of several oil companies, and low estimated net margins, of nearly 2 euro cents per liter. The bulk of the then recently observed price increase was attributed to oil price fluctuations, and to the multiplication of costly environmental constraints. The report did however bring little microeconomic evidence regarding competition between gas stations, focusing essentially on profitability. Beyond the price difference of approximately 8 euro cents per liter between supermarket and other gas stations, net margins were estimated to be low for all retailers.[1].

This paper investigates competition on the French retail gasoline market, using a change in strategy implemented by the largest gas station operator in France, Total S.A. In 2011, the group indeed significantly expanded its discount offer through the creation of a new chain, "Total Access". Between September 2011 and December 2014, approximately 600 existing gas stations were consequently rebranded. Among these, 250 used to be operated under the Elf brand, originating from a prior merger between Total S.A. and Elf, and set retail prices which were competitive with those of competing supermarket gas stations. The other gas stations were operated under the Total brand and their prices were decreased by 10 euro cents per liter on average, in order to match the discount strategy of the newly created chain. Using difference-in-difference regressions, we analyse the diesel price reaction of competitors within markets affected by the rebranding.

Following suspicions of asymmetries in upward and downward retail price adjustments to oil price fluctuations, a rich literature on retail gasoline markets has emerged (Eckert (2013)). Several papers have sought to use merger and acquisition evaluations to shed light on market competitiveness[2]. Difference in differences analyses have yet yielded contrasted results. Hastings (2004) and Taylor et al. (2010) studying the same

[1] This spread is still to be observed in the gross margin, which measures the difference between before tax price and wholesale cost (including transportation). The net margin, however, once distribution costs are taken into account, is estimated to be similarly low across retailers as the lower gross margin of supermarkets is offset by higher volumes.

[2] Einav and Levin (2010) and Angrist and Pischke (2010), regarding recent advances in applied industrial economics, express a disagreement as to which methodologies are most appropriate. Einav and Levin (2010) advocate structural approaches which allow to estimate the demand function and perform merger simulations. Angrist and Pischke (2010) criticize the overwhelming use of such 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", citing in particular Hastings (2004), to which our approach is very similar, as we do not have data on volumes.

concentration operation in California, albeit with different data samples, have reached significantly different conclusions: a 5 cent increase per gallon in the former, and only a 1 cent increase in the latter. Houde (2008) has simulated a merger between two retailers in Canada and evaluated its impact ex-post using a difference in differences approach. The paper stresses the importance of heterogeneity across stations and markets, which make results highly sensitive to biases in data sampling. Findings generally suggest weak price adjustment by competitors, conditioned by the structure of competition (presence of independent gas stations in Hastings (2004), road traffic and local market power in Houde (2008)).

The data used in this paper allow to overcome the main problems raised by the literature. Price records have a daily frequency which ensure a reasonably fine observation of pricing strategies. In addition, sellers are observed in a virtually exhaustive way, including rebranded stations, other stations operated by Total S.A., and competitors. We first estimate aggregate effects using difference in differences regressions by defining markets through radiuses of 1, 3 or 5 km around each rebranded gas station. We then estimate the reaction of each gas station, in order to measure heterogeneity and relax the constraints previously imposed by our market definition. We find that supermarkets have generally implemented small price decreases in reaction to a nearby rebranding, while oil companies and independent gas stations have rather slightly increased prices. The heterogeneity of reactions leads to stress the necessity to observe a large sample of gas stations. Focusing on the Paris metropolitan area, we observe that prices have been adjusted upwards within Paris and nearby, but downwards in gas stations located further away (and for France in general). This heterogeneity can be accounted for by the low penetration of supermarket gas stations in Paris. Besides contributing to the understanding of competition in the French retail gasoline market, our paper illustrates how results obtained through difference in difference regressions to evaluate a merger can diverge, depending on market structure and concentration in the observed sample.

Regarding competition in the French retail gasoline market, the results lead to stress the existence of market segmentation and frictions. The mild aggregate response to a large shock in prices may indeed seem surprising if one refers to standard Hotelling competition. The small price decreases implemented by supermarkets are yet consistent with intense competition and low pre-existing gross margins. Conversely, more expensive gas stations appear to have kept addressing a less price sensitive demand of loyal or captive customers, for whom the launch of Total Access may indirectly have led to an increase in prices. In this regard, the absence of data on entries and exits is unfortunate.

The first part of the paper describes the French retail gasoline market as well as the rebranding operation implemented by Total S.A. The second part provides an overview

of the data with descriptive statistics. Finally, the last part details the estimation strategy and the main results.

2 Context and Data

2.1 The French retail gasoline market

The data cover a three-year time span, from September 2011 to December 2014. Over the period, retail diesel price increases until April 2012 (approximately 20 cents to reach 1.40 euro per liter after tax), and decreases thereafter. Diesel accounts for 75% of household gasoline consumption. Three main types of retailers operate on the market: supermarket chains, oil companies, and independent networks. The price difference between a supermarket gas station and a competitor from another type is generally close to 8 euro cents per liter. The data do not contain information about vertical relations, hence their potential impact is not taken into account in the analysis.

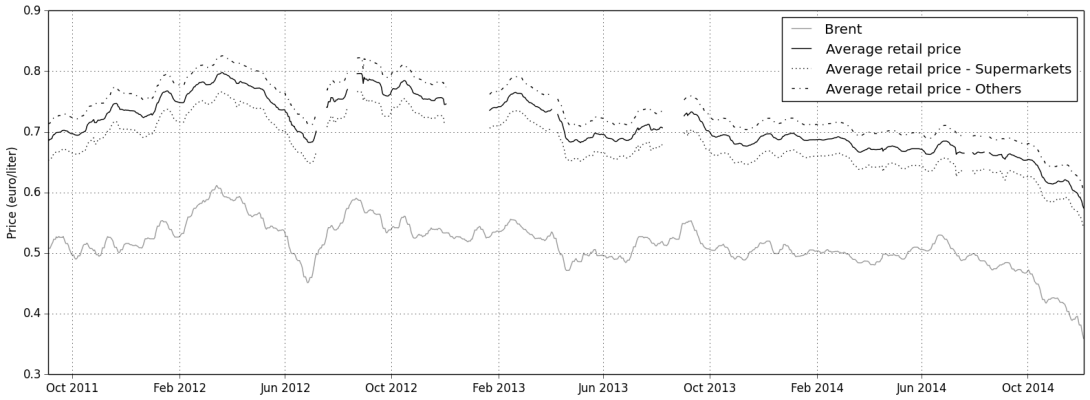
Oil refining has undergone significant restructuring over the last years (acquisitions and shutdowns e.g. Petroplus in Normandy in 2012), and these changes could have had some impact on prices. Taxes account for nearly 60% of retail prices. They consist of fixed components (exhibiting slight variations across regions) and the VAT.

The Bellec et al. (2012) report relies on the evolution of the gas station network as well as estimations of profitability (net margin of 0.2 to 2 euro cents per liter) to support the thesis of a highly competitive market. The total number of gas stations has indeed plummeted from roughly 40,000 to 12,000 between 1980 and 2012. Meanwhile, the market share of oil companies has kept decreasing, while supermarkets were expanding their networks. Gas stations operated by supermarkets account for approximately 50% of the total number of gas stations over the studied period, while their market share is close to 60%. The Bellec et al. (2012) report notes that there is no reliable data source regarding the number of operational gas stations. Between 200 and 300 gas stations are assumed to close permanently every year, and about 100 to open, mainly within supermarket chains.

2.2 Data

Data come from the price comparison website www.prix-carburants.gouv.fr, on which French gas stations are legally required to keep prices up to date since 2007. The accuracy of the data is ensured by regular controls which can lead to financial penalties. Small gas stations are not subject to this legal obligation. As a consequence, 10,200 gas stations out of 12,000 over France are observed. Other price comparison website exist but tend to be less exhaustive as the legal obligation imposed on retailers only concerns the governmental price comparison website.

Figure 2.1: Evolution of diesel and Brent prices



Reading Note Diesel prices are average retail prices excluding taxes. Some periods are missing as prices could not be collected for technical reasons.

Source: Governmental website for diesel prices; UFIP website for Brent.

The analysis is based on daily price records spanning slightly more than three years, from September 4, 2011 to December 4, 2014. Some short periods are missing due to technical issues encountered during data collection (cf. Figure 2.1). Gas stations located on the island of Corsica were excluded (130 gas stations) as well as gas stations located on highways (less than 500) since they are part of specific markets. Some gas stations for which observations were too short or price dynamics looked suspicious, mainly because of excessive rigidity, were also discarded. We also drop gas stations which already compete with a Total Access gas station as of September 2011 (i.e. when the rebranding occurred prior to the beginning of our study) or which compete with a gas station for which the observed price change is below 4 euro cents per liter (500 gas stations). Former Total stations for which the price change is small are likely gas stations which were used during the test phase, meaning that the price decrease has occurred before the rebranding (cf. comments on Figure 2.3). In order to estimate a potential effect of the rebranding on other gas stations operated by Total S.A., we include them in the analysis. About 300 are located within a distance of 5 km of a Total-Total Access, and 100 within 5 km of an Elf-Total Access. Finally, we only keep gas stations which are located within 10 km from a rebranded gas station (cf. part 4 on the estimation strategy). About 4,700 gas stations are dropped based on this filter, leaving us with approximately 3,100 gas stations and 3 million price observations. About 1,300 are considered to compete with a Total-Total Access gas station and 400 with an Elf-Total Access. Other gas stations located within 5 to 10 km from a Total Access are used as a control group in our difference in differences analysis, as they are assumed to not directly compete with Total Access gas stations. Our data include gas station characteristics and locations. Census data and other publicly available data

Table 2.1: Key determinants of prices and Total gas station brands

	(1) Price OLS	(2) Price OLS	(3) Price OLS	(4) Price OLS	(5) Total Access Probit	(6) Elf Probit
Supemarkets	-7.15***	-7.10***	-7.09***	-5.66***		
(Ref. Oil companies)	(0.31)	(0.33)	(0.33)	(0.42)		
Dist. to closest		0.07**	0.05**	0.07***	-0.011*	-0.009***
supermarket (km)		(0.03)	(0.02)	(0.02)	(0.006)	(0.003)
Nb competitors		-0.14***	-0.11***	-0.10***	0.015***	-0.000
within 3 kms		(0.05)	(0.03)	(0.04)	(0.005)	(0.003)
Nb same group gas stations		0.35***	0.32***	0.25***	-0.011	0.007***
with 3 kms		(0.08)	(0.06)	(0.05)	(0.009)	(0.003)
Rural area			0.49***	0.53***	-0.074**	-0.049**
(Ref. Other urban area)			(0.08)	(0.09)	(0.032)	(0.022)
Urban area of more than			-0.47***	-0.47***	0.084***	0.025*
100,000 inhabitants			(0.13)	(0.13)	(0.032)	(0.015)
(except for Paris)			1.11**	1.13**	-0.106**	-0.056***
Paris and surroundings			(0.40)	(0.41)	(0.040)	(0.014)
Pay at the pump				-0.73***	0.171***	0.158***
				(0.12)	(0.020)	(0.011)
Premium gasoline				1.46***	-0.070	-0.258***
				(0.13)	(0.065)	(0.053)
Shop				0.46***	0.114*	0.030
				(0.14)	(0.047)	(0.045)
Car repair				0.54***	0.015	-0.075*
				(0.15)	(0.072)	(0.047)
Intercept	139.60***	139.66***	139.68***	138.12***		
	(0.18)	(0.35)	(0.30)	(0.51)		
Nb. Obs.	7417	7417	7417	7417	1582	1770
Adj. R2	0.560	0.575	0.582	0.607	0.131	0.222
Reference probability					0.219	0.112
Sensitivity					0.094	0.139

Reading Note: Columns 1 to 4 account for the regressions of gasoline prices (cents, including taxes) on September 9, 2011. Column 5 describes the determinants of a Total-Total Access rebranding (marginal effects of a probit model run with all Total gas stations). Column 6 accounts for the specificities of gas stations belonging to the Elf network at the beginning of the period (marginal effects of a probit model run with all Total and Elf gas stations). Regional controls are included. Errors are clustered at the regional level. Significance threshold: *** 1%, ** 5%, * 10%.

were added based on the municipality of each gas station. For each of them, we built several variables describing competition such as the distance to the closest competitor and the number of competitors located within a 1, 3 or 5 km distance.

Table 3 describes price level determinants in level on a given day[3]. Supermarket gas stations are on average 5 to 7 cent cheaper than other gas stations and this single distinction accounts for 56% of the price variance on a given day. The presence of a shop, car services (e.g. car wash) or the availability of premium gasoline generally imply a 0.5 to 1.5 cent higher price. On the other hand, the presence of automated pumps is associated with a 0.7 cent lower price. Prices are generally higher when the

[3]Retail prices are primarily determined by crude oil prices. Studying prices on a given day allows to shed lights on the influence of observable variables describing gas stations and their markets.

population density is lower but also in Paris and its close surroundings. Competition is found to have a significant impact (distance to the closest supermarket, number of competitors within 3 km), yet of low magnitude and therefore likely not noticeable for consumers. Prices are also slightly higher when concentration is higher. Results are consistent with Hosken et al. (2008), who find gas station prices in Washington DC to be largely accounted for by gas station chains, characteristics, market demand structure, and to a lesser extent by competition. Zimmerman (2012) has studied the impact on prices of the development of supermarket gas stations in the US. However, the situation is yet different as these account for less than 10% of gas stations in the US.

2.3 Competition definition

The goal of this paper is to study the price reaction of competitors to the creation of Total Access via a difference in difference analysis. This requires to carefully define competition so as to obtain adequate treatment and control groups. The usual procedure implemented in the literature consists in relying on a measure of distance, which can also be defined in several ways, and to check the robustness of results to changes in the definition. Hastings (2004) uses a radius of one mile (1.8 km), based on interviews with retailers, to study the impact on prices of a merger between two chains in California. The robustness of results is checked with alternative radiuses of 0.5 and 1.5 miles. The control group is composed by non affected gas stations located within the urban area (San Diego or Los Angeles). Bruzikas and Soetevent (2015) measure the price effect of the installation of automated gas stations in the Netherlands. The intensity of competition is captured by the number of gas stations within 2 and alternatively 5 km. In order to study a merger between two chains in Canada, Houde (2008) uses a structural approach based on a Hotelling model. The definition of markets essentially relies on commuting patterns. Results of the structural approach are compared with results obtained via a difference in differences analysis which use a distance based on time (30 seconds, 1 or 1.5 minute). Results are overall consistent but sensitive to the market structure and the analyzed sample.

We therefore consider a 5 km radius to be a relatively consensual definition of competition and perform robustness checks with 1 and 3 km radiuses. We also perform a specific analysis for the Paris metropolitan area, where we can distinguish the city of Paris, its close surroundings, and the whole region.

2.4 Total Access: strategy and network development

Upon creating Total Access, Total S.A. has stated that its goal was to recapture market shares lost to supermarkets. The idea was to associate the development of a discount

offer with a quality brand image. Before the creation, Total S.A. distributed gasoline through three chains: Total, Elf and Elan. The Elan chain is essentially to be found in areas where both demand and competition are low, typically in rural areas. It was not impacted by the rebranding operation. The Elf chain was inherited from a prior merger, and was until then the discount chain of Total S.A. The stations belonging to the Elf chain were rebranded Total Access and the pricing policy was left unchanged. Some minor modifications may have affected stations such as the addition of a premium gasoline. Finally, as regards former Total gas stations, the rebranding was accompanied by a significant decrease in prices, and occasionally some upgrades such as the the addition of some pumps. Gas station amenities may have been marginally reduced, for instance through a shrinkage in the number of products available at the gas station shop.

While Total S.A. has largely advertised the creation of its new chain, the identity of affected gas stations was not disclosed until their respective rebranding. The pace of the development of the network has been fairly regular following the announcement. Required renovation works last one to two weeks and were reported to cost between one and two million euros for each gas station. Total S.A. has also mentioned that the magnitude of the price decreases was decided based on each gas station's market environment. At the end of the observed period, Total Access gas stations could be found across all regions of France.[4].

Table 2 describes the competitive environment of Total-Total Access gas stations. Gas stations which are rebranded are generally located in markets exhibiting above-average competitive pressure. Their closest competitors are on average 1.2 km away, while they are 1.8 km away for those not rebranded. Their number of competitors is higher regardless of the size of the considered radius (1, 3 or 5 km). In "Ile-de-France", the administrative region which contains Paris, the city and its close surroundings contrast sharply with the broader surroundings. On the one hand, the number of competitors within 5 km is more than twice higher and the distance to the closest competitor is about twice smaller in the former than in the latter. On the other hand, the distance to a supermarket gas station is greater, with respective 2.2 km and 1.6

[4]Philippe Callejon, in charge of the creation of the chain, has given several interviews to explain the choice to convert classic Total stations to Total Access. "A test was implemented over 18 months within 45 cities. This test period was crucial. Prices, through a 9 cent decrease, were aligned on market prices. We have then examined at the impact of the price decrease on the economic model of gas stations. Conclusion: in order to make up for the price decrease, we must increase volumes. Very practically, we need gas stations big enough to have several pumps. They must be located on highly frequently roads so that there are enough potential customers. As a matter of fact, within the Total network, some gas stations are located on roads were the traffic is low. We have no influence on such elements, as well as the fact that it is nearly impossible to obtain the authorization to open a new gas station. The test has thus confirmed the necessity to create a network with two offers: gas stations with competitive prices, in periphery, where higher volumes resulting from potential demand can make up for the decrease in price, and proximity gas stations, with classic prices."

Table 2.2: Competition in Total gas station markets

	Nb.	Distance to closest			Nb competitors within		
	Obs.	competitor	supermarket	Total	1 km	3 km	5km
TOTAL	1787	1.64	2.05	5.55	0.71	3.39	6.96
- not Total Access	1413	1.75	2.18	6.01	0.69	3.17	6.50
- Total Access	374	1.21	1.55	3.85	0.81	4.23	8.70
ELF	269	1.17	1.60	2.57	0.86	4.68	10.28
Paris & close suburbs	137	0.83	2.22	1.07	1.11	9.36	24.76
Other Paris suburbs	180	1.64	2.08	3.00	0.52	3.51	8.54

Reading note: On average, the 1787 Total gas stations (whether rebranded or not) are located 1.64 km away from their closest competitor. Their average number of competitors whithin 3 km is 3.39. The 137 Elf and Total gas stations (whether rebranded or not) located in Paris and its closest surroundings are on average 2.22 km away from a supermarket.

km average distance. The market structure is therefore different in and close to Paris as it is characterized by a higher gas station density but less supermarket gas stations. The broader surroundings of Paris are rather similar to France as a whole in terms of market structure.

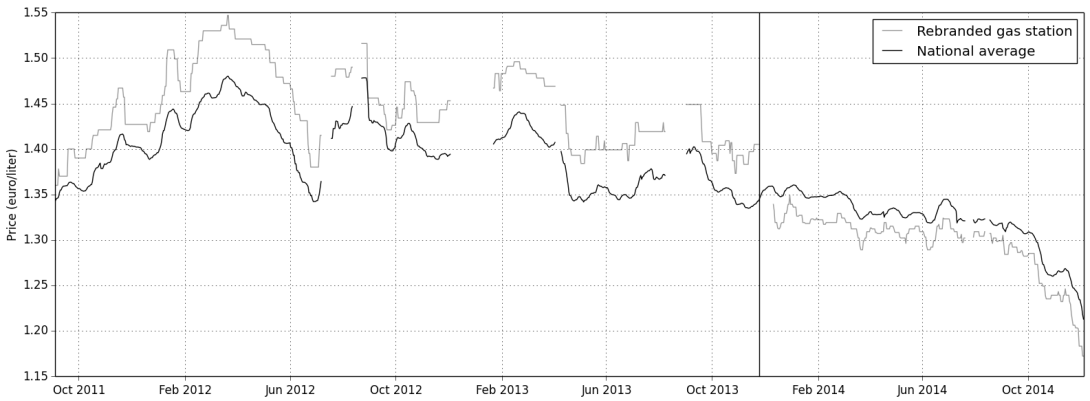
Tables 3 accounts for the impact of observed gas station characteristics on the decision by Total S.A. to rebrand its Total gas stations Total Access or not. Findings allow to confirm part of the communication of Total S.A. Changes are more frequent in large urban areas (except for Paris), when the number of competitors is high and the distance to a supermarket is low. Changes are less likely inside Paris and in rural areas. On average, a gas station located in Paris or its close surroundings is 10% less likely to be rebranded.

The possibility to pay at the pump, generally associated with above average volumes, implies a 17% increase in the probability of rebranding. Converted Elf gas stations are also generally located in urban areas, in particular around Paris, and are often equipped with pay at the pump systems. Their number of competitors is relatively normal compared to the overall Total gas station population. They however differ significantly as regards the presence of premium gasoline (26% less likely in an Elf gas station). The switch from Elf to Total Access could therefore have meant an increase in perceived quality, if premium gasoline has been made available. All these characteristics do not allow to predict which stations were going to be rebranded (the sensitivity i.e. the percentage of changes correctly predicted is 10%). This supports the hypothesis that competitors could not anticipate which gas stations were likely to be rebranded Total Access.

Gas station chain changes are imperfectly observed in the data. They were validated with publications of Total S.A. and the analysis of price series. A renovation work period is associated with each rebranding operation and lasts about 15 days, during which the gas station remains closed. The measure of the price effect via a difference in differences for Total-Total Access and Elf-Total Access gas stations as well as for other

Total S.A. gas stations correspond to the direct effect, namely a decision by Total S.A.[5]. The direct effect analysis is meant to verify that price decreases were actually implemented together with the rebranding, and whether Total S.A. may have changed its pricing policy for other gas stations at the same time without communicating about it.

Figure 2.2: Detection of a change in pricing policy for a Total Access gas station



Reading Note: The grey line represents the prices P_t of a Total Access gas station. The black line represents the national average price $\overline{P}_t$. The detected date for the change is represented by the vertical line.

The price series of a rebranded Total gas station is shown on Figure 2.2. Prices are initially above the national average. From December 2013 on, following an approximate 8 euro cent decrease in the pricing policy, prices are consistently below the national average. The temporal discontinuity is easily seen graphically and can been detected statistically as well. In the data, we observe 325 chain changes from Total to Total Access accompanied by a price decrease of 4 cents or more (i.e. 2 euros for a 50L tank), 250 chain changes from Elf to Total Access, and 350 other changes (within 5 km of a Total Access gas station). Approximately 200 are associated with a change from Esso to Esso express, for which a small price decrease is observed (below 2 cents). For other gas stations (except for very few of them), we do not observe any adjustment in prices. Changes essentially result from harmonization within supermarket chains. We excluded these gas stations from the analysis and checked that results were not affected. The conversions of Total and Elf gas stations occur progressively over the studied period. Figure 3 displays the distributions of Total SA gas station prices at the beginning and at the end of the period. In the latter, the distribution can be seen to be bimodal, reflecting the existence of the discount networks besides other gas stations with standard Total prices. The graph also displays gas stations which already set

[5]Only difference in difference results obtained with a national control are presented. Robustness checks were performed with regional controls and at a series level (cf part 4.1) and provided similar results.

Table 2.3: Total Access rebranding aggregate direct effets (euro cents) - DD regressions

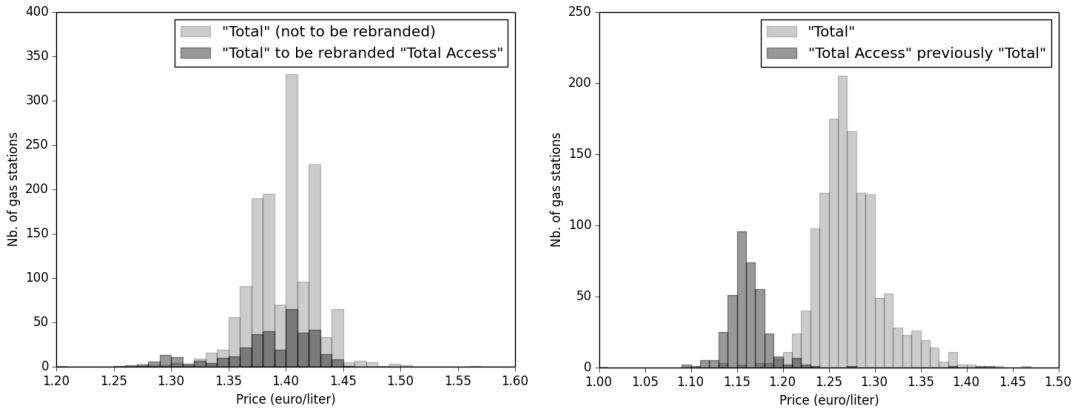
	(1) Price 5 km	(2) Price 3 km	(3) Price 1 km
Total-Total Access	-9.89***	-9.88***	-9.87***
	(0.17)	(0.17)	(0.16)
Total SA gas stations close to Total-Total Access	-0.17	-0.19*	-0.09
	(0.10)	(0.10)	(0.25)
Elf-Total Access	-0.08*	-0.08	-0.07
	(0.05)	(0.05)	(0.05)
Total SA gas stations close to Elf-Total Access	0.08	-0.01	-0.14
	(0.12)	(0.13)	(0.13)
Intercept	136.10***	135.81***	135.15***
	(0.17)	(0.17)	(0.17)
Nb. Obs.	301.500	282.231	249.040
Adj. R2	0.967	0.967	0.967

Reading note: Prices include VAT. Columns 1, 2 and 3 correspond to the following regression: $P_{it} = \gamma \cdot Treatment_{it} + \mu_i + \eta_{t,c} + \varepsilon_{it}$ with $\eta_{t,c}$ the temporal effects of a control group c, μ_i gas station fixed effects, γ the effect of the rebranding ($Treatment_{it}$ is a dummy which takes value 1 after the rebranding). The control group includes gas stations within a 5 to 10 km distance from Total Access gas stations. Errors are clustered at the regional level. Significance thresholds: *** 1%, ** 5%, * 10%.

low prices in September 2011 before their conversion, and were likely part of the test phase. These gas stations were identified through difference in differences regressions and excluded from the analysis. Results of the difference in differences analysis for rebranded Total and Elf gas stations, as well as for nearby Total gas stations, are reported in Table 1. Rebranded Total gas station are found to drop prices by an average 10 cents. No substantial decrease is observed for former Elf gas stations and for nearby Total S gas stations. Table 6 (cf. infra) details the distribution of price policy changes measured through an individual level analysis. The median price decrease for former Total stations is 10 cents, the ninth decile is 7 cents and the first decile 12 cents. Regarding former Elf gas stations, only one observation displays a price policy change above 2 cents (3 cent increase).

Market price variations over time are captured by daily fixed effects. Their estimates are almost perfectly correlated with wholesale diesel cost. We perform our analysis with prices including VAT in order to obtain the effect for consumers. Regressions with pre-tax prices yield similar results (closer to the gross margin of retailers), approximately 20% smaller in value due to the VAT.

Figure 2.3: Distributions of Total prices before and after rebranding



(a) September 4, 2011 (b) December 4, 2014

Reading Note: In black, Total gas stations which are rebranded Total Access. In grey, Total gas stations which are not rebranded.

3 Estimation and results

3.1 Estimation strategy

We perform a difference in differences analysis, close to the approach of Hastings (2004). We compare the evolutions in prices, before and after the rebranding, between competitors of rebranded gas station and control gas stations. We include all price records available until December 2014 to estimate reactions.

Prices are sticky (Gautier and Le Saout (2015)), typically remaining unchanged during 5 to 7 days. We perform our difference in differences regressions with weekly data (Friday prices) as a conservative approach limiting the overconfidence typically induced by stickiness. Competition is successively defined by circles of radius 1, 3 and 5 km.[6] We first estimate an aggregate model with treatment dummies $Treatment_{it}$ for Total-Total Access and Elf-Total Access changes, that we also cross with retailer type (oil company and independent vs. supermarkets). Regarding gas stations which are found to compete with several Total-Total Access gas stations, we take into account the date of the first conversion.[7] When a gas station is found to compete both Total-Total Access and Elf-Total Access gas stations, we do not take into account the Elf-Total Access rebranding. The expected impact of an Elf-Total Access conversion is indeed limited, given the absence of change in the pricing strategy. The model writes as follows, for a gas station i and a date t:

[6] We obtain similar results when we consider all gas stations within 10 km to be competitors, and include all others in the control group. Performing the analysis with daily data (or another day of week) also does not substantially affect results. It essentially increases the significance of estimated coefficients.

[7] Taking into account the date of change of the closest gas station does not affect results.

$$\mathrm{P}_{it} = \gamma \cdot Treatment_{it} + \mu_i + \eta_{t,c} + \varepsilon_{it} \tag{2.1}$$

The fixed effect μ_i controls for unobserved gas station characteristics. Including such fixed effects is necessary to avoid biases that could arise if treated gas stations are more likely to be of a certain type (e.g. oil or independent) associated with a specific pricing strategy. Time fixed effects $\eta_{t,c}$ are included for each date t and control group c. They control for shocks common to all gas stations, whether they are treated or belong to the control group. Finally, ε_{it} accounts for unobserved idiosyncratic shocks.

The validity of the estimation strategy relies on the existence of a trend common to the treatment and control groups. In the absence of treatment, price trends should thus remain similar for the control and the treatment groups. This means that $\mathbb{E}\left(Treatment_{it} \cdot \varepsilon_{it} \,|\mu_i, \eta_{t,c}\right) = 0$. French gasoline price variations depend closely on oil price fluctuations. This generally implies a strong correlation between any gas station price series. However, since the converted gas stations are chosen based on market characteristics, we want to include to compose our control group with gas stations that are far enough not to be affected but are close enough to face relatively similar market conditions. Our baseline specification thus includes all gas stations which are separated by more than 5 km and less than 10 km in the control group. We also consider building several control groups for each of the 13 French administrative regions as they might better capture local shocks such as the closure of a refinery. Finally, since gas stations are converted progressively over the whole period, we can implement an internal control, namely define the control group at any point in time as the gas stations which have not been converted to Total Access yet. Since all these methods give similar results, we only include outputs obtained with the baseline specification of the control group.

A major shortcoming of the aggregate approach is the fact we obtain an average effect, which is therefore sensitive to the definition of markets and the studied sample. With two well identified retailer types, and therefore various local market structures, it is meanwhile reasonable to expect heterogeneity among affected gas stations. We therefore estimate individual treatment effects with the following specification:

$$\mathrm{P}_{it} = \gamma_i \cdot Treatment_{it} + \mu_i + \eta_{t,c} + \varepsilon_{it} \tag{2.2}$$

The coefficient of interest γ_i is therefore estimated for each gas station, thanks to the richness of the date and the close relationship generally observed between gas station prices and their main determinant, the Rotterdam wholesale gasoline price. The individual approach allows to investigate heterogeneity in reactions and to discuss the robustness of results to our definition of market. In terms of methodology, it is close to a temporal discontinuity regression (Auffhammer and Kellogg (2011)), for which the randomization hypothesis boils down to the necessity for the conversion not to have

been expected by competitors. This hypothesis seems reasonable since Total has not announced specifically which gas stations were going to be rebranded.

In order to deal with temporal and geographical correlation of difference in differences estimators, we allow the error terms to be correlated within each region and on each date (Bertrand et al. (2004)), or simply within each region (Cameron and Miller (2015)). We have also performed the analysis by running one regression for each gas station with the Newey-West estimator in order to allow for temporal autocorrelation. Results displayed come from the aggregate estimation, with correlation allowed within regions. They are robust to the previously mentioned variations in specification. Generally, all measured increases or decreases whose value is above 1 cent are found to be statistically significant at the 5% level.

3.2 Results

Difference in differences analysis performed at an aggregate level for competitors of Total-Total Access gas stations (Table 4) reveal a slight average decrease, of 0.2 euro cents per liter, for all competitors located within 5km (respectively 0.3 cents within 3km, and 0.4 cents within 1 km, which means 20 euro cents for a tank of 50 liter). The impact thus decreases with distance. Regressions performed with competitors of Elf-Total Access gas stations yield a null aggregated effect. The weakness of these aggregated effects is illustrated by the comparison of the distributions at the beginning and the end of the studied period (Figure 4). While a significant change can be observed for Total-Total Access gas stations (Figure 3), no similar bi-modality can be observed. This analysis nevertheless hides significant heterogeneity related to retailer types. Supermarket gas stations are indeed found to have decreased prices by 0.6 cents when they were affected by a Total-Total Access conversion, and 0.2 cents for an Elf-Total Access conversion (3km radius). Other gas stations, on the other hand, have slightly increased prices in both cases. Within a 5 km radius, only supermarkets are found to have reacted. This is consistent with supermarkets reaching consumers further than other gas stations.

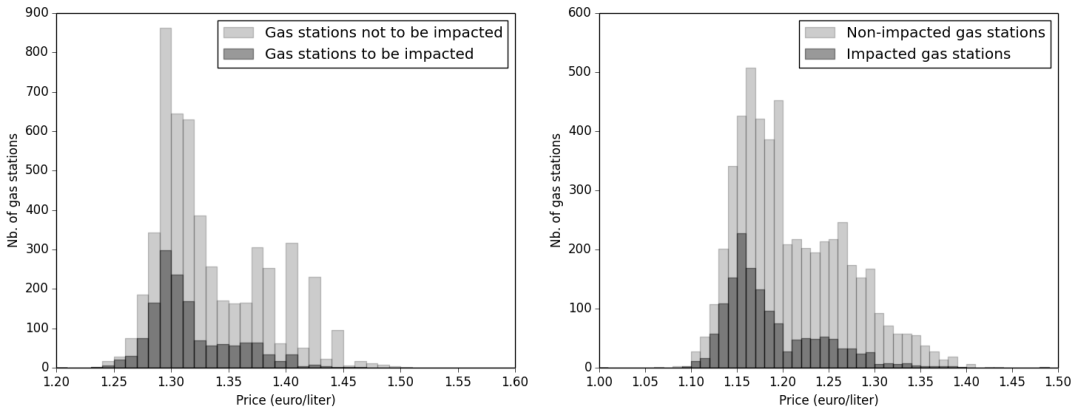
On top of the heterogeneity among retailer types, the location of gas station location is also correlated with prices and characteristics as shown in the descriptive statistics. Table 5 details the impact of the rebranding in Paris region, distinguishing Paris and is closest surroundings from the broader surroundings. A difference in reaction between supermarkets and other retailers is again observed. However, while results in broader Paris surrounding are similar to the ones obtained at the national level, Paris and its closest surroundings exhibit some differences. In these areas, the aggregate effect is an increase of 0.2 euro cents but supermarkets are found to decrease prices by 1.4 euro cents on average. Given the lower penetration of supermarkets, these results suggest that competitive pressure is lower in these areas. They also illustrate the need to

Table 2.4: Aggregate reactions to Total Access rebranding (cents) - Difference in differences regressions

	(1) Price 5 km	(2) Price 5 km	(3) Price 3 km	(4) Price 3 km	(5) Price 1 km	(6) Price 1 km
Total-Total Access competitors	-0.23**	-0.55***	-0.31***	-0.59***	-0.41***	-0.53***
	(0.08)	(0.13)	(0.07)	(0.07)	(0.14)	(0.08)
Oil company		0.40		0.80***		0.87***
(Ref. Supermarkets)		(0.26)		(0.15)		(0.13)
Elf-Total Access competitors	0.17	-0.19	0.16	-0.18***	-0.08	-0.15**
	(0.10)	(0.15)	(0.12)	(0.06)	(0.18)	(0.06)
Oil company		0.25		0.85***		0.85***
(Ref. Supermarkets)		(0.34)		(0.19)		(0.15)
Intercept	133.19***	133.20***	133.34***	133.34***	133.72***	133.72***
	(0.20)	(0.19)	(0.19)	(0.18)	(0.17)	(0.17)
Nb. Obs.	444.338	444.338	360.743	360.743	209.883	209.883
Adj. R2	0.954	0.954	0.954	0.954	0.956	0.956

Reading Note: Columns 1 to 6 provide the results of the regression $P_{it} = \gamma \cdot Treatment_{it} + \mu_i + \eta_{t,c} + \varepsilon_{it}$ with $\eta_{t,c}$ temporal effects, function of a control group c, μ_i gas station fixed effects, and γ the impact of a Total-Total Access or Elf-Total Access conversion on competitor prices (within a 1, 3 or 5 km distance). Treatment (equal to 1 after a nearby conversion has occurred) is crossed with retailer type in columns 2, 4 and 6. The control group is composed by gas stations located within 5 to 10 km of a Total- or Elf-Total Access gas station. Error terms are clustered at the regional level. Significance thresholds *** 1%, ** 5%, * 10%.

Figure 2.4: Distributions of Total Access competitor prices vs .others



(a) September 4, 2011 (b) December 4, 2014

Reading Note: In black, competitors (within 5 km) of gas stations converted from Total to Total Access. In grey, gas stations are not affected by such changes (distance above 5 km).

measure and investigate the heterogeneity of reactions, in the light of local market characteristics.

Estimated individual treatment effects (Table 6) reveal a significant heterogeneity among reactions, and a limited number of large adjustments. Regarding competitors of Total - Total Access gas stations, 10% are estimated to have decreased prices by 1.5

Table 2.5: Aggregate reactions to Total Access rebranding in Paris region (cents) - Difference in differences regressions

	(1) Price 3 km Paris and surroundings	(2) Price 3 km Paris and surroundings	(3) Price 3 km IdF Other	(4) Price 3 km IdF Other
Total-Total Access competitors	0.22***	-1.43***	-0.39***	-0.59***
	(0.05)	(0.12)	(0.04)	(0.04)
Oil company		1.91***		0.53***
(Ref. supermarkets)		(0.11)		(0.07)
Elf-Total Access competitors	0.46***	-0.71***	0.30***	-0.13***
	(0.06)	(0.09)	(0.04)	(0.04)
Oil company		1.43***		1.06***
(Ref. supermarkets)		(0.10)		(0.06)
Intercept	136.01***	136.01***	133.02***	133.02***
	(0.20)	(0.19)	(0.12)	(0.11)
Nb. Obs.	15.120	15.120	32.992	32.992
Adj. R2	0.965	0.966	0.957	0.958

Reading Note: Columns 1 to 4 provide the results of the regression $P_{it} = \gamma \cdot Traitement_{it} + \mu_i + \eta_{t,c} + \varepsilon_{it}$ with $\eta_{t,c}$ temporal effects, function of a control group c, μ_i gas station fixed effects, and γ the impact of a Total-Total Access or Elf-Total Access conversion on competitor prices (within 3 km). Treatment (equal to 1 after a nearby conversion has occurred) is crossed with retailer type in columns 2 and 4. The control group is composed by gas stations located within 5 to 10 km of a Total- or Elf-Total Access gas station. Significance thresholds *** 1%, ** 5%, * 10%.

euro cents per liter or more, and another 10% to have increased prices by a similar amount. The maximum estimated decrease and increase are respectively of 6 and 5 cents. Reacting to a decrease of 10 cents, 90% of gas stations adjust by an amount which is smaller than 2 cents (upward or downward). Examples of price adjustments are provided in Figure 2.5.

The distinction between supermarket gas stations and oil company and independent gas stations appears to be relevant (Table 7), as could be expected. Supermarket gas stations account for 86% of the 298 gas stations which have decreased prices by 1 cent or more, and only 27% of the 215 gas stations which have increase prices by 1 cent or more. Price decreases implemented by supermarkets are generally of low amplitude, as could be expected given the relatively low pre-existing prices and margins. A small number of oil company and independent gas station also decrease prices. Most of the gas stations which operated at a price level close to the competing Total gas station have not implemented a similar price cut in return upon the conversion.

Significant reactions are also observed at the individual gas station level among competitors of rebranded Elf gas stations. Contrary to Total-Total Access conversions, the number of increases exceeds the number of decreases, which could be explained by an increase in differentiation linked to the disappearance of the Elf brand. Several instances of significant decreases nevertheless also suggest that the rebranding, even

Table 2.6: Gas station level price changes (cents) upon or in reaction to Total Access rebranding - Differences-in-differences estimations

	Nb. Obs.	Avg.	Std. Dev.	Min.	Q10	Q50	Q90	Max.
Total-Total Access (T-TA)	325	-9.92	1.83	-14.37	-12.22	-9.88	-7.30	-5.47
T-TA competitors	1332	-0.19	1.27	-6.34	-1.53	-0.31	1.45	5.28
Elf-Total Access (E-TA)	249	-0.04	0.75	-1.97	-0.87	-0.11	1.04	3.01
E-TA competitors	395	0.21	1.51	-13.37	-1.21	0.14	1.67	8.96

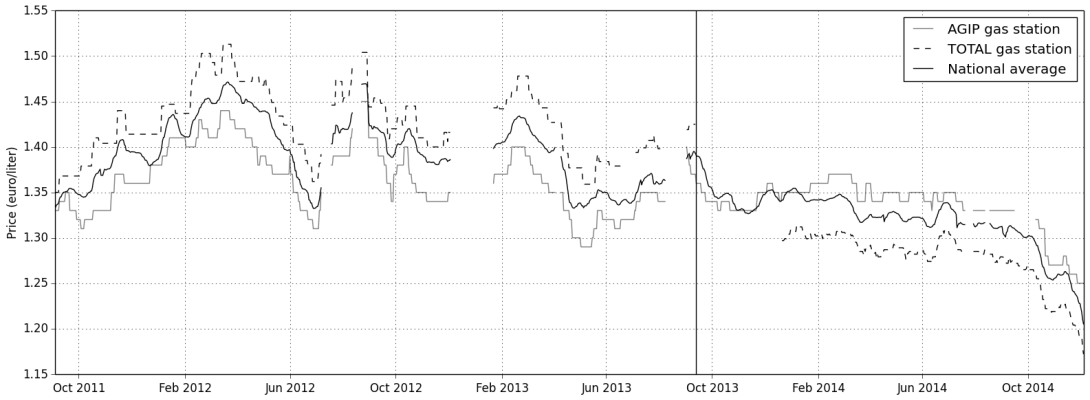
Reading Note: Results come from the regression $P_{it} = \gamma_i \cdot Treatment_{it} + \mu_i + \eta_{t,c} + \varepsilon_{it}$ with $\eta_{t,c}$ temporal effects, function of a control group c, μ_i gas station fixed effects and γ_i the individual treatment effect of a nearby Total-Total Access or Elf-Total Access conversion (within 5km). The control group is composed by gas stations located within 5 to 10 km of a Total- or Elf-Total Access gas station. Significance thresholds *** 1%, ** 5%, * 10%.

Table 2.7: Distributions of Total Access competitor price reactions

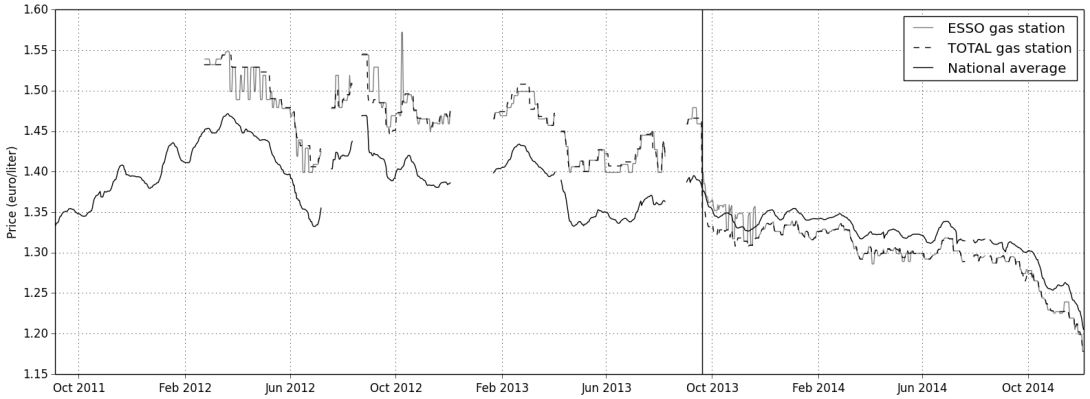
		Price adjustment (euro cents per liter)							
		< - 1	**> 1**	**< - 2**	**> 2**	**< - 3**	**> 3**	**< - 4**	**> 4**
Total-Total Access competitors	Nb. Obs.	298	215	65	57	20	16	10	3
(1332)	%	22	16	5	4	2	1	1	0
Supermarkets	Nb. Obs.	256	57	49	9	12	0	3	0
(952)	%	86	27	75	16	60	0	30	0
Elf-Total Access competitors	Nb. Obs.	57	93	15	24	3	12	3	6
(395)	%	14	24	4	6	1	3	1	2
Supermarkets	Nb. Obs.	52	31	12	2	0	0	0	0
(278)	%	91	33	80	8	0	0	0	0

Reading note: Among the 1,332 competitors of Total-Total Access gas stations (within 5 km), 298 (22%) exhibit a significant price decrease of 1 cent or more. Among these 298 gas stations, 10 are found to cut price by 4 cents or more, and 256 are supermarket gas stations (86%).

Figure 2.5: Examples of price series of Total Access competitors



(a) Price increase



(b) Price decrease

Reading Note: The continuous black line represents the average national price, the dashed black line accounts for a Total Access gas station and the grey line illustrates the prices of one of its competitors. The date of the change in brand is marked by a vertical line.

when it was not accompanied by a shock on prices, may have increased consumer attention and thus the intensity of competition. Consumers may also have perceived the rebranding as a quality improvement.

These results are consistent with the existence of segmentation in the market, between weakly differentiated gas stations addressing a large demand on the one hand, and oil company and independent gas stations focusing on loyal or captive consumers. For these consumers, the creation of Total Access may have actually resulted in an increase in price, depending on the market structure. The robustness of results has been confirmed by numerous specification variations regarding the treatment date (earliest change or closest competitor), the market definition (1, 3 or 5 km), the control group (national or regional), error term correlation, and data frequency (weekly, daily, day of week).

3.3 Discussion

The retail gasoline market has been the topic of a rich literature, motivated at first by questions regarding the reactions of retail prices to variations in crude oil prices. Gas stations were indeed suspected to perform downwards adjustments at a slower pace than upwards adjustments. Access to increasingly large and precise data has allowed to test the various models of oligopolistic competition, legitimated by a strong heterogeneity in observed price patterns. Eckert (2013) has documented the existence of Edgeworth cycles, not caused by wholesale cost fluctuations but related to the market structure. Chandra and Tappata (2011) have provided evidence of significant price dispersion, reflecting a lack of information about prices by consumers. Finally, evaluating the impact of a merger, Houde (2008) has shown the impact of commuter flows, and thus transportation costs, on competition.

The French market differs in a significant way from previously cited markets due to the strong penetration of supermarkets. It is indeed composed, in virtually equal shares, of relatively low and high price price gas stations. The first are mostly operated by supermarkets and offer few amenities, while the second seek to differentiate through services and enjoy more brand awareness (except for few completely independent gas stations). The studied Total-Total Access conversions are accompanied by an 8 to 10 euro cent per liter decrease in price, which essentially corresponds to the difference in price observed between low and high price gas stations. At the time of the creation of Total Access, the Total group has stated that it had selected gas stations for which higher volumes resulting from price decreases were expected to make up for the loss of margin. Given the proximity of supermarket competitors, this leads to think that such gas stations were likely selling to loyal or captive customers, and are supposed, upon rebranding, to benefit from a larger demand, including more price sensitive customers.

There are no significant elements which support the idea that Total Access has effectively reduced its operating costs within rebranded gas stations. It seems therefore reasonable to rather expect a change in demand to account for Total's strategical choice. Consumers can indeed be expected to have become increasingly price sensitive. The three main differentiation dimensions of the market can be reminded to be location, brand and amenities (presence of a shop, car wash facility etc.).

Taking into account the characteristics and dynamics of the French market leads to contemplate several explanations regarding the reactions of competitors of converted gas stations. In a standard Hotelling model with quadratic transportation costs, the price of each seller is a decreasing function of its competitor's price in equilibrium. In our case, as regards supermarkets, the relatively low estimated margins of supermarkets do not allow for large price cuts. The case of oil company and independent gas stations raises mores questions. Capacity constraints related to the number of pumps, or an unfavorable location, may hamper the ability of a gas station to be profitable at

prices set by supermarkets (resp. Total Access) gas stations. An affected gas station may thus be left with a demand composed virtually exclusively by captive or loyal customers, whereas it previously attracted some mildly price sensitive customers. In such a situation, an increase in prices could potentially result from a nearby conversion. In the long run, it seems however unlikely that such gas stations can remain profitable, hence the necessity for further research regarding the impact of Total Access on exits from the market.

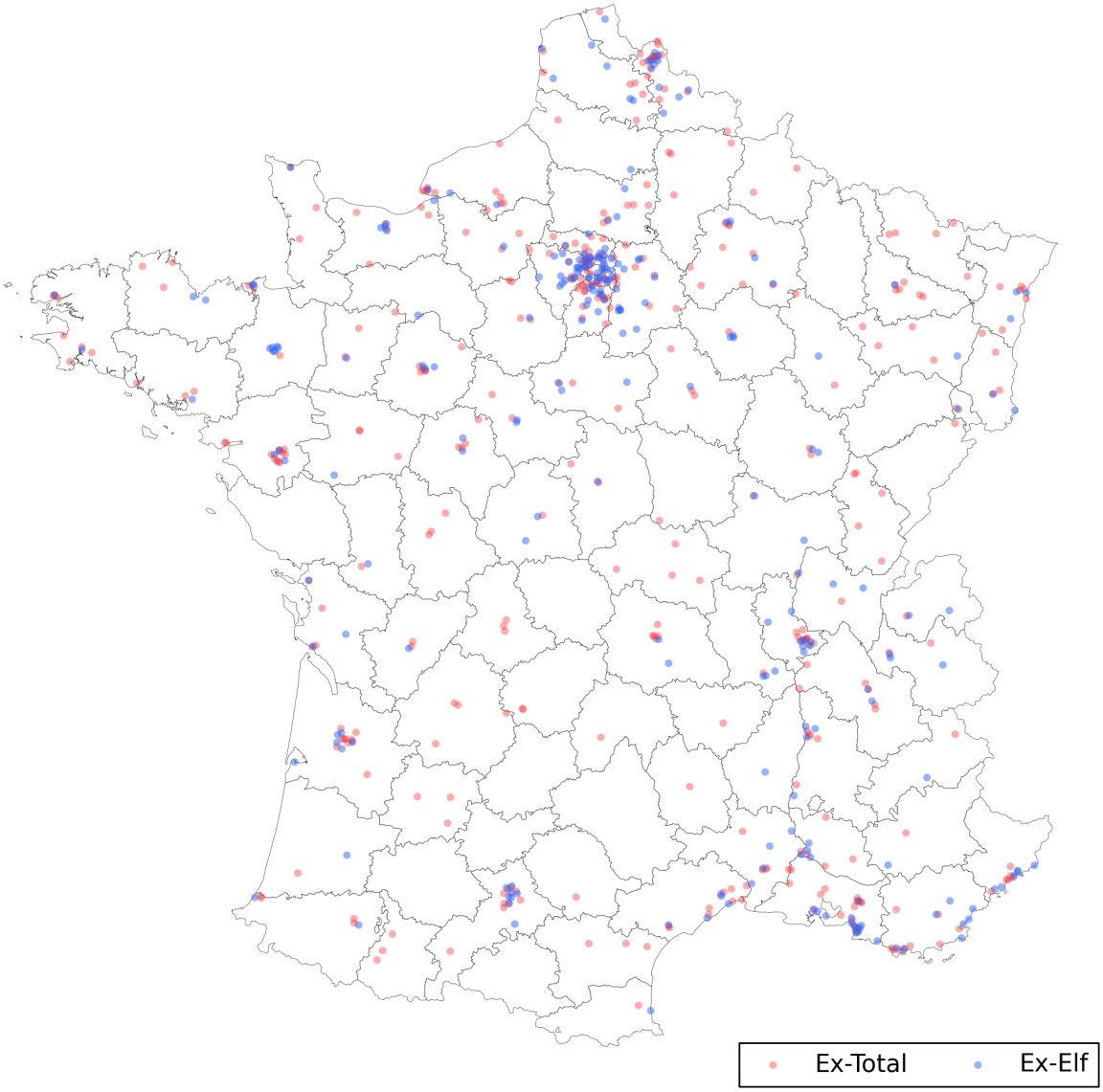
4 Conclusion

The analysis of the impact of the creation of the low price chain Total Access leads to stress heterogeneity in adjustments beyond the absence of a significant reaction at an aggregate level. Findings are consistent with a strong role of market segmentation. A major determinant of the decision by Total to develop its discount offer has been a relative decrease of the demand coming from loyal or captive customers, as a result of fierce competition leveled by developing supermarket gas station networks. Mild consumer reactions can reasonably be accounted for by the low pre-existing margins of supermarket gas stations on the one hand, and the fact that high price gas stations have generally stuck to their policy of focusing on loyal and captive customers. Beyond the period studied in the present paper, the overall share of captive of loyal customers, to which high price gas station cater, can be expected to further decrease. In the long term, the pro competitive effects may be offset by a decreasing gas station density, as a result of exits from non profitable gas stations. In this prospect, it seems desirable to monitor not only prices but also the evolution of the national gas station network.

From a methodological point of a view, the paper leads to stress the limitations of an aggregate analysis of a shock on competition (e.g. merger) across local markets. Effects can indeed be strongly heterogeneous depending on local market structures (e.g. differentiation, demand segmentation), making the outcomes sensitive to the observed sample, as is the case with Paris in our study. Such heterogeneity has already been found to generate discrepancies in evaluations of concentrations in the literature, hence we strongly advocate the use of a flexible individual approach over the measurement of aggregate treatment effects.

5 Appendix

Figure 2.6: Total Access gas stations end of 2014



Chapter 3

Price dispersion on the French retail gasoline market

Abstract. Using a large panel of daily French diesel prices over three years, this paper finds support for a relation between price dispersion and imperfect consumer information. The volatility in price rankings between pairs of competitors is indeed positively correlated with the distance that separates them, namely a measure of consumer search costs. Furthermore, supermarket gas stations, which operate at relatively low markups, are more likely than others to strictly align prices with those of competitors. The study of price dynamics leads to identify a significant number of gas stations acting as leaders or followers in their market. The chain affiliation of sellers largely determines their price strategy. Variations of price dispersion across local markets and over time suggest that a higher diesel cost is associated with lower dispersion, and a higher seller density with increased dispersion for high markup gas stations, namely independent and oil company gas stations. Overall, findings suggest that the latter generally address the needs of customers characterized by relatively strong loyalty or high search costs, while supermarkets compete for a well informed highly price sensitive demand.

Key Words: *Consumer search, price dispersion*

1 Introduction

The understanding of retail gasoline prices has motivated a rich academic literature, marked by suspicions of collusion or insufficient competition. The first widely cited paper of the field, Bacon (1991), tests an hypothesis put forward by a report from the British Monopolies and Mergers Commission. Retailers are indeed suspected of adjusting prices upward faster than downward, a phenomenon commonly referred to as the "rocket and feather" effect. Using a partial adjustment model with fortnightly UK prices between 1982 and 1989, the paper finds support for the existence of asymmetry. Borenstein et al. (1997) makes a significant methodological improvement by introducing the Error Correction Model[1] in the literature and also finds evidence of asymmetry in the US, in particular between city branded terminal prices and retail prices. During many years, most likely for lack of adequate data, microenomic empirical investigations however lag behind so that little is actually known on the intensity of competition between retailers. Hastings (2004), in a context where vertical integration raises increasing concerns, studies a merger in California with gas station level data. The paper finds that independent gas stations foster competition, while an increase in the share of branded gas stations is associated with higher price levels. More closely connected to the "rocket and feather" literature, several papers investigate price dynamics at the city or gas station level and find evidence of Edgeworth cycles, namely short cycles unrelated to cost variations[2]. Remarkably, the diversity of price patterns observed in the gasoline industry questions our ability to understand competition in the field.

Stigler (1961) has initiated a large body of literature by highlighting the link between the "ignorance in the market", namely a lack of consumer information, and price dispersion, which can be defined as the persistence over time of different prices for a homogeneous good. A major theoretical contribution was made by Varian (1980) through the modeling of price dispersion as a result of a Mixed Strategy Equilibrium, namely sellers drawing their price randomly from a distribution in equilibrium. The paper notes that this phenomenon can be interpreted as "temporal" price dispersion, typically comparable to "sales". The imperfection of consumer information relaxes competition so that the famous "law of one price" does not apply and the very same good can actually be purchased at various prices depending on the visited retailer. The development of the Mixed Strategy Equilibrium paradigm has led to learn that consumer search costs and other structural parameters can be uncovered via the observation of an empirical price distribution. In this regard, the relevance of search and price dispersion models has an important practical interest.

[1] Cf. Frey and Manera (2007) for a survey on econometric models of asymmetric price transmission.

[2] Eckert (2002), Eckert (2003), Eckert and Douglas (2004a), Eckert and Douglas (2004b), Noel (2007a), Noel (2007b) and Noel (2008) in Canada, Lewis (2009) and Lewis and Noel (2011) in the US. Cf. Eckert (2013) for a survey on Edgeworth cycles in retail gasoline markets

Using a large panel of daily French diesel prices over three years, this paper quantifies price dispersion and discusses its relation with consumer information issues. A remarkable specificity exhibited by the French market is the presence of large persistent price differences between gas stations, with supermarket and discount gas stations setting prices generally 8 to 10 euro cents per liter cheaper than oil company and independent gas stations.

The paper finds support for a connection between consumer search and price dispersion. Among pairs of competing stations, the volatility in price information over time is indeed found to significantly increase with the distance that separates gas stations, namely a measure of consumer search costs. The existence of such volatility is consistent with price randomization, and its correlation with distance suggests that it reflects information frictions. The study of volatility in competing pair price rankings yet also leads to note that supermarket gas stations, which operate at relatively low markups, frequently align prices with their competitors. Among them, a significant number can be identified as leaders or followers in their market based on the observation of price dynamics. The chain affiliation largely determines the level of prices as well as their stickiness.

These results leads to consider that the relevant market definition, when it comes to studying price dispersion, involves distance as well as chain affiliation. A consumer indeed cannot ignore the information revealed by chain affiliation, and it seems unlikely that the price differences merely reflect variations in offered utility. The heterogeneity of price dispersion across local markets and time suggest that a higher diesel cost is associated with lower dispersion, and a higher seller density with increased dispersion for high markup gas stations, namely independent and oil company gas stations. Overall, findings suggest that the latter generally address the needs of customers characterized by a strong loyalty, high search costs or stringent time constraints. Further research may therefore allow to quantify these constraints via structural estimations.

2 Literature

The retail gasoline market is an interesting candidate when it comes to studying the impact of consumer search costs on competition. Consumers indeed purchase only one relatively homogeneous product and typically face significant costs to remain informed about prices.

Barron et al. (2004) were the first to investigate price dispersion in the retail gasoline market, using a data set of nearly 3,000 gas station prices within four US areas on a single day, in 1997. The non observation of price dynamics implies a limited ability to control for the impact of station-specific characteristics on prices, and the necessity to

consider both static and dynamic theoretical explanations of price dispersion[3]. Under monopolistic competition, price dispersion related to heterogeneity in seller demand or cost should decrease when seller density increases, and so should the average price. Under a search-theoretic approach, the average price can either decrease or increase[4], but seller density and price dispersion should be negatively correlated. This effect can yet be mitigated or reinforced depending whether seller density influences search costs. In particular, Varian (1980) finds that a higher proportion of informed customer can lead to an increase or decrease in the variance of prices, depending on the model's parameters. Barron et al. (2004) measure the density of sellers by the number of gas stations within a 1.5-mile radius around each station. Price dispersion is measured by unexplained variations in prices, namely the squared residuals of the regression of the log of prices on market characteristics, including seller density. An increase in the number of nearby gas stations is found to be associated with a reduction in price dispersion.

Hosken et al. (2008) provide some insights about price dispersion[5] with weekly prices from 272 gas stations around Washington DC between 1997 and 1999. They first regress prices on week time indicators, common to all gas stations, and use the residuals to study the persistence of gas station pricing policies. They then add station fixed-effects to the regression so that residuals reflect deviations from each station's typical price level. Controlling for station fixed-effects accounts for much of the persistence in prices, meaning that a significant amount of dynamic price dispersion is observed once gas station long term pricing policies are taken into account. The data and method employed offer an improvement over Barron et al. (2004) as they shed light on price dynamics which require to go beyond models of static price dispersion. The determinants of observed price dispersion are yet not investigated.

Lewis (2008) reconciles the two previous approaches by using station level fixed-effects to control for differentiation, and investigating the relationship between price dispersion and local market characteristics. Data include price records of 327 gas stations in the San Diego area on each Monday in 2000 and 2001 (91 weeks). The paper finds a negative relationship between seller density and price dispersion, in line with Barron et al. (2004), and refines this result by introducing a distinction between high-brand groups, composed by premium branded stations, and low-brand groups, which

[3] In a theoretical context, price randomization by sellers in a mixed strategy equilibrium is typically interpreted as dynamic price dispersion. Varian (1980), for instance, notes that "It is common to observe retail markets where stores deliberately change their prices over time-that is, where stores have sales". Static price dispersion simply refers to the use of heterogeneous pure price strategies in equilibrium.

[4] It decreases in Carlson and McAfee (1983), in which price dispersion is static, and increases in Varian (1980), which has dynamic price dispersion

[5] They focus on the explanation of gas station mark up levels, the main determinant of which is found to be brand affiliation, and observe many changes in mark up levels on a yearly basis

include discount brand and unbranded stations. The relationship between the density of low-brand sellers and price dispersion is found to be negative, while high-brand sellers have a weakly positive or insignificant impact. Lewis (2008) however observes that a more localized measure of dispersion can lead to find a positive relationship between density and price dispersion, which suggests a complex relationship between seller heterogeneity and price dispersion.

Finally, Chandra and Tappata (2011) make two significant contributions to the literature. Working with US daily gas station prices spanning one year and a half, they introduce a formal test regarding the relationship between price dispersion and consumer search, using distance between competing gas stations as a proxy for consumer information, and then use price dispersion measured at the market level to investigate the relationship between price dispersion and market characteristics[6].

3 Context and data

3.1 The French retail gasoline market

According to the French Union of Petroleum Industries (UFIP), diesel fuel accounted for 81% of total French gasoline consumption in 2013, while its share was only 31% in 1980. Meanwhile, the size of the French gas station network had decreased at a steady pace, from roughly 40,000 sellers in 1980 to nearly 12,000 in 2013. Unlike most other European countries, the French market was characterized by a strong presence of supermarket gas stations. Supermarket chains indeeed accounted for 43% of the total number of gas stations, and over 50% of retail gasoline distribution.

Virtually all gas stations are affiliated to a chain, and a large majority of them belongs to the company that operates the chain. From an operational viewpoint, these are usually either operated by company staff or through a "location-gerance" contract, according to which the manager receives a commission on gasoline sold. For instance, Total SA, the largest gas station operator in France, reported in 2012 that 214 out of its nearly 2,000 Total, Total Access and Elf gas stations freely determined their prices[7]. Industry margins are widely acknowledged to have decreased significantly over the last decade[8], as a result of competition by supermarket chains and increasingly stringent environmental regulations. This has led some oil companies to exit the market (Shell and BP) or to engage in significant divestitures (Esso).

Key cost components are the cost of wholesale gasoline, including delivery fees,

[6] Lewis (2008) uses price residuals as independent variables to investigate the potential determinants of market dispersion, while Chandra and Tappata (2011) regress empirical measures of market dispersion computed from price residuals.

[7] This information was communicated in a context where the government had initiated discussions with retailers aiming at achieving price reductions.

[8] Cf. Bellec et al. (2012)

gas station operating expenses, and taxes. Taxes included a fix part called TICPE, which slightly varies between regions, and the classical Value-Added Tax (19.6% over the period studied, which bear on cost and TICPE). The period studied in the paper is however marked by a temporary tax reduction. On August 29, 2012, following a promise made during the presidential election campaign, the government announced a 3 euro cent per liter tax cut and called on gas station operators to reduce prices by 6 euro cents per liter, through an additional temporary margin reduction. The potential impacts of this shock on price dispersion are discussed in each relevant section.

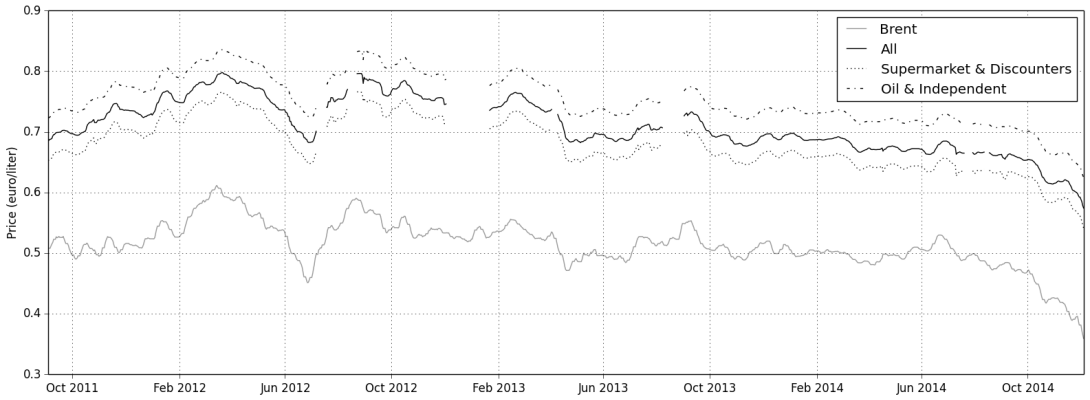
At an aggregate level, two kinds of consumers can be distinguished: businesses and individual customers. Businesses are typically offered card programs which allow them to monitor employee consumption and obtain rebates. An important implication is that the price of the gas station is irrelevant (or only partly relevant) to a significant number of transactions in the market. Individual consumers pay the posted price, and can get information from a variety of sources: at gas stations, on their gps, mobile phone applications (e.g. Zagaz, Carbeo, Essence Free) and on a computer or mobile phone browser (Prix-Carburants.gouv.fr).

Since 2007, French gasoline retailers are required by law[9]to keep prices updated on the governmental price comparison website prix-carburants.gouv.fr. Two other comparison websites, Carbeo and Zagaz[10] were created respectively in 2005 and 2006, relying on user provided information. End of 2014, both had given up their exclusive crowd-sourcing philosophy. Carbeo had started purchasing and using data from the government as early as 2009, while Zagaz resisted until 2014. The governmental website traditionally suffers from significant shortcomings compared to its private counterparts. As of December 2014, the website did not have a mobile application and users were not provided with a way to report errors such as outdated prices or wrong gas station locations. Furthermore, the website did not allow to display rivals of a given gas station on a map, or to access all gas station prices on a given highway. Consequently, it cannot be excluded that the creation of the governmental website may have had an adverse impact on consumer information as it diverted users from promising comparison websites at a time when they crucially needed to grow their user base.

[9]Stations having sold over $500m^3$ gasoline the previous year are exempted from this obligation. According to UFIP, there were nearly 1,500 such gas stations in 2013, while the medium volume sold by a gas station was between 1,000 and $3{,}000m^3$. Given the absence of a national register, the governmental report Bellec et al. (2012) on the French retail gasoline market remarks that "nobody knows precisely the number of gas stations operating in the market".

[10]The governmental body in charge of town and country planning used Zagaz data in 2012 as it looked for the most comprehensive source to study the evolution of the gas station network.

Figure 3.1: Evolution of Brent and average French diesel retail prices excluding taxes



3.2 Data

The governmental comparison website was visited on a daily basis between September 4, 2011 and December 4, 2014, hence a period of nearly 3 years, to collect gas station prices and brand affiliation. Further information collected from the website included the address, the gps coordinates, and gas station amenities such as the presence of a car wash or a shop. Data thereby obtained include 10,180 gas stations, of which 437 were located on highways, 124 on the island of Corsica, and 402 were found to have insufficient or suspicious price data. The analysis was thus performed with a total number of 9,217 gas stations. Gas station municipality was used to enrich the data with socio-demographic variables obtained from the French National Institute of Statistics and Economic Studies (INSEE).

Figure 3.1 provides an overview of the evolution of Brent and French average diesel prices excluding taxes. Discontinuities in diesel price series correspond to short periods of missing price records due to technical data acquisition issues. Variations in aggregate retail prices closely follow the evolution of Brent prices[11]. This close correlation is consistent with the conclusions of Gautier and Le Saout (2015) who use the same source of data to study retail price dynamics between 2007 and 2009. They indeed find that wholesale cost variations are fully transmitted to prices in about 10 days, with no significant upward or downward asymmetry. The difference in prices related to the type of gas stations previously described can be seen to remain fairly stable over time.

Table 3.1 provides an overview of gas station chains and prices on the last day of the period included in the data, namely December 04, 2014. The first column distinguishes three main types of chains: "Oil & Independent" refers to chains operated by oil companies, large intermediaries and independent gas stations, “Oil discount” regroups two low price chains operated by oil companies (thereafter simply referred to as discount

[11] Cf. Lewis (2011) for an illustration of asymmetric gasoline price adjustments in the US.

Table 3.1: Overview of gas station prices by chain on December 04, 2014

Type	Chain	Gas stations		Prices in euro cents			
		Nb	Share	Mean	Std	$\frac{Q75}{Q25} - 1$	$\frac{Q90}{Q10} - 1$
Oil/Independent							
Oil	Total	1 281	15%	1.27	0.03	3%	7%
Oil	Elan (Total)	233	3%	1.32	0.04	5%	8%
Oil	Agip	116	1%	1.25	0.03	2%	6%
Oil	BP	262	3%	1.26	0.04	3%	7%
Oil	Esso	144	2%	1.27	0.05	5%	11%
Independent	Avia	375	4%	1.27	0.05	3%	7%
Independent	Dyneff	55	1%	1.26	0.04	5%	7%
Independent	Other	360	4%	1.24	0.06	7%	12%
Total - Oil/Independent		2 826	32%	1.27	0.04	4%	8%
Oil discount							
Oil discount	Total access	621	7%	1.16	0.02	2%	4%
Oil discount	Esso express	318	4%	1.16	0.02	2%	4%
Total - Oil Discount		939	11%	1.16	0.02	2%	4%
Supermarkets							
Large	Carrefour	200	2%	1.15	0.02	2%	5%
Large	Auchan	118	1%	1.16	0.03	2%	5%
Large	Cora	111	1%	1.18	0.04	4%	8%
Large	Geant Casino	97	1%	1.16	0.02	3%	4%
Large/medium	Intermarche	1 389	16%	1.17	0.03	3%	6%
Large/medium	Systeme U	770	9%	1.16	0.03	3%	6%
Large/medium	Leclerc	585	7%	1.15	0.02	3%	6%
Medium/small	Carrefour market	716	8%	1.18	0.03	3%	5%
Small	Carrefour contact	233	3%	1.20	0.03	3%	5%
Small	Simply (Auchan)	222	3%	1.20	0.03	4%	7%
Small	Casino	200	2%	1.21	0.03	3%	6%
Small	Intermarche contact	112	1%	1.20	0.03	4%	7%
Other	Other	209	2%	1.20	0.04	5%	8%
Total - Supermarkets		4 962	57%	1.17	0.03	4%	7%
Total		**8 727**	**100%**	**1.20**	**0.06**	**7%**	**12%**

Sub-classification of type for supermarkets is meant to reflect what consumers can infer from chain name (as provided on the price comparison website).
Gas stations are considered independent when they are neither operated by a supermarket nor part of a chain operated by an oil company. BP and Esso (including Esso Express) gas stations have an intermediary status: they have been sold to third-party companies with a supply agreement and the right to exploit the brand name.

chains), and "Supermarkets" contains gas stations which are operated by supermarket chains, generally next to a store. Discount and supermarket gas stations set prices significantly lower than these of oil company and independent gas stations, by an average 10 euro cents per liter. Among supermarket gas stations, chain affiliation typically conveys additional information regarding price levels. For instance, the average gas station operated by a Carrefour store (hypermarket), has a diesel price of 1.15 euros per liter vs. 1.18 for a "Carrefour market" (large supermarket) and 1.20 for a "Carrefour contact" (small supermarkets, often located in city centers). The average difference with a Total gas station thus varies depending on exact chain affiliation within the Carrefour group. The last two column indicate that the 50% (80%) of the prices in the middle of the national distribution are within 7% (12%) of each other. Within most chains, the third price quartile does not exceed the first price quartile by more than 3%, which roughly represents a difference of 1.8 euros for 50 liter.

Gas station affiliations are essentially stable over time, except for the creation by Total S.A. of a new discount chain, Total Access, through the rebranding of existing gas stations. The nearly 300 gas stations of the previous discount chain of Total S.A., Elf, are thus converted over the period, as well as a similar number of Total gas stations. For the latter, the rebranding is accompanied by a significant change in pricing policy to roughly align prices with nearby supermarket or discount gas stations. The impact of this operation is studied in a companion paper, Chamayou and Le Saout (2016), which finds that the large price cuts implemented by former Total gas stations turned into Total Access triggered very few and limited reactions from nearby competitors.

Table 3.2 describes gas station pricing behaviors and market environment depending on their type. Since price observations are not available for all gas stations on each day[12], there is an average of 7,895 prices observed per day. Statistics accounting for prices, markups and price variations are obtained by first computing the average for each gas station over time, then taking the average over gas stations. The average oil or independent gas station has an average price of 1.42 euro per liter over the whole period, respectively 11 and 9 euro cent higher than the average discount and supermarket gas station. The average price increase for all gas stations is 1.30 euro cents per liter, slightly lower than its negative counterpart which is 1.47 euro cents per liter. Price changes implemented by discount gas stations are the smallest in value, both average positive and negative variations being smaller than 0.8 cent. These statistics are nearly twice as high for oil and independent gas station. Supermarket price variations are a bit smaller than those of the latter. The daily probability for a gas station to change its price is 18%, which translates in slightly less than a change per week, and an average

[12]Gas station can be closed for maintenance or post prices which are inconsistent and thus dropped from the data. Also, gas stations can cease to report prices if they are no more concerned by the legal obligation due to lower volumes. This means that data do not allow to identify gas stations which are shutting down.

Table 3.2: Station-level summary statistics

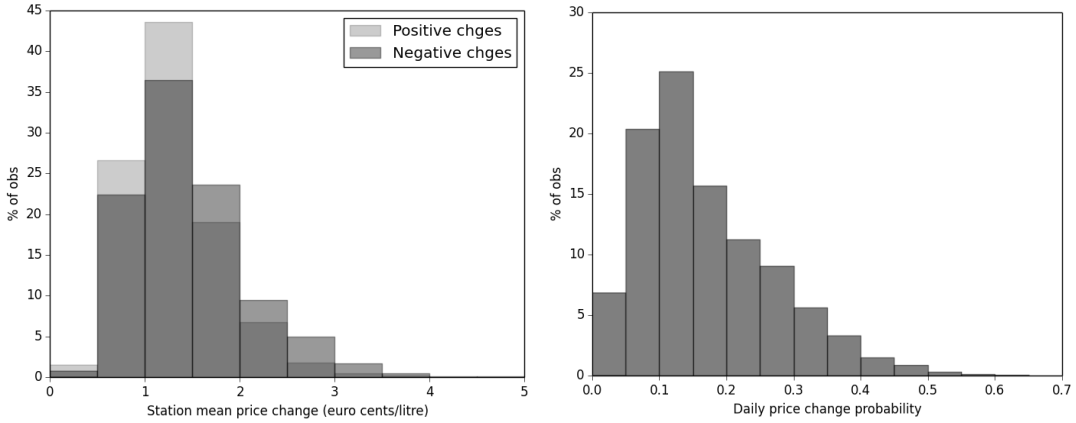
	Oil & Ind	Discount	Supermarkets	All
Nb stations				
All periods	3 177	985	5 055	9 217
Nb daily observations	2 632 (134)	863 (9)	4 400 (55)	7 896 (122)
Price and Markup (euro / litre)				
Price after tax	1.42 (0.04)	1.31 (0.02)	1.33 (0.02)	1.36 (0.05)
Price excl. Tax	0.75 (0.03)	0.66 (0.02)	0.68 (0.02)	0.70 (0.04)
Markup over wholesale cost	0.15 (0.03)	0.07 (0.01)	0.08 (0.01)	0.10 (0.04)
Price changes (euro cent / litre)				
Daily price change probability	0.11 (0.05)	0.29 (0.07)	0.18 (0.10)	0.17 (0.10)
Avg. price increase	1.49 (0.50)	0.74 (0.20)	1.29 (0.45)	1.30 (0.50)
Avg. price decrease	1.67 (0.63)	0.79 (0.22)	1.48 (0.58)	1.47 (0.62)
Rivals				
Nb within 1 km	0.77 (0.94)	0.94 (0.98)	0.55 (0.82)	0.67 (0.89)
Nb within 3 km	3.76 (4.05)	5.18 (3.68)	2.51 (2.92)	3.23 (3.54)
Nb within 5 km	7.92 (9.59)	11.01 (8.77)	4.89 (6.23)	6.59 (8.09)
Distance to closest	1.95 (2.67)	1.08 (1.03)	2.64 (3.18)	2.23 (2.89)

Standard errors in parentheses. Price statistics are obtained by i) computing the average for each station over time ii) taking the average over stations. Costs of transportation and distribution are to be subtracted from the markup, leading to a net margin generally estimated at nearly 1 euro cent per liter.

of 1,500 price changes per day over France. Supermarket gas stations change prices more frequently than oil company and independent gas stations, but less frequently than discount gas stations. Empirical observations on the magnitude and frequency of prices changes by type are consistent with varying degrees of price rigidity, but do not support the presence of Edgeworth cycles on the French retail gasoline market. While Edgeworth cycles are characterized by relatively scarce large price increases and numerous small price decreases, the distributions of positive and negative price variations in euro cents can be observed to have similar shapes on figure 3.2a. Negative variations actually tend to have somewhat larger values. Figure 3.2b provides an overview of the large heterogeneity in the frequency of price changes among gas stations.

Market characteristics and pricing behaviors are also correlated with gas station location. Retailers in Paris metropolitan area tend to change prices more frequently, with smaller variations, and are closer to their closest rival than others. Chamayou and Le Saout (2016) perform regressions of gas station prices on variables accounting for socio-demographic factors, competition intensity, gas station amenities and chain affiliation. Findings are consistent with Hosken et al. (2008) who find that chain affiliation is a very strong price predictor, while coefficients of variables accounting for competition tend to be insignificant or negligible in value. On the French market, this result is fairly unsurprising given the large differences in prices despite the fact than many supermarkets are located in the close vicinity of oil and independent gas stations. For instance, more than half non supermarket gas stations are located within

Figure 3.2: Histograms of price change frequencies and values



(a) Station average price variations

(b) Station daily price change probabilities

1.4 km from a gas station operated by a supermarket. These findings lead to suspect the existence of a significant market segmentation, reflecting a strong heterogeneity in consumer preferences or constraints. The implications of such a potential segmentation on competition are discussed further in the paper.

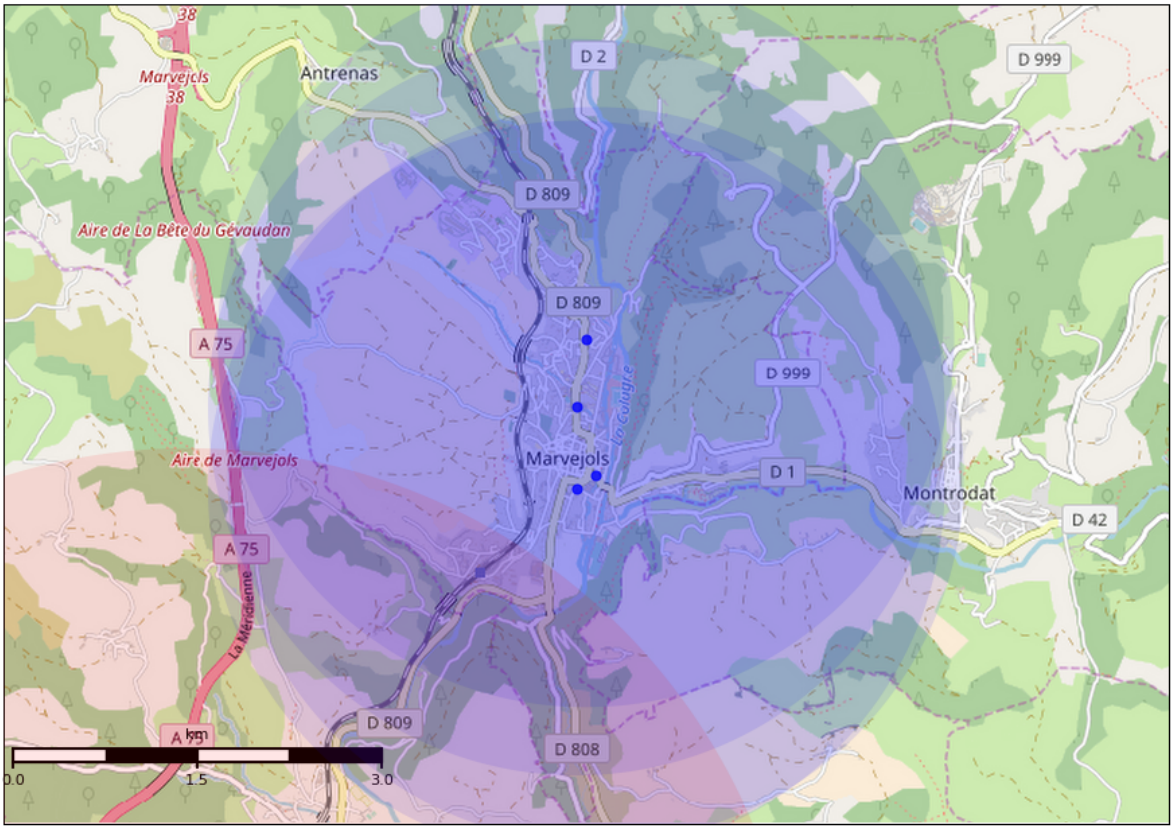
3.3 Competition definition

A crucial issue when it comes to studying competition between gas stations is the definition of catchment areas. It has been traditionally addressed in the literature by considering circles of various radiuses around gas stations, usually measuring distance as the crow flies. Barron et al. (2004) use a radius of 1.5 miles (2.4 km) to study dispersion within four urban areas in California and Arizona. The average number of rivals varies from 8.3 to 10.6 depending on the area, while the distance to the closest rival respectively fluctuates from 0.3 to 0.5 miles (0.5 to 0.8 km). Chandra and Tappata (2011) analyses price dispersion across California, Florida, New Jersey and Texas using distances of 1 and 2 miles (respectively 1.6 and 3.2 km). Depending on the state, the average number of rivals varies from 4.3 to 4.6 with 1 mile and from 13.5 to 14.2 with 2 miles. Houde (2008), uses travel times of 0.5, 1 and 1.5 minutes to study a merger in Quebec city. In the Netherlands, Bruzikas and Soetevent (2015) use distances of 2 and 5 km to study the impact of conversions of manned to unmanned gas stations on prices. Pennerstorfer et al. (2015), in Austria, combine a driving distance of 2 miles and commuting patterns to build markets. Though travel distance or time can be seen as an improvement over distance as the crow flies, their merits must to be kept in perspective. On small distances, they can magnify the usual inaccuracies in retailer locations. On larger distances, they are more likely to make a difference, but still provide largely incomplete information absent data on commuting patterns or driving

habits. Our analysis adopts a distance as the crow flies of 3 km as a baseline to define competition, and discusses the use of additional criteria such as the type or the price level of gas stations.

Table 3.2 shows that supermarkets tend to have less nearby competitors than gas stations of other types. Within 3 km, the average supermarket has 2.5 competitors, while oil and independent gas stations have 3.8 rivals. Discount gas stations appear to be even more exposed with 5.2 competitors within 3km. As can be expected, the average distance to the closest rival follows an inverse order: 1.1 km for discount, 2.0 km for oil and independent and 2.6 km for supermarkets.

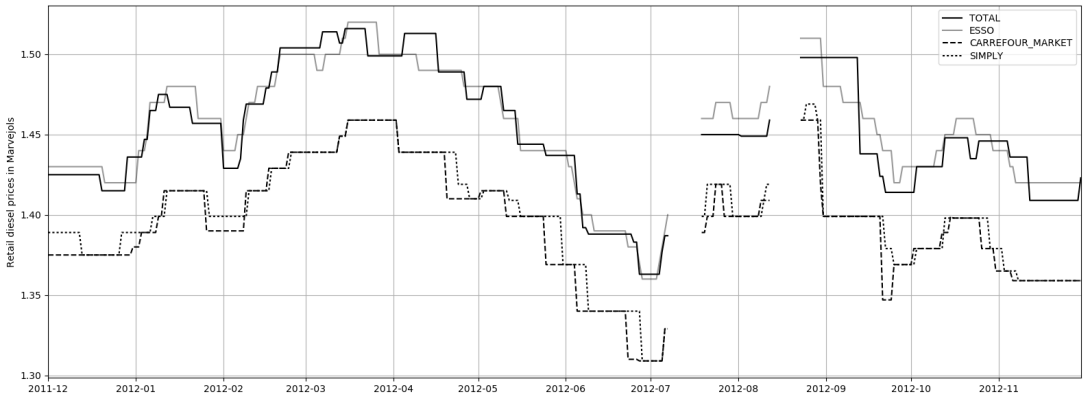
Figure 3.3: Illustration of local competition: Marvejols



Background map from OpenStreetMap. Gas stations in the municipality of Marvejols (Lozère)

The first part of the paper focuses on pairs of competing gas stations. More precisely, we consider all pairs of gas stations separated by a distance of less than 3 km and which are not operated by the same company. Figure 3.3 provides the example of small municipality with four gas stations. Blue circles of 3 km radius are drawn around each gas station. Based on our definition of competition, all gas stations in the city compete with each other, which yields 6 pairs of competing gas stations. The part of the red

Figure 3.4: Retail diesel prices in Marvejols between November 2011 and November 2012



circle in the lower left corner belongs to a 5km radius circle drawn around the gas station which is the closest to any of the gas stations in Marvejols. As a consequence, they are considered not to compete with gas stations outside of the municipality, even if we consider a distance of 5 km to define competition. The second part of the paper investigates price dispersion at the market level. Following a procedure previously used in the literature, we start by considering each gas station as the center of a market defined by a 3 km radius. In the case of Marvejols, this leads to define 4 markets which all contain the same outlets. We thus describe a simple procedure in section which allows to identify well defined markets such as the city of Marvejols.

Figure 3.4 displays diesel prices in Marvejols between November 2011 and November 2012. Gas stations located the farthest north and south respectively belong to the Total and Esso chains. Their price levels can be seen to be relatively similar, and consistently higher than those of the Carrefour Market and Simply gas stations, which are both operated together with a supermarket. Prices at Carrefour Market and Simply are frequently perfectly aligned, while this is rarely the case for Total and Esso prices.

4 Rank reversals, consumer information and leadership

The goal of this section is to evaluate whether price dispersion a la Varian (1980) matches with price patterns observed in the French retail gasoline market.

4.1 Rank reversals

Randomization of prices in Varian (1980) means that a seller is cheaper or more expensive than a competitor with equal probabilities. The existence of such volatility in price rankings can be directly verified in the data. Furthermore, Chandra and Tappata

(2011) suggest that we can test whether price dispersion and consumer search are connected based on the assumption that distance between sellers can be used as a proxy for search costs. The logic is that when a gas stations operates in the close vicinity of another, a higher share of consumers is likely to perfectly observe both prices than when sellers are separated by a higher distance. If the share of uninformed consumers is negligible, sellers can be expected to compete a la Bertrand, or Hotelling, and retail prices should essentially match wholesale price fluctuations. Conversely, if consumers are largely uninformed, dynamic dispersion can arise following the intuition exposed in Varian (1980). Chandra and Tappata (2011) measure temporal price dispersion between two stations as the probability that the gas station which is in general cheaper, in terms of day count, turns out to be more expensive. Formally, denoting two gas stations i and j, and $t \in T$ each day on which price records p_{it} and p_{jt} are available, the rank reversals statistic r_{ij}[13]writes:

$$r_{ij} = \min \left\{ \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}_{p_{it} > p_{jt}}; \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}_{p_{jt} > p_{it}} \right\}$$

Rank reversals can reach a maximum value of 50% when gas station i is strictly cheaper than its competitor j during half of the period, and strictly more expensive during the other half. If one seller is always more expensive than the other, or both always set the same price, r_{ij} takes value 0. In order to account for price alignment, we compute the percentage of days on which both sellers set the very same price:

$$sp_{ij} = \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}_{p_{it} = p_{jt}}$$

Figure 3.5 provides an example of two gas stations displaying large rank reversals. The Esso gas station is indeed successively more expensive and cheaper than its BP competitor between January and July 2014. The standard deviation of the daily price spread can be used as a proxy for dispersion that is robust to persistent price differences[14]. Standard deviation is however more sensitive to outliers that can result from extraordinary promotions or erroneous price records. Formally, denoting $s_{ijt} = p_{it} - p_{jt}$ the price spread between gas station i and j on day t, and its average $\bar{s}_{ij} = \frac{1}{T} \sum_{t=1}^{T} s_{ijt}$,

[13] All the statistics we use to account for price dispersion between pairs of gas stations are symmetric (in particular $r_{ij} = r_{ji}$).

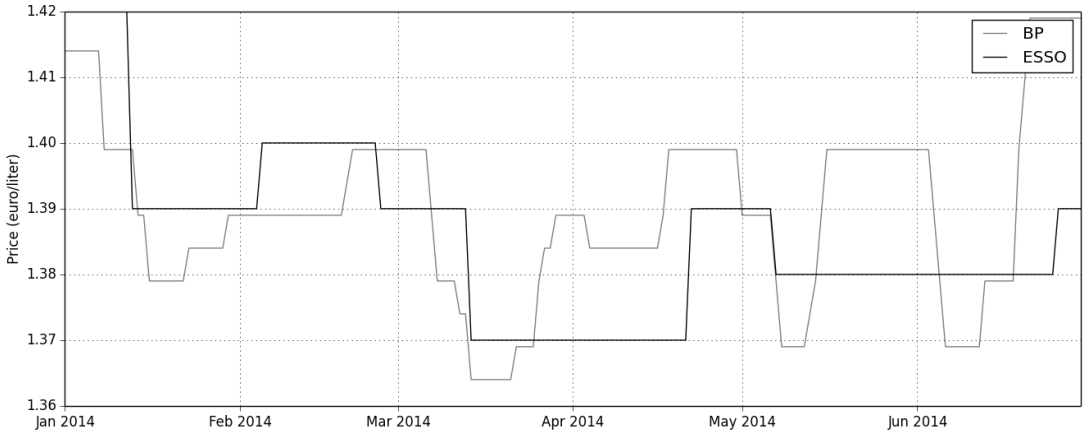
[14] Similarly as in Chandra and Tappata (2011), we only perform this analysis with raw prices as rank reversals measured from price residuals tend to be highly sensitive to the process employed to obtain price residuals.

the standard deviation over the period is computed as follows:

$$\sigma_{ij} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[s_{ijt} - \bar{s}_{ij}]^2}$$

We refer to $|\bar{s}_{ij}|$ as the long term average price difference or the mean price spread between two gas stations[15].

Figure 3.5: Price series of two competing gas stations



The formal test proposed by Chandra and Tappata (2011) consists in regressing measures of price dispersion on a dummy variable which identifies competitor pairs for which a low separating distance implies low search costs. Denoting Pair dispersion$_{ij}$ the price dispersion indicator between two gas stations i and j, $\mathbb{1}[\text{Short distance}]_{ij}$ an indicator for whether the stations are separated by a short distance, and X_{ij} a set of control variables[16], the estimated equation:

$$\text{Pair dispersion}_{ij} = \beta_0 + \beta_1 \mathbb{1}[\text{Short distance}]_{ij} + \beta_2 X_{ij} + \epsilon_{ij} \tag{3.1}$$

Price dispersion, measured trough rank reversals or the standard deviation in prices, is expected to be lower when the separating distance is shorter, since information frictions are reduced. In the case where information is perfect, prices should be perfectly aligned as a result of Bertrand competition.

[15] This indicator is also symmetric: $|\bar{s}_{ij}| = |\bar{s}_{ji}|$.
[16] Accordingly, β_2 is a vector of coefficients.

4.2 Persistent price differences and spurious rank reversals

Standard measures of market price dispersion are the range, namely the difference between the maximum and the minimum price, and the standard deviation of prices[17]. As regards the French retail gasoline market, a naive use of these statistics is unlikely to be informative about search related price dispersion. Indeed, Total or BP gas stations are commonly expected to be much more expensive than a Leclerc or Carrefour gas station. The treatment of persistent price differences is a sensitive issue in the analysis of dispersion. It has been frequently addressed in the literature by working with price residuals, namely after controlling for seller (and occasionally time) fixed effects. However, beyond the practical errors inherent to the statistical treatment, this approach is only supported by theory to the extent that static price dispersion mirrors heterogeneity in offered utility[18]. In our case, the large differences in prices and the relatively low predictive power of seller characteristics suggest that this represents a very strong assumption. Therefore, we consider an alternative hypothesis according to which a large persistent price difference between two retailers implies that they do not cater to the same customers.

Two sources of large rank reversals unrelated to consumer information are identified in the data. One is the conversion by Total S.A. of nearly 300 Total gas stations to its discount brand, Total Access, accompanied by a sharp change in the pricing policy. Converted gas stations decrease prices by approximately 10 euro cents per liter, which has been found to trigger very moderate adjustments by competitors (Chamayou and Le Saout (2016)). A converted Total gas station is thus often consistently more expensive than a competitor at the beginning of the period and becomes cheaper once the conversion has occurred. Figure 3.8, in appendix, provides such an example, for which a naive measure of rank reversals over the whole period is equal to 47%. This issue is addressed by dropping all pairs of competitors which involve a Total gas stations converted to Total Access from our analysis[19]. Another observed source of rank reversals is the occasional use by a few gas stations of dynamic price discrimination. The latter is detected by the regularity of successive inverse price changes at the gas station level. The most commonly observed pattern is a surge in prices during weekends. Such a pricing policy, whenever it is implemented unilaterally by a gas station in one market, can generate rank reversals which are unrelated to the use of mixed strategies. Figure 3.9, in appendix, provides an illustration of the phenomenon. However, as very

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

[18] Cf. Wildenbeest (2011).

[19] More generally, a similar phenomenon may arise whenever a shock affects a gas station differently from the way it affects its nearby competitors. The robustness of our results is tested i) by dropping gas stations which are found to implement a significant change in pricing policy over the period ii) by dropping pairs which exhibit high rank reversals but the price series of which scarcely cross each other. Details are available upon author request

few observations are affected, dropping pairs of competitors which involve a gas station implementing dynamic price discrimination does not affect our results.

4.3 Measured rank reversals

For each pair of competing gas stations separated by a maximum distance of 3 km, we compute the percentages of rank reversals, same price, as well as the mean price spread and its standard deviation. This yields a database of 11,754 observations which are described in Table 3.3[20]. The average rank reversals is 7.1% which is arguably small. Considering only pairs in which both retailers are either independent or operated by an oil company, the average rank reversals increases to 10.5%. Within this sample, if we focus on pairs which exhibit a maximum mean price spread of 2 cents per liter, we are left with 646 pairs which exhibit an average 22.2% rank reversals. These pairs can be noted to set the very same price on average 8.5% of the time, while pairs of supermarket rivals (under the same spread restriction) respectively display a 32.2% average probability to set the same price on a given day, and 14.6% average rank reversals. Given the strong product homogeneity and the widely admitted competitive nature of the market, a reasonable baseline hypothesis for the strong price alignment displayed by many supermarket pairs is that involved gas stations compete à la Betrand. By contrast, pairs of high markup retailers, namely oil company and independent gas stations, often exhibit significant rank reversals, which make them good candidates for a theoretical representation in terms of MSE. By comparison, Chandra and Tappata (2011) report average rank reversals ranging from 11.9% to 14.9% with a maximum separating distance of 1 mile, and from 12.3% to 15.9% with 2 miles[21].

Figure 3.10, in appendix, displays the percentage of pairs of gas stations whose price rank is reversed on each day of the period studied. Among pairs of gas stations built with a maximum distance of 3 km, the percentage of reversed pairs fluctuates between 5.4% and 15.3% (mean 8.1%). The "No differentiation" series results from a focus on pairs which exhibit an average price difference below 2c/l. Among these pairs, the minimum percentage of reversed pairs fluctuates between 13.8% and 29.3% (mean 19.5%). From a consumer viewpoint, this translates in one chance in five to pay the highest price upon patronizing among two competitors of similar price level the one which is cheaper most of the time.

[20] Cf. Table 3.10 in Appendix for replications with a distance of 5 km

[21] They do not introduce restrictions on type or average price spread.

Table 3.3: Overview of pairs price dispersion (distance $\leq$ 3 km)

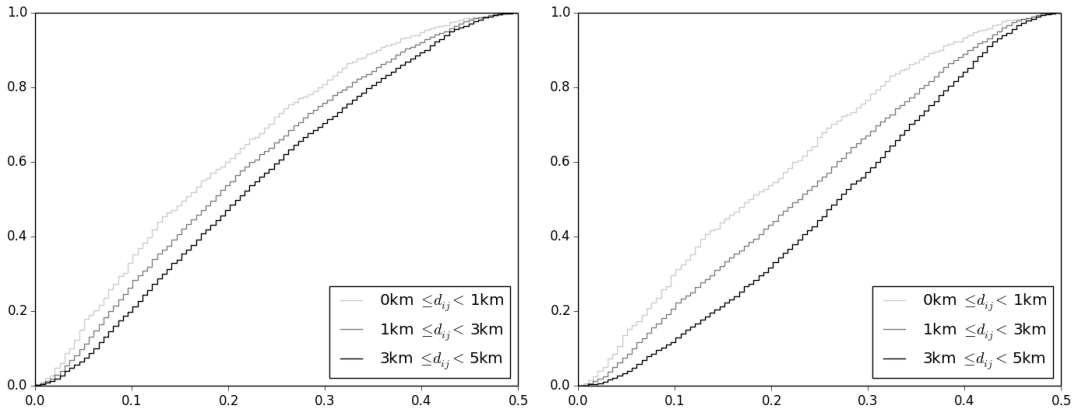
	Nb pairs	Rank reversals	Same price	Price spread Mean	Price spread Std.
No price spread restriction					
All types	11 754	7.1 (10.8)	7.8 (17.0)	5.3 (4.3)	1.6 (0.6)
- Oil & Ind	1 679	10.5 (12.0)	3.9 (9.0)	3.2 (2.6)	1.9 (0.6)
- Supermarkets & Discounters	3 706	15.4 (11.3)	21.7 (23.6)	1.0 (1.0)	1.3 (0.5)
- Supermarkets	2 232	12.9 (10.3)	27.3 (27.6)	1.0 (1.1)	1.3 (0.5)
- Discounters	157	28.2 (9.5)	13.6 (6.5)	0.4 (0.4)	1.1 (0.3)
- Supermarkets vs. discounters	1 317	18.1 (11.4)	13.1 (11.7)	1.0 (0.9)	1.3 (0.5)
Price spread <= 2 cent per liter					
All types	4 171	18.1 (11.3)	21.3 (23.0)	0.7 (0.6)	1.3 (0.5)
- Oil & Ind	646	22.2 (10.7)	8.5 (13.1)	1.1 (0.5)	1.7 (0.4)
- Supermarkets & Discounters	3 176	17.1 (11.1)	24.9 (24.0)	0.6 (0.5)	1.2 (0.4)
- Supermarkets	1 867	14.6 (10.3)	32.2 (27.6)	0.6 (0.5)	1.2 (0.5)
- Discounters	156	28.3 (9.4)	13.7 (6.5)	0.4 (0.4)	1.1 (0.3)
- Supermarkets vs. discounters	1 153	19.8 (11.0)	14.6 (11.7)	0.7 (0.5)	1.2 (0.4)
Price spread <= 1 cent per liter					
All types	2 928	20.3 (11.8)	27.4 (24.6)	0.4 (0.3)	1.1 (0.4)
- Oil & Ind	297	29.0 (10.5)	12.1 (17.0)	0.5 (0.3)	1.6 (0.4)
- Supermarkets & Discounters	2 423	19.0 (11.5)	30.2 (24.9)	0.4 (0.3)	1.1 (0.4)
- Supermarkets	1 404	15.9 (10.8)	40.0 (17.3)	0.3 (0.3)	1.1 (0.4)
- Discounters	140	30.0 (8.2)	14.4 (6.4)	0.3 (0.2)	1.0 (0.3)
- Supermarkets vs. discounters	879	22.2 (10.9)	17.1 (11.9)	0.5 (0.3)	1.1 (0.3)

Except for the first column, all figures are averages of statistics computed at the pair level (standard errors in parentheses). "Mean Price spread" refers to $|\bar{s}_{ij}|$ and accounts for the existence of persistent price differences (0 if both gas stations are equally expensive over the long term), while the standard deviation, "Std. Price spread", measures dispersion around the long term price difference (0 if both gas stations always set the same price). "Rank reversals" and "Same price" are percentages.

4.4 Test of the relation with consumer information

A clear ranking of empirical distribution functions of rank reversals can be observed among pairs of gas stations depending on distances in Figure 3.6. This is consistent with the idea that nearby gas stations compete in a virtually complete information setting, in which there is no reason to expect rank reversals. Conversely, distance can create an information issue for other pairs, preventing the existence of an equilibrium in pure strategy.

Figure 3.6: Empirical distribution functions of rank reversals



(a) Pair average price spread $\leq$ 2 cent/litre (b) Pair average price spread $\leq$ 1 cent/litre

Table 3.11 reports the estimations of β_1 obtained with all competitor pairs as well as subsamples based on gas station types, using both rank reversals and standard deviation in spread as proxies for pair price dispersion. Overall, pairs of gas stations which are separated by a distance of 1 km or less exhibit rank reversals which are 2.3 points smaller than pairs which are separated by a distance comprised between 1 and 3 km. This result is close to the 2.7 point difference found by Chandra and Tappata (2011). The average effect is stronger for pairs of oil and independent gas stations, and smaller for supermarkets. For the latter, no significant impact is found when standard deviation in spread is used to account for price dispersion. Considering the relative strong price alignment displayed by supermarket pairs, this leads to put into perspective the impact of consumer search cost on their markets[22].

Overall, empirical findings provide support for the existence of price dispersion related to consumer search, but we also observe that most low markup gas stations tend to largely align prices with those of nearby competitors. While this could simply reflect almost perfect consumer information, one must take into account the fact that

[22]Theoretically, rank reversals with strong price alignment can be obtained in an extension of Varian (1980) that requires sellers to choose prices from a grid. Occasional rank reversals could also simply result from temporary frictions linked to wholesale cost variations, or short periods of price wars.

Table 3.4: Regressions of pair price dispersion (average price spread $\leq$ 1 euro cent per liter)

	Dependent Variable	OLS	Quantile regressions 25%	50%	75%	90%
All pairs	r_{ij}	-2.28***	-1.70***	-3.60***	-2.70***	-2.10***
(N = 2 928)		(0.51)	(0.66)	(0.86)	(0.80)	(0.81)
	σ_{ij}	-0.04**	-0.05***	-0.03	-0.02	-0.01
		(0.02)	(0.02)	(0.02)	(0.03)	(0.04)
Oil & Ind	r_{ij}	-3.29**	-3.80**	-3.00*	-1.90	-0.50
(N = 297)		(1.34)	(1.82)	(1.74)	(1.84)	(1.84)
	σ_{ij}	-0.22***	-0.25***	-0.21***	-0.17*	-0.10
		(0.05)	(0.06)	(0.06)	(0.09)	(0.12)
Supermarkets	r_{ij}	-2.12***	-1.40**	-2.50***	-2.30	-1.90
(N = 1 404)		(0.68)	(0.60)	(0.91)	(1.41)	(1.60)
	σ_{ij}	-0.01	-0.02	-0.01	0.04	-0.02
		(0.02)	(0.03)	(0.03)	(0.04)	(0.05)
Discounters	r_{ij}	-2.29	-2.80	-1.10	-3.10	-2.80
(N = 140)		(1.69)	(2.96)	(2.44)	(2.15)	(2.23)
	σ_{ij}	-0.28***	-0.19***	-0.24***	-0.38***	-0.52***
		(0.06)	(0.07)	(0.08)	(0.08)	(0.10)
Supermarkets vs. Discounters	r_{ij}	-2.70***	-3.40**	-2.80**	-2.60**	-4.30***
(N = 879)		(0.91)	(1.49)	(1.34)	(1.32)	(1.31)
	σ_{ij}	-0.02	-0.05*	-0.03	0.00	0.02
		(0.03)	(0.03)	(0.03)	(0.04)	(0.07)

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

supermarkets may use gasoline to attract consumers in stores, thus setting prices in a way that does not purely reflect fundamental local market characteristics. We discuss this question through an analysis of price alignment and chain affiliation.

4.5 Leadership: the importance of chain affiliation

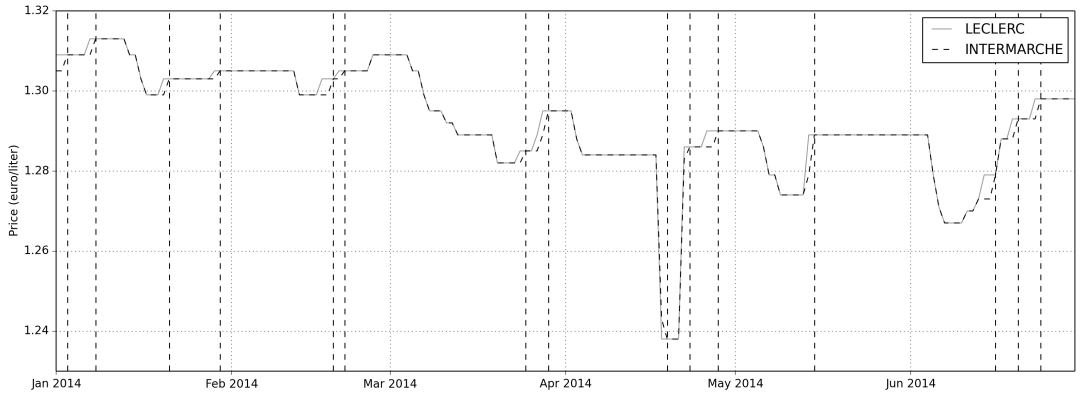
Several papers in the literature have emphasized how chain strategies can shape market prices and dynamics. Hosken et al. (2008) remark that gas stations of a specific chain nearly always match the lowest price in their local market. Lewis (2012), working on markets which exhibit Edgeworth cycles, show that gas stations of two chains exhibit a much higher probability of jumping on the first day of the restoration than most of their competitors. They also show that cycles tend to be absent from almost every city in which these two chains have little or no market presence.

In order to capture the potential existence of leadership in price changes, we count, within each pair, the number of times each seller matches the price of its competitor. Formally, the number of times seller i matches the price of seller j writes:

$$\sum_{t=2}^{T} \mathbb{1}_{p_{it}=p_{jt} \text{ with } p_{it-1}\neq p_{jt-1} \text{ and } p_{jt}=p_{jt-1}}$$

We then test the hypothesis that each gas station is equally likely to match its competitor's price through a binomial test of equality of the two statistics[23]. Whenever the equality hypothesis is rejected, the gas station which displays the highest figure is considered as a leader, and the other as a follower[24]. Regarding pairs for which the equality hypothesis is not rejected, we want to consider only pairs which actually often set the same price. We thus drop pairs which do not meet a threshold regarding the percentage of days on which both gas stations set the same price. Figure 3.7 displays the price series of a pair in which leadership is detected. The Intermarche gas station can be seen to frequently match the price of a Leclerc competitor. Each vertical dashed line corresponds to one such adjustment.

Figure 3.7: Example of two gas stations identified as leader and follower



The Intermarche gas station can be seen to frequently match the price of its Leclerc competitor. Each vertical dashed line corresponds to one such adjustment.

Results are aggregated for each gas station in order to distinguish those which have at least one competitor with similar prices, and among those to evaluate which ones can be identified as leaders or followers in their market. A gas station is labeled as a relative leader, respectively follower, when it is found to lead, respectively follow, in at least one pair. Furthermore, all gas stations belonging to pairs in which no leader can be identified are labeled as uncertain. Importantly, these three labels are not mutually exclusive: a gas station can identified as a leader in one pair, a follower in another, and neither as a leader nor a follower in a third. Finally, an absolute leader is defined as a seller which belongs to the set of relative leaders but not to the sets of relative followers and uncertain. Similarly, an absolute follower is defined as a seller which belongs to the set of relative followers but not to the sets of relative leaders

[23] Cf. Seaton and Waterson (2013) for a similar approach with British supermarket prices.

[24] Sellers are often observed to simultaneously adopt the same price on a given day. Such observations are not taken into account in the test.

Table 3.5: Overview of leadership test results by chain

Type	Chain	Nb stations	Similar prices	Leader Rel.	Leader St.	Follower Rel.	Follower St.	Uncertain Rel.	Uncertain St.
Oil/Independent									
Oil	Total	1 365	7	1	1	1	1	4	4
Oil	Elan (Total)	279	5	1	1	0	0	4	4
Oil	Agip	132	14	1	1	6	5	9	8
Oil	BP	290	7	1	1	1	1	6	5
Oil	Esso	156	15	3	2	2	2	12	10
Independent	Avia	403	18	2	2	5	3	13	11
Independent	Dyneff	64	14	5	0	5	2	11	6
Independent	Other	390	27	6	3	5	3	22	17
Oil discount									
Oil discount	Total access	319	34	19	16	8	6	12	8
Oil discount	Esso express	326	63	6	2	45	29	31	15
Supermarkets									
Large	Carrefour	198	94	37	1	73	19	62	12
Large	Auchan	119	83	63	24	29	6	49	8
Large	Cora	115	43	24	9	14	3	30	10
Large	Geant Casino	97	96	34	6	81	19	63	5
Large/medium	Intermarche	1 392	43	19	12	16	7	22	11
Large/medium	Systeme U	773	47	19	9	19	9	28	14
Large/medium	Leclerc	591	75	58	35	9	3	35	12
Medium	Carrefour market	704	62	8	2	47	31	26	11
Small	Carrefor contact	236	11	2	1	2	1	9	8
Small	Simply (Auchan)	226	24	10	7	6	4	13	8
Small	Casino	206	24	4	2	15	12	9	7
Small	Intermarche contact	116	24	7	5	5	4	15	12
Other	Other	224	29	6	3	9	6	21	14

and uncertain. Gas stations which exclusively belong to pairs for which no leader is identified are labeled as absolutely uncertain. The daily frequency of the data imposes a limitation on leadership detection. Many pairs which are labeled uncertain by the present analysis may be found to involve leadership if the exact time of price changes was taken into account. Unfortunately, such inaccuracies tend to propagate across observations. When a gas station A matches the prices of a competitor B so closely that no leader is identified, if a seller C loosely follows B, both A and B are labeled as relative leaders in their comparison with C. Results analyzed at the chain level however suggest that the present analysis, despite its current limitations, conveys meaningful information.

Table 3.5 reports results aggregated at the chain level. The column "Similar prices" provides the percentage of gas stations within each chain which have at least one competitor exhibiting similar prices, namely with which they set the same price more than 30% of the time. The "Relative Leader" column provides the percentage of gas stations within each chain which are identified as relative leaders in at least one pair. The "Strict Leader" column provides the percentage of gas stations within each chain which are identified as strict leaders in their market. Oil company and independent gas stations are confirmed to be relatively less likely than others to align prices with

competitors. Conversely, many gas stations operated by large supermarkets have competitors which frequently adopt the same prices. For instance, respectively 73% and 58% of gas stations operated by Carrefour and Leclerc[25], are found to compete with at least one seller with which they share the same price 30% of the time or or more. The same is true for oil discounters, while smaller supermarket chains exhibit figures which are significantly lower, yet remaining higher than those of oil company gas stations. Leclerc, which is widely acknowledged to be the most aggressive large retailer regarding prices, is by far the chain under which the share of absolute leaders is the highest (35%). The only other chain to exhibit a share of absolute leaders above 10% is Auchan (23%). Carrefour gas stations, on the other hand, appear much more likely to act as followers (26% are labelled absolute followers for large supermarkets, and respectively 29% for medium size supermarkets), as are Geant Casino stores (33% absolute followers). Results for Systeme U and Intermarche are more balanced, which is likely reflects the strong heterogeneity in store formats. These two chains also differ from Auchan, Carrefour and Geant by being franchises, as opposed to integrated groups. Finally, results suggest that Total Access are much more likely to lead, conversely to Esso Express gas stations which are more frequently found to follow. A possible explanation may lie in the ownership structure of Esso Express gas stations, as slower price adjustments could be linked to vertical contracting, or to the creation of the chain Total Access over the period. Prices at converted gas stations were indeed more likely to be under scrutiny, thus creating an incentive to price relatively aggressively[26].

To conclude, the analysis of price leadership suggests that gas station chain affiliation often plays an important role in shaping market prices. Concerned gas stations are mostly operated by supermarkets, which is consistent with gasoline prices being used to attract consumers in stores. This leads to be cautious regarding the possibility to conclude that strong price alignment reflects virtually perfect consumer information. Nevetheless, such a reserve does not applied to discount gas stations. Overall, low markup gas stations thus appear to generally compete for a well informed highly price sensitive demand.

5 Price dispersion, cost and number of firms

The following section investigates how variations in cost and competition intensities relate to market price dispersion. The first approach consists in working with measures of price dispersion for each local market and day. As it requires relatively strong assumptions regarding the definition of markets and the reliability of dispersion mea-

[25]Carrefour and Leclerc are the two leading food retailers in France

[26]Under the current test, a gas station which cuts prices quickly when the cost of wholesale diesel drops, and does not lag behind when it goes up, will typically be identified as a leader (even though it may only initiate price cuts).

sures, the richness of the data is used to measure price dispersion directly at the gas station level, from each price distribution. This allows to gather additional evidence on the link between competition and dispersion.

The first approach closely follows Chandra and Tappata (2011). Each gas station is successively considered as the center of a market delimited by a circle of a given radius. Price dispersion is then measured on each day as the empirical range and standard deviation of local prices. This allows to investigate how price dispersion varies with cost variations over time on the one hand, and with the intensity of competition across markets on the other hand. As noted by Chandra and Tappata (2011), results regarding competition must yet be analyzed with caution. The number of gas stations within an area indeed reflects demand and thus may not provide an accurate measure of the competitiveness of the market. Also, considering each gas station as the center of a market leads to attribute a lot of weight to markets which exhibit high seller densities. To tackle this issue, a simple algorithm is used to obtain non overlapping markets. We impose that i) each gas station in the market competes directly, or indirectly via another seller, with all other retailers in the market, using a 3 km threshold to define competition ii) no gas station which is not part of the market should be within 5 km of a gas station in the market. The first criterion thus ensures that all retailers in the market are close enough to each other, while the second requires the market to be relatively isolated from the closest sellers which are not part of it. This algorithm leads to identify 508 local markets, with no gas station being part of two markets.

Finally, we deal with persistent price differences by working with residual prices, namely prices net of station-specific fixed effects. Furthermore, besides following an approach which supposes that all gas stations compete in the same market, we consider a scenario involving market segmentation. Oil and independent gas stations are then assumed to compete in high price markets, while supermarket and discount gas stations form low price markets.

Table 3.6 reports descriptive statistics of price dispersion at the market level. All measures of price dispersion can be seen to drop significantly when residual prices are used and under the market segmentation scenario. For instance, under the simple 3 km radius market definition, gains from search, are estimated to be 1.25 euro cents per liter with residual prices while they were 3.93 euro cents per liter with raw prices. Under segmentation, dispersion is higher within high price gas station markets (2.90 vs. 0.97 euro cents per liter gains from search with raw prices). The variations observed between dispersion measures computed with raw and residual prices confirm than high price gas stations tend to be more differentiated than low price gas stations.

With $\text{Market Dispersion}_{jt}$ a measure of price dispersion on market j at date t, $Trend_t$ a trend variable, MC_t a measure of the marginal cost (wholesale diesel in euro per liter) on date t, N_j the number of gas stations on market j, we estimate the

Table 3.6: Market-level summary statistics

	All	Low	High	No overlap
Nb observations	3 838 661	1 345 825	509 245	539 548
Nb markets	3 605	1 262	492	508
Nb sellers	5.09 (2.61)	3.96 (1.34)	4.75 (2.43)	4.17 (1.82)
Nb sellers observed	4.88 (2.51)	3.85 (1.35)	4.47 (2.31)	3.99 (1.80)
Raw prices (euro cents per liter)				
Range	9.17 (4.44)	2.14 (1.86)	5.47 (3.31)	8.54 (4.58)
Standard deviation	3.71 (1.76)	0.87 (0.75)	2.11 (1.20)	3.65 (1.96)
Gain from search	3.68 (2.08)	0.94 (0.85)	2.71 (1.67)	3.38 (2.02)
Residual prices (euro cents per litre)				
Range	2.26 (1.62)	1.60 (1.34)	2.53 (1.68)	1.96 (1.49)
Standard deviation	0.84 (0.57)	0.64 (0.53)	0.96 (0.61)	0.78 (0.58)
Gain from search	1.13 (0.87)	0.79 (0.70)	1.25 (0.90)	0.98 (0.81)

Standard errors in parentheses.

parameters of the following equation for each market definition:

$$\text{Market Dispersion}_{jt} = \beta_0 + \beta_1 Trend_t + \beta_2 MC_t + \beta_2 N_j + \epsilon_{jt} \tag{3.2}$$

Considering that the government intervention between August 2012 and January 2013 may have had a relatively strong temporary impact on dispersion, in a period of high oil prices, estimation outputs are provided in Table 3.7 for the period starting February 1, 2013 (628 days). They are consistent with results obtained by Chandra and Tappata (2011): dispersion is found to increase with the number of firms and decrease with cost. Cost appears to have a stronger impact on dispersion displayed by high markup gas stations than low markup gas stations. When the whole period is considered (cf. Table 3.7 in appendix), cost is estimated to have a significant impact only for high markup gas stations. This result is consistent with fierce price competition among low markup gas stations, and more relaxed competition, in a context of imperfect information, for oil company and independent gas stations.

This approach however has significant shortcomings. First, it leaves out gas stations which have no or few competitors around, and either tends to give much weight to large markets, or to restrict the analysis to a limited number of markets, the definition of which remains uncertain. The empirical range is a relatively volatile statistic, and the empirical standard deviation is generally a biased estimator (absent knowledge of the distribution), which typically leads to underestimate actual standard deviation when samples are small. As a consequence, we develop another approach which consists in measuring dispersion directly from the residual price distribution of each seller. An important merit lies in the fact that potential outliers can be dropped easily (either by trimming the price distribution, or by dropping seller dispersion statistics ex-post). Measures of price dispersion thereby obtained are reported in Table 3.8.

Table 3.7: Regressions of market dispersion

	All		Low		High		No overlap	
	Range	Std	Range	Std	Range	Std	Range	Std
Constant	6.21***	2.49***	3.55***	1.62***	6.62***	2.77***	5.98***	2.54***
	(1.51)	(0.58)	(1.38)	(0.56)	(1.20)	(0.49)	(1.57)	(0.64)
Trend	-0.00**	-0.00**	-0.00	-0.00	-0.00***	-0.00***	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cost	-7.94***	-2.90***	-4.38*	-1.76*	-8.50***	-3.28***	-7.55***	-2.95***
	(2.19)	(0.84)	(2.01)	(0.82)	(1.76)	(0.71)	(2.26)	(0.91)
Nb firms	0.24***	0.04***	0.27***	0.06***	0.27***	0.05***	0.22***	0.04***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)
R2	0.19	0.07	0.12	0.06	0.21	0.09	0.10	0.04
N	632 765	632 765	222 438	222 438	83 076	83 076	88 565	88 565

Standard errors in parentheses, clustered by market and date.
Significance thresholds: * p<.1, ** p<.05, *** p<.01.
Region fixed effects are included in all specifications.

Table 3.8: Gas station residual price distributions

	Oil/Ind	Discount	Supermarkets
Nb stations	1 457	476	3 995
Std	1.11 (0.39)	1.00 (0.29)	1.06 (0.28)
Kurtosis	1.15 (3.79)	1.43 (3.81)	2.77 (3.91)
Skewness	0.07 (0.69)	0.20 (0.74)	-0.55 (0.87)
Range	6.65 (2.04)	6.24 (1.77)	7.53 (2.03)
Trimmed range 5%	4.25 (1.43)	3.87 (1.13)	4.14 (1.16)
Trimmed range 10%	3.52 (1.23)	3.18 (0.96)	3.29 (0.93)

Standard errors in parentheses.

Price distributions of supermarket gas station exhibit smaller standard deviations but fatter tails than these of oil and independent gas stations. They are generally left-skewed, which reflects the (scarce) use of promotions, often implemented at the chain level. Conversely, no systematic skew is observed for oil and independent gas stations. In order to investigate the relation between price dispersion and seller density, the empirical standard deviation and the range are regressed on the number of competitors within 3 km.

Table 3.9: Regressions of price dispersion measured at the gas station level

	Oil		Discount		Supermarkets	
	Tr. range	Std	Tr. range	Std	Tr. range	Std
Constant	3.83***	1.00***	3.51***	0.92***	3.94***	1.03***
	(0.03)	(0.01)	(0.11)	(0.03)	(0.02)	(0.01)
Nb firms	0.05***	0.02***	0.04	0.01	0.01	0.00
	(0.01)	(0.00)	(0.04)	(0.01)	(0.01)	(0.00)
R2	0.11	0.13	0.11	0.12	0.04	0.08
N	1 735	1 735	476	476	3 995	3 995

Standard errors in parentheses. Significance thresholds: * p<.1, ** p<.05, *** p<.01.
Region fixed effects are included in all specifications.

Regression results in Table 3.9 lead to question the results obtained with the previous specification. Seller density is no more found to have an impact on price dispersion for supermarkets and discounters. A significant positive relationship is yet still found for oil and independent gas stations. Estimates do not significantly differ whether seller density is accounted for by the number of competitors of the same type or by all competitors. Such results are relatively consistent with previous findings on the link between consumer information and dispersion. Indeed, supermarkets were found to often be engaged in tough price competition, leaving little room for the extraction of an informational rent. Conversely, consumer search seemed to play a bigger role in the case of oil and independent gas stations, so that a positive relationship between dispersion and competition intensity could be expected based on a model a la Varian (1980).

6 Conclusion

This paper analyzes price dispersion in the French retail gasoline market, taking into account the presence of large persistent price differences between gas stations, with supermarket and discount gas stations setting prices generally 8 to 10 euro cents per liter cheaper than oil company and independent gas stations. Overall, rank reversals are found to be less frequent for pairs of retailers separated by a short distance, namely competitors whose prices are easy to compare for consumers. This result supports the hypothesis of a connection between consumer information and price dispersion. Pairs of competitors which operate at low markups, namely supermarkets and discounters, exhibit less rank reversals than those which sell at higher markups. Pairs of competitors operating at low markups are often observed to adopt the very same prices with a high probability. The study of price dynamics allows to unveil large scale price leadership, which is strongly correlated with chain affiliation. Conversely, high markup gas stations

tend to exhibit significant dynamic price dispersion. This leads us to conclude that the latter generally address the needs of customers characterized by relatively strong loyalty or high search costs, while supermarkets compete for a well informed highly price sensitive demand.

At the market level, due to large persistent differences in gas station markups, working with raw prices leads to largely overestimate price dispersion potentially related to search costs. With price residuals, price dispersion is found to generally decrease with cost, and to increase with seller density for high markup gas stations, namely these for which competition a la Varian (1980) seems most relevant.

Regarding the fit between Varian (1980) and the data, it is worth noting that an important aspect of the dynamics remains unexplained. In the model, while firms are ex-ante indifferent between all prices in the support of the equilibrium price distribution[27], indifference obviously disappears once prices are posted on the market. The cheapest firm attracts shoppers but would be better off increasing its price to almost match the second cheapest price. Other firms would rather increase their price to consumer reservation price, or undercut the cheapest firm. In the retail gasoline market, it is thus not clear why sellers would generally wish to keep prices unchanged for a week or more, as can be observed in the data. A possible explanation may be that firms refrain from changing prices too often for fear of triggering more search by consumers and thus more intense competition. There may as well be some tacit collusion regarding the need for each gas station to regularly assert its competitiveness. Further theoretical and empirical investigations are required to understand the sources of the observed rigidity.

[27] It also holds in terms of randomization over utilities.

7 Appendix

7.1 Examples of rank reversals unrelated to mixed strategies

Figure 3.8: Example of "spurious" rank reversals: Total Access conversion

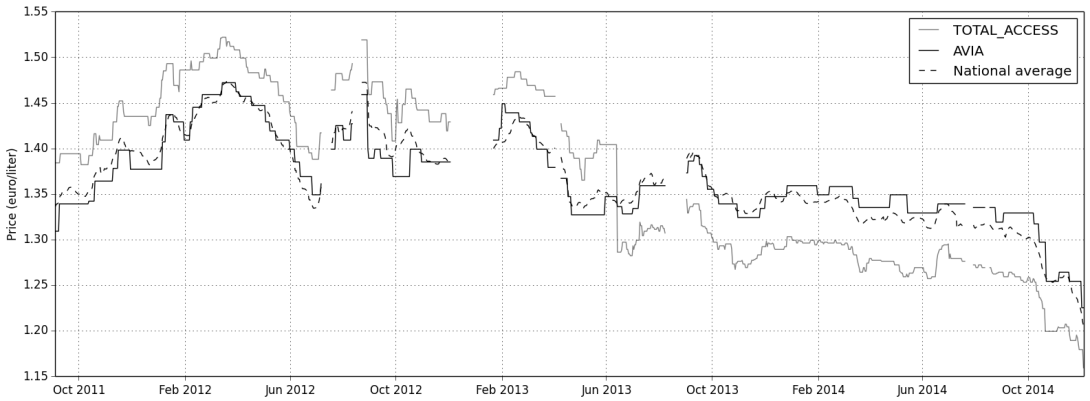
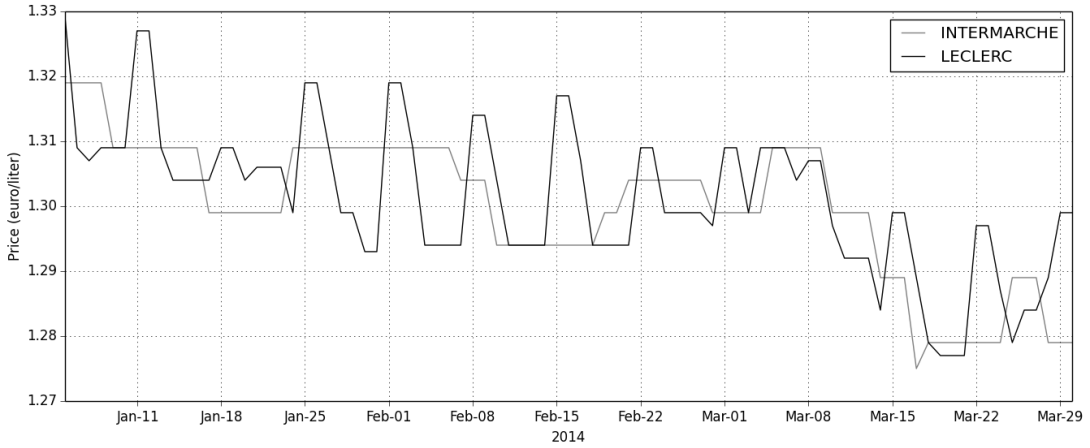
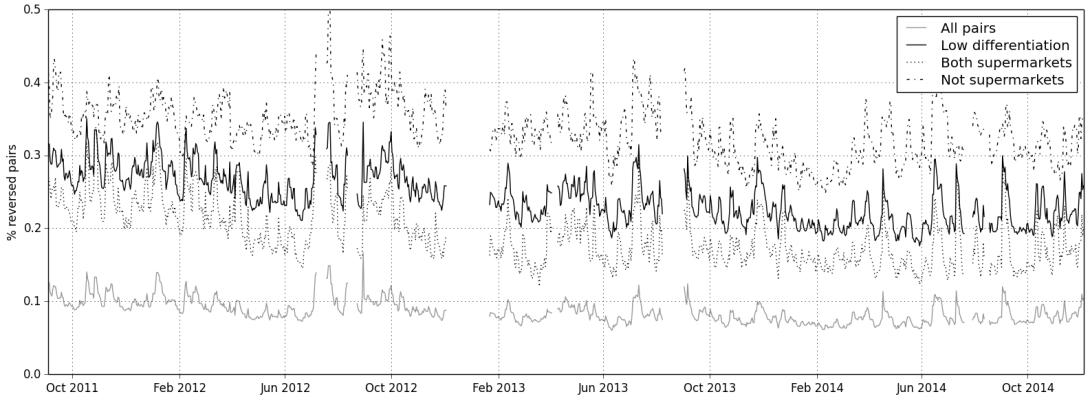


Figure 3.9: Example of "spurious" rank reversals: deterministic price variations



7.2 Rank reversals over time

Figure 3.10: Percentage of rank reversals among pairs



Series represent for each day the percentage of pairs observed where the usual price order is not respected (reversed rank). No differentiation implies that pairs exhibit an average price difference below 1c/l.

7.3 Robustness checks on competitor pairs

Table 3.11: Regressions of pair price dispersion (average price spread ≤ 2 euro cent per liter)

	Dependent Variable	OLS	Quantile regressions 25%	50%	75%	90%
All pairs	r_{ij}	-1.32***	-1.10**	-1.70***	-1.50**	-2.00**
(N = 4 171)		(0.43)	(0.45)	(0.65)	(0.75)	(0.83)
	σ_{ij}	-0.08***	-0.09***	-0.09***	-0.08***	-0.06
		(0.02)	(0.02)	(0.02)	(0.03)	(0.04)
Oil & Ind	r_{ij}	0.00	-2.10*	0.50	1.00	0.90
(N = 646)		(1.02)	(1.25)	(1.44)	(1.65)	(2.10)
	σ_{ij}	-0.17***	-0.20***	-0.14***	-0.09	-0.16
		(0.04)	(0.04)	(0.04)	(0.06)	(0.10)
Supermarkets	r_{ij}	-1.30**	-1.10**	-1.60**	-1.60	-0.40
(N = 1 867)		(0.59)	(0.47)	(0.74)	(1.08)	(1.56)
	σ_{ij}	-0.05**	-0.08***	-0.06*	-0.04	-0.07
		(0.03)	(0.03)	(0.03)	(0.04)	(0.06)
Discounters	r_{ij}	-0.11	2.40	0.50	-2.20	-2.70
(N = 156)		(1.92)	(3.67)	(2.83)	(2.36)	(2.51)
	σ_{ij}	-0.33***	-0.20**	-0.31***	-0.42***	-0.59***
		(0.06)	(0.08)	(0.09)	(0.09)	(0.10)
Supermarkets vs. Discounters	r_{ij}	-2.18***	-2.40**	-2.90**	-2.30*	-4.50***
(N = 1 153)		(0.82)	(1.08)	(1.27)	(1.29)	(1.34)
	σ_{ij}	-0.05	-0.07**	-0.06*	-0.04	-0.01
		(0.03)	(0.03)	(0.03)	(0.05)	(0.07)

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

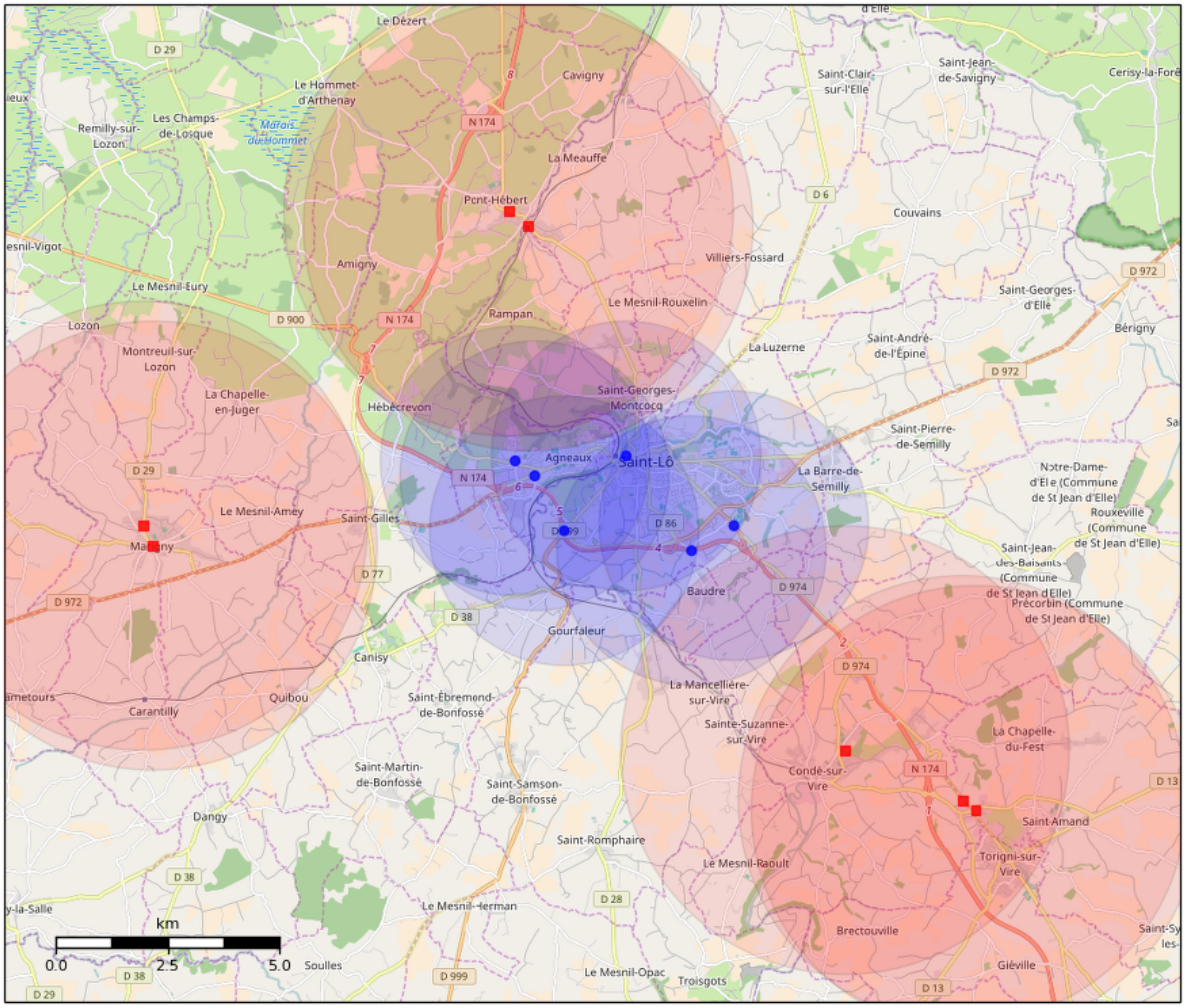
Table 3.10: Overview of pairs (distance $\leq$ 5 km)

	Nb pairs	Rank reversals	Same price	Price spread Abs. Mean	Price spread Std.
No price spread restriction					
All types	23 824	7.1 (11.0)	6.1 (14.5)	5.4 (4.3)	1.7 (0.6)
Oil & Ind	3 780	9.6 (11.5)	3.0 (7.4)	3.4 (2.6)	1.9 (0.6)
Supermarkets & Discounters	7 244	16.1 (11.8)	17.4 (21.2)	1.1 (1.1)	1.3 (0.5)
- Supermarkets	4 292	13.7 (10.9)	21.9 (25.3)	1.1 (1.2)	1.4 (0.5)
- Discounters	296	27.5 (10.6)	11.3 (6.4)	0.5 (0.6)	1.2 (0.3)
- Supermarkets vs. discounters	2 656	18.7 (11.9)	10.8 (10.5)	1.0 (0.9)	1.4 (0.5)
Price spread <= 2 cent per liter					
All types	7 973	19.1 (11.5)	17.4 (20.7)	0.8 (0.6)	1.3 (0.5)
Oil & Ind	1 277	22.2 (10.6)	7.1 (11.4)	1.1 (0.5)	1.8 (0.4)
Supermarkets & Discounters	6 050	18.3 (11.5)	20.5 (21.9)	0.7 (0.5)	1.2 (0.4)
- Supermarkets	3 473	16.0 (10.8)	26.7 (15.8)	0.7 (0.6)	1.2 (0.5)
- Discounters	286	28.1 (10.0)	11.5 (6.3)	0.5 (0.5)	1.1 (0.3)
- Supermarkets vs. discounters	2 291	20.7 (11.4)	12.3 (10.6)	0.7 (0.5)	1.3 (0.4)
Price spread <= 1 cent per liter					
All types	5 257	22.1 (11.9)	23.4 (22.8)	0.4 (0.3)	1.2 (0.4)
Oil & Ind	530	29.8 (10.0)	10.4 (15.4)	0.6 (0.3)	1.6 (0.4)
Supermarkets & Discounters	4 365	20.9 (11.8)	25.8 (23.3)	0.4 (0.3)	1.1 (0.4)
- Supermarkets	2 493	17.9 (11.4)	34.3 (26.5)	0.4 (0.3)	1.1 (0.4)
- Discounters	241	30.8 (8.2)	12.6 (6.2)	0.3 (0.3)	1.1 (0.3)
- Supermarkets vs. discounters	1 631	24.0 (11.1)	14.8 (11.0)	0.5 (0.3)	1.1 (0.3)

Except for the first column, all figures are averages of statistics computed at the pair level (standard errors in parentheses). Abs. Mean refers to $|\bar{s}_{ij}|$ and accounts for the existence of persistent price differences (0 if both gas stations are equally expensive over the long term), while the standard deviation, Std., measures dispersion around the long term price difference (0 if both gas stations always set the same price). Rank reversals and Same price are percentages.

7.4 Example of market

Figure 3.11: Example of market identified by the algorithm described in section 5



Background map from OpenStreetMap. Gas stations in the market are represented by blue circles. All gas stations within 10 km of a gas station in the market are represented by red squares. A disk of radius 3 km is drawn around each gas station within the market (5 km for those which are outside). It can be seen that the gas station at the far right of the market is not within 3 km of the gas station at the far left. Their disks nevertheless intersect with the disk of a gas station located in the middle of the market. No gas station in the market is in the disk of a gas station outside the market, which confirms that the separating distance is at least 5 km. Our algorithm leads to identify the city of Saint-Lô (Normandy) as a well-defined geographic market with 6 sellers.

7.5 Robustness checks on market dispersion

Table 3.12: Regressions of market dispersion (September 04, 2011 to December 4, 2014)

	All		Low		High		No overlap	
	Range	Std	Range	Std	Range	Std	Range	Std
Constant	2.12***	1.02***	0.28	0.32	3.16***	1.45***	1.72***	0.89***
	(0.71)	(0.27)	(0.66)	(0.27)	(0.60)	(0.23)	(0.76)	(0.30)
Trend	-0.00***	-0.00***	-0.00	-0.00	-0.00***	-0.00***	-0.00**	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cost	-1.33	-0.48	-0.87	-0.36	-2.96***	-1.14***	-0.76	-0.28
	(1.11)	(0.42)	(1.03)	(0.42)	(0.91)	(0.36)	(1.16)	(0.47)
Nb firms	0.25***	0.04***	0.28***	0.06***	0.27***	0.05***	0.24***	0.04***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)
R2	0.20	0.07	0.11	0.05	0.20	0.09	0.10	0.03
N	1 086 668	1 086 668	380 944	380 944	144 225	144 225	152 744	152 744

Standard errors in parentheses, clustered by market and date.
Significance thresholds: * p<.1, ** p<.05, *** p<.01.
Region fixed effects are included in all specifications.

Chapter 4

Competition between French supermarkets: Evidence from a price comparison website

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.

Key Words: *Chain-store retailers, uniform pricing, price dispersion*

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

[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.

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.

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

[2]Source: Kantar Worldpanel 2015.

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.

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,

Table 4.1: 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%

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 4.1 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

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

Table 4.2: 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.

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 4.2 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 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. (2016a), 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.

[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.

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 4.3. 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 4.4 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,

Table 4.3: 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 4.4: 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

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.

3.3 Price comparison results

Results of chain level comparisons performed according to the website methology are reported in Table 4.5. 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 4.6. 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 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

Table 4.5: 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.

Table 4.6: Comparisons between Leclerc stores and their competitors by chain

	Nb	Comparison of Leclerc stores vs. chain competitors						
	pairs	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.

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.

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

[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 4.7: 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.

statistics over longer periods. Indeed, product turnover generally does not allow meaningful direct comparisons between non successive price records. Table 4.6 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, 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].

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:

$$\log P_{ij} = \alpha_i \text{Product}_i + \beta_j \text{Store}_j + \epsilon_{ij} \tag{4.1}$$

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 4.8.

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

[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.

Table 4.8: 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 4.1)

differences in competition intensity. Denoting Store price index$_i$ the price index of store i and Store characteristics$_i$ a vector of variables accounting for store characteristics[8], we estimate the coefficients of the following equation:

$$\text{Store price index}_i = \mu + \beta \text{ Store characteristics}_i + \epsilon_i \tag{4.2}$$

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. 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 4.9.

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

Table 4.9: 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.

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 and Le Saout (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, 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 4.10 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 4.11 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

Table 4.10: 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 4.11: 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.

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. (2016b) 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. (2016a) 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.

[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".

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 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:

$$r_{ij} = \min \left\{ \frac{1}{L} \sum_{t=1}^{L} \mathbb{1}_{p_{il} > p_{jl}}, \frac{1}{L} \sum_{t=1}^{L} \mathbb{1}_{p_{jl} > p_{il}} \right\} \tag{4.3}$$

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 4.12 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

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

[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.

Table 4.12: 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.

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 4.13. 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.

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:

$$\mathrm{r}_{ij} = \mu + \alpha \ \text{Nearby}_{ij} + \beta X_{ij} + \epsilon_{ij} \tag{4.4}$$

In a first specification, distance as the crow flies is used, with a threshold of 5 km

[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.

[13]Cf. Wildenbeest (2011) for a model of search and price dispersion with vertical differentiation.

[14]Accordingly, β is a vector of coefficients

Table 4.13: 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.

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 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.

Table 4.14: 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.

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 (4.1). Descriptive statistics are provided in Table 4.15. 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:

$$\text{Product dispersion}_i = \mu + \alpha \text{ Price}_i + \beta_j \text{Section}_{ij} + \epsilon_i \tag{4.5}$$

Different measures of price dispersion are used depending whether the regression is

Table 4.15: National price dispersion by product section

	Nb Prod.	Raw prices Mean	CV	$\frac{Q75}{Q25}-1$	$\frac{Q95}{Q5}-1$	Residuals 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.

performed with raw prices or price residuals. Results are reported in Table 4.16. 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. 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 (4.1). 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 4.16: 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 (4.1). 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 4.17 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

Table 4.17: 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.

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:

$$\text{Product market dispersion}_{ij} = \mu + \alpha_i \text{ Product}_i + \beta \text{ Market characteristics}_{ij} + \epsilon_{ij} \tag{4.6}$$

Results are reported in Table 4.18. 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 chain index[17].

While the HHI, the population size and the population revenue are all estimated to

[15] The same can be observed with product families (available upon author request).
[16] Accordingly, β is a vector of coefficients.
[17] Alternatively, we considered a simple average of store price indexes and obtained similar results.

Table 4.18: 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.

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.

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.

Conclusion

The goal of this dissertation was to investigate how detailed price observations can be used to assess the competitiveness of a market. In order to achieve this, I have built two large datasets, almost exclusively based on publicly available data. The code used to collect, prepare and analyse the data has been published together with the data. It is my hope that this work will contribute towards developing a culture of systematic market monitoring, both with a view to increase transparency for consumers and to assist competition authorities in their missions.

Bibliography

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

Angrist, J. and Pischke, J.-S. (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.

Auffhammer, M. and Kellogg, R. (2011). Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review*, 101(6):2687–2722.

Bacon, R. (1991). Rockets and feathers: The asymmetric speed of adjustment of u.k. retail gasoline prices to cost changes. *Energy Economics*, XIII:211–18.

Barron, J., Taylor, B., and Umbeck, J. (2004). Number of sellers, average prices and price dispersion. *International Journal of Industrial Organization*, 22:1041–1066.

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

Baye, M. R., Morgan, J., and Scholten, P. (2004). Price dispersion in the small and in the large: Evidence from an internet price comparison site. *The Journal of Industrial Economics*, 52(4):463–496.

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

Bellec, G., Bolliet, A., Meslin, O., Ravignon, B., and Serres, H. (2012). Les prix, les marges et la consommation de carburants. *Rapport au Ministre de l'Économie et des Finances*.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How must should we trust differences in differences estimates? *Quaterly Journal of Economics*, 119(1):249–275.

Bertrand, M. and Kramarz, F. (2002). Does entry regulation hinder job creation? evidence from the french retail industry. *Quarterly Journal of Economics*, 117(4):1369–1413.

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

Borenstein, S., Cameron, A., and Gilbert, R. (1997). Do gasoline price respond asymmetrically to crude oil price changes? *The Quarterly Journal of Economics*, 112:305–399.

Bruzikas, T. and Soetevent, A. R. (2015). The proliferating of unmanned fuel retailing and cost pass-through: Empirical evidence from the netherlands. *Working Paper*.

Burdett, K. and Judd, K. L. (1983). Equilibrium price dispersion. *Econometrica*, 51(4):955–969.

Cameron, C. A. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–373.

Carlson, J. A. and McAfee, R. P. (1983). Discrete equilibrium price dispersion. *Journal of Political Economy*, 91(3):480–493.

Chamayou, E. and Le Saout, R. (2016). Changement d'enseigne sur le marché français des carburants: Effets sur la concurrence locale d'une baisse des prix. *Revue Economique (forthcoming)*.

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

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory*, 3(2):156–158.

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

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

Eckert, A. (2002). Retail price cycles and response asymmetry. *The Canadian Journal of Economics*, 35:52–77.

Eckert, A. (2003). Retail price cycles and the presence of small firms. *International Journal of Industrial Organization*, 21:151–170.

Eckert, A. (2013). Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys*, 27(1):140–166.

Eckert, A. and Douglas, S. W. (2004a). Retail gasoline price cycles across spatially dispersed gasoline stations. *Journal of Law and Economics*, 47:245–273.

Eckert, A. and Douglas, S. W. (2004b). A tale of two cities: Price uniformity and price volatility in gasoline retailing. *Annals of Regional Science*, 38:25–46.

Einav, L. and Levin, J. (2010). Empirical industrial organization: A progress report. *Journal of Economic Perspectives*, 24(2):145–162.

Frey, G. and Manera, M. (2007). Econometric models of asymmetric price transmission. *Journal of Economic Surveys*, 21(2):349–415.

Gautier, E. and Le Saout, R. (2015). The dynamics of gasoline prices: Evidence from daily french micro data. *Journal of Money, Credit and Banking*, 47(6):1063–1089.

Hastings, J. (2004). Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in southern california. *The American Economic Review*, 94(1):317–328.

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

Houde, J.-F. (2008). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, 105(5):2147–82.

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

Lewis, M. (2008). Price dispersion and competition with differentiated sellers. *The Journal of Industrial Economics*, 56(3):654–678.

Lewis, M. (2009). Temporary wholesale gasoline price spikes have long-lasting retail effects: The aftermath of hurricane rita. *Journal of Law and Economics*, 52(3):581–605.

Lewis, M. (2011). Asymmetric price adjustment and consumer search: An examination of the retail gasoline market. *Journal of Economics and Management Strategy*, 20(2):409–449.

Lewis, M. (2012). Price leadership and cooperation in retail gasoline markets with price cycles. *International Journal of Industrial Organization*, (2):409–449.

Lewis, M. and Noel, M. (2011). The speed of gasoline price response in markets with and without edgeworth cycles. *The Review of Economics and Statistics*, 93(2):672–682.

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

Noel, M. D. (2007a). Edgeworth price cycles, cost-based pricing and sticky pricing in retail gasoline markets. *Review of Economics and Statistics*, 89:324–334.

Noel, M. D. (2007b). Edgeworth price cycles: Evidence from the toronto retail gasoline market. *Journal of Industrial Economics*, 55:69–92.

Noel, M. D. (2008). Edgeworth price cycles: Computational dynamic markov equilibria. *Journal of Economics and Management Strategy*, 17(2):345–377.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., and Yontcheva, B. (2015). Information and price dispersion: Theory and evidence. *Working Paper*.

Rothschild, M. (1973). Models of market organization with imperfect information: A survey. *Journal of Political Economy*, 81(6):1283–1308.

Seaton, J. s. and Waterson, M. (2013). Identifying and characterising price leadership in british supermarkets. *International Journal of Industrial Organization*, 31:392–403.

Stahl, D. (1989). Oligopolistic pricing with sequential consumer search. *American Economic Review*, 79(4):700–712.

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

Tappata, M. (2009). Rockets and feathers: Understanding asymmetric pricing. *RAND*, 40(4).

Taylor, C. T., Kreisle, N. M., and Zimmerman, P. R. (2010). Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in southern california: Comment. *American Economic Review*, 100(3):1269–76.

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.

Zimmerman, P. R. (2012). The competitive impact of hypermarket retailers on gasoline prices. *Journal of Law and Economics*, 55(1):27–41.

Titre: Dispersion des prix et recherche des consommateurs: analyses du marché des carburants et de la grande distribution en France

Mots clefs: concurrence, dispersion des prix, carburant, supermarchés

Résumé: Cette thèse est un travail empirique sur la dispersion des prix, c'est-à-dire le fait qu'un bien identique puisse être vendu à des prix différents, en violation de la célèbre loi du prix unique. L'approche s'inscrit dans une litérature initiée par Stigler (1961), qui note que "la dispersion des prix est (...) la mesure de l'ignorance dans le marché". Il en découle que de simples observations de prix peuvent révéler beaucoup d'information sur le fonctionnement d'un marché. Le premier chapitre étudie l'impact de la création d'une enseigne à bas prix sur le marché français de la distribution de carburant. Cette création implique que de nombreuses stations sont confrontées à une baisse de prix importante d'un proche concurrent. La réaction agrégée mesurée au niveau national est faible, mais masque des hausses et des baisses en proportions équivalentes. L'hétérogénéité des réactions souligne l'importante segmentation du marché. Le second chapitre, utilisant le même jeu de de données, examine la pertinence de modèles qui identifient la dispersion des prix à des équilibres en stratégies mixtes. Empiriquement, on observe que l'ordre des prix des stations concurrentes tend en effet à varier dans le temps, et que sa volatilité croit lorsque la distance qui sépare les points de vente augmente. La dispersion est donc croissante d'un coût recherche supporté par les consommateurs. Par ailleurs, l'enseigne des stations détermine largement leur stratégie de prix. Les stations qui pratiquent les prix les plus bas sont relativement plus susceptibles de maintenir des prix parfaitement alignés sur ceux des proches concurrents, tandis que la dispersion mesurée entre les stations plus onéreuses est corrélée négativement avec le coût du diesel et positivement avec le nombre de stations présentes sur le marché. Ainsi, les résultats renforcent la thèse de la coexistence d'un marché proche d'une concurrence à la Bertrand avec un marché moins concurrentiel, où des stations tirent parti de frictions importantes. Le dernier chapitre s'intéresse à la grande distribution, s'appuyant sur des données collectées sur un comparateur de prix en ligne. J'observe que les comparaisons réalisées entre chaînes de magasins par le site sont relativement peu informatives compte tenu de l'hétérogénéité des résultats au niveau local. L'échantillon de produits retenu peut en outre conduire à des résultats largement différents. La volatilité des comparaisons augmente avec la distance séparant les supermarchés, ce qui, comme dans le cas du carburant, dénote la présence de coûts de recherche. A l'échelle locale, le niveau des prix ne croit pas avec la concentration approximée par les parts de marché, ce qui conduit à remettre en question la pertinence de cet indicateur en matière de politique publique. La dispersion est positivement corrélée au niveau des prix, ce qui suggère que l'imperfection de l'information permet effectivement aux supermarchés de pratiquer des prix plus élevés qu'en concurrence parfaite.

Title: Price dispersion and consumer search: Evidence from the retail gasoline market and the supermarket industry in France

Keywords: competition, price dispersion, gasoline, supermarkets

Abstract: This thesis is an empirical study of price dispersion, namely the fact that a homogenous good can typically be purchased at various prices, in violation of the famous law of one price. The approach belongs to a literature initiated by Stigler (1961), which notes that "price dispersion is (...) the measure of ignorance in the market". A noteworthy consequence is that simple price observations can be very informative about competition in a market. The first chapter analyses the impact of the creation of a discount chain on the French retail gasoline market. This creation implies that many gas stations are confronted with a sharp price decrease by a competitor. The aggregate reaction, measured at the national level, is weak but it conceals increases and decreases in equivalent proportions The heterogeneity of measured reactions highlights an important market segmentation. Using the same data, the second chapter explores the relevance of models which identify price dispersion with an equilibrium in mixed strategies. Empirically, the rank of competing gas stations is indeed observed to vary over time, and its volatility is positively correlated with the distance that separates the outlets. Dispersion thus increases with a search cost incurred by customers. The chain affiliation of gas stations largely determines their pricing strategies. Retailers which have low price policies are more likely than others to keep prices aligned with nearby competitors, while dispersion measured between more expensive gas stations is negatively correlated with diesel cost and positively with the number of sellers in the market. Results thus support the idea that a market close to Bertrand competition coexists with a less competitive market, where gas stations take advantage of significant frictions. The last chapter focuses on grocery stores, using data collected from an online price comparison website. Aggregate national chain comparisons that are displayed on the website are found to provide information of little value to consumers given the heterogeneity observed within store level comparison results. Furthermore, these can vary significantly depending on the set of compared products. Volatility tends to increase with the distance that separates supermarkets, which, as in the case of gasoline, suggests that search cost influence competition. Within local markets, the measured concentration is negatively correlated with price levels. This leads to question its relevance in terms of public policies. Price dispersion is found to increase with market price levels, which is consistent with sellers taking advantage of consumer search costs to post higher prices.