

Price dispersion on the French retail gasoline market

(Very preliminary)

Etienne Chamayou*

CREST and Department of Economics, Ecole Polytechnique

February 8, 2016

Abstract: This paper studies static and dynamic price dispersion in the French retail gasoline market. Using a large panel of daily gas station prices to produce statistics accounting for local competition, it finds support for a link between consumer information and price dispersion. Indeed, among competitors exhibiting similar price levels, changes in price rankings are found to be positively correlated with the distance separating both sellers, namely with the difficulty for consumers to compare prices. Supermarkets, which generally operate at low markups, are over-represented among pairs of competitors with similar price levels as well as among pairs of competitors which regularly set at the very same price. This is consistent with them addressing demand from well informed price sensitive customers, as opposed to oil company and independent gas stations generally more focused on less informed and/or more loyal or captive customers.

Keywords: Consumer search, Price dispersion

JEL Classification Numbers: L13

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

1 Introduction

The seminal paper of the consumer search literature, Stigler (1961), draws attention to the relation between the "ignorance in the market", namely consumers' lack of information, and price dispersion i.e. the persistence over time of different prices for a homogeneous good. While the paper only provides evidence of static price dispersion e.g. retailers posting different prices for a given car model in a suburb, thus leaving room for unobserved differentiation, the empirical literature would soon make the case stronger through the observation of price dynamics. The ranking of sellers in price distributions would indeed be observed to vary significantly over time, making it difficult for consumers to find the cheapest price.

The retail gasoline market offers interesting case studies for empirical industrial organization. Indeed, consumers generally buy only one product which is essentially homogeneous, suggesting a potentially tough price competition whenever the density of sellers is high enough. More generally, transportation costs can be expected to relax competition, as well as uncertainty about prices, all the more as oil cost variations generate constant retail price variations.

Competition in the retail gasoline market has attracted a lot of attention in France over last years. In 2006, a price comparison website was launched by the Ministry of Economics with a view to increase transparency in the market. In 2011, a large inquiry was ordered by the government as retail price adjustments to oil cost variations were suspected to occur faster upwards than downwards (the "Rocket and feathers" hypothesis first investigated by Bacon (1991) in the UK). Little evidence was yet found and the report rather emphasized signs pointing toward a significant degree of competitiveness. More recently, in August 2012, following an election promise, the government of the then newly elected French President Francois Hollande implemented a 3 Euro cents per litre tax cut during 3 months, asking gasoline retailers on the market to produce a similar effort. The French retail gasoline market thus tends to be thoroughly scrutinized.

More generally, in a context where dynamic prices generate heightened debates (e.g. Amazon), the idea to monitor competition by constraining firms to increase transparency enjoys strong popularity among politicians. Nevertheless, economic theory shows that measures affecting information can have adverse effects. Costless monitoring indeed makes collusion easier, while aligned prices provide little incentive for consumers to become or remain informed.

The following paper elaborates on the methodology used by Chandra and Tappata (2011) to study price dispersion, in order to deal with the differentiation observed in the French market and discuss the sources of observed price dispersion. Among pairs of competing stations exhibiting little differentiation, price dispersion is found to significantly increase with distance separating stations (very close competitors vs. competitors located further away), providing support for the connection between consumer information and price dispersion. The effect is significantly weaker for differentiated competitors. At the market level, price dispersion is found to increase with the

number of competitors and decrease with price level.

2 Literature

3 The theory of price dispersion

A major contribution of the consumer search literature was made by Varian (1980) through the modeling of price dispersion as a result of mixed pricing strategies. According to the paper, price dispersion thereby obtained could be interpreted as "temporal" price dispersion, typically in the form of "sales". This provided a rationale for rank reversals i.e. a seller being cheaper or more expensive than a competitor with positive probabilities.

The dynamic interpretation of the model is not unambiguous however. In the model, firms are ex-ante indifferent between all prices in the support of the equilibrium price distribution (also holds in terms of randomization over utilities). Ex-post, indifference obviously no more exists. The cheapest firm attracts shoppers but would be better off increasing its price to match the second cheapest price. Other firms would rather increase their price to consumers' reservation price, or try to undercut the cheapest firm. In the retail gasoline market, while it may not be possible for firms to change prices on too frequent a basis, periods of significant oil price variations reveal that gas stations can adjust prices on a daily basis. If one admits that gas stations play according to Varian (1980) on a given day, it is thus not clear why sellers would wish to keep prices unchanged the following day. 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. Whether this can be obtained in a fully competitive setting or would require some kind of collusion (possibly at the retail chain level or at the local level) remains an open question.

4 Empirical analysis

Chandra and Tappata (2011) work with a data set including daily prices of gas stations within four states in the US over one year and a half. They argue that gas stations located across the same corner should display less rank reversal vs. those a bit further, since consumer information on prices should decrease with distance. This is shown to be true and all the more convincing as close gas stations display a lower average spread. Regressions of various measures accounting for price dispersion on marginal costs and the number of firms in the market (built by considering circles of varying radiuses) yields results consistent with the extension of Varian (1980) proposed by the paper (resp. negative and positive impacts). All the analysis is performed with raw prices, which, I argue, is bound to bias some results.

A simple way to measure temporal price dispersion between two stations with daily observations is to consider the probability that the station which is in general cheaper (in terms of day count) turns out to be more expensive. Formally, considering the prices p_{it} and p_{jt} of two stations i and j over T_{ij} days, such that $p_{it} \geq p_{jt}$ is observed most of the time, the rank reversals statistic writes:

$$r_{ij} = \frac{1}{T_{ij}} \sum_{t=1}^{T_{ij}} \mathbb{1}_{p_{jt} > p_{it}}$$

5 The French retail gasoline market

5.1 Retail gasoline distribution

Diesel consumption currently accounts for c. 80% of retail gasoline sales. A significant increase in the diesel share was achieved over the last decades through a lower tax on diesel. Meanwhile, the size of the French gas station network has been decreasing at a steady pace, from c. 40,000 in 1980 to c. 12,000 currently. Unlike most other European countries, the French market is characterized by a strong competitive pressure generated from gas stations operated by supermarket chains. They currently represent c. 50% of sales in retail gasoline.

According to the French Union of Petroleum Industries (UFIP), there were 11,662 gas stations operating in France (of which 4,947 operated by supermarket chains) in 2012 and 11,476 in 2013 (4,979 for supermarkets). Regarding volumes, it was reported that 1,506 gas stations sold less than 500m³ in 2012 (1,433 in 2013), with the median gas station selling between 1,000 and 3,000m³ (same for 2013). As of May 20, 2014, the price comparison website Zagaz listed 12,832 gas stations, but no price has been posted recorded over the last months/years for many of them so that this figure most likely overestimates the actual number of gas stations.

Gas stations are essentially either owned and operated by a chain or with a "location-gerance" contract according to which the manager receives a commission on gasoline sold (e.g. only 200 gas stations set prices independently among gas stations from Total brand). Margins are largely acknowledged to have decreased significantly over the last decade, 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 reduce significantly the size of their network (Esso, Total).

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

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 employees' consumptions 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 can get information about prices from a variety of sources: at gas stations, on their gps, on mobile phone applications (e.g. Zagaz, Carbeo, Essence Free) and on a computer or mobile phone browser (Prix-Carburants.gouv.fr).

The period covered by the paper is marked by two significant events of different natures. The first is the creation by the largest gas station operator in the country, Total, of a discount brand with a view to recapture market shares lost to supermarket gas stations. This creation was achieved through the rebranding of c. 600 gas stations between 2011 and 2014, accompanied for about half of them by a 10 euro cents per litre decrease in prices. The second event is of political nature. On August 29, 2012, following an election promise made by Francois Hollande, the government announced a decrease in price of 6 euro cents per litre, (to be) achieved by a decrease of tax of 3 Euro cents per litre and an equivalent "effort" by gas station operators.

5.2 Consumer information and data sources

Since 2007, French gasoline retailers are required by law to keep prices updated on the price comparison website prix-carburants.gouv.fr. Small gas stations¹ are exempted from this obligation hence c.10,000 gas stations are observed out of an estimated total number of 12,000 retailers². Significant limitations of the governmental comparison website include the fact that users have never been provided a way to report errors such as wrong or out of date prices or wrong gas station locations, and that price comparison functionalities have always remained very poor: e.g. one cannot view prices for a given highway, nor the rivals of a given gas station on a map. As a consequence, it does not seem far-fetched to suspect that the creation of this website was actually detrimental to consumer information, as it diverted drivers from other comparison websites such as Zagaz at a time when they crucially needed to grow their user base.

5.3 Descriptive statistics

Prices and brand changes were collected from prix-carburants.gouv.fr on a daily basis between September 4, 2011 and December 4, 2014, hence a period of about 3 years, interrupted by some gaps related to data acquisition issues. Over this period, prices of 10,180 gas stations (after duplicate reconciliations) were recorded, of which 437 were located on highways, 124 on the island of Corsica,

¹Stations having sold over 500m³ gas the previous year

²A 2012 governmental report on the French retail gasoline market notes that "nobody knows precisely the number of gas stations operating in the markets". Two other comparison websites, carbeo.com and zagaz.com, were created in 2005 and 2006, relying on user provided information. Zagaz has stuck to its "crowdsourcing" philosophy until 2014, while Carbeo started purchasing licences from the government in 2009. In 2012, the governmental body in charge of town and country planning worked with Zagaz data to study the French retail gas station network.

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. Due to the fact that price observations are not always available (gas station maintenance, probable mistake, stations which are not required by law to keep prices posted), there is actually an average of 7,895 prices observed on a daily basis. On average, c. 1,500 gas stations (c. 18% of gas stations observed and retained) change prices within a day. The average gas station changes price a little less than every week.

Static price dispersion in French Data is easily seen in any single day price cross section. The resulting price distribution is bimodal, largely reflecting the presence of two dominant types of actors: supermarkets and oil companies. Supermarkets typically state that they use gasoline to attract customers, while oil companies mention the quality of the service to justify higher prices. Oil companies also traditionally enjoy more demand from businesses.

5.4 Price dispersion and consumer information

A database is built in which observations are pairs of stations assumed to compete within the same market. In the following analysis, distance as the crow flies is used with an upper distance limit of 3km (figures are provided for a limit of 5km in appendix). This results in a database describing price competition between XXXXX pairs of rival gas stations (XXXXX pairs with 5km).

A sensitive issue in the empirical analysis is the treatment of differentiation, broadly understood as lasting differences in prices. Such an issue is traditionally addressed by working with price residuals i.e. after subtraction of gas station (and time, depending on the application) fixed effects, which is bound to be imperfect. The analysis is thus performed with raw prices whenever differentiation is found to be relatively small, and with residual prices when differentiation requires it.

Figure 3 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 3km, 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.

A clear ranking of empirical distribution functions of rank reversals can be observed among pairs of gas stations depending on distances. This is consistent with the idea that nearby gas stations compete in a (virtually) complete information setting, where there is no particular reason to expect rank reversals. Conversely, distance can create an information issue for other pairs, leading to the inexistence of an equilibrium in pure strategies.

Table 1: Pair rank reversals

	Raw prices		Residual prices		
	All	No Total Access	Price difference $\leq 2c/l$ (mean)	Price difference $\leq 2c/l$ (mean)	Price difference $\geq 2c/l$ (mean)
Max d_{ij} : 3km					
Nb pairs	16,142	13,846	5,166	5,166	8,680
Mean	0.09	0.08	0.19	0.39	0.45
Std	0.13	0.12	0.13	0.11	0.05
Median	0.02	0.01	0.17	0.43	0.46
Max d_{ij} : 5 km					
Nb pairs	33,478	28,433	10,314	10,314	18,119
Mean	0.09	0.08	0.21	0.41	0.45
Std	0.13	0.12	0.13	0.10	0.04
Median	0.02	0.01	0.19	0.44	0.46

Pairs including a Total Access gas station are excluded in the three last columns

In order to evaluate the hypothesis according to which information is connected to price dispersion, two measures of price dispersion: rank reversals and standard deviation in spread, are regressed on a proxy for consumer information: distance between stations.

Table 2: Regression of rank reversals

	No differentiation raw prices	No differentiation residual prices	Differentiation residual prices
OLS: distance	0.014*** (0.002)	0.020*** (0.002)	0.001 (0.001)
OLS: same corner	-0.024*** (0.007)	-0.034*** (0.006)	-0.006*** (0.002)
Quantile 0.25: same corner	-0.027*** (0.008)	-0.067*** (0.010)	-0.007*** (0.003)
Quantile 0.50: same corner	-0.024* (0.011)	-0.030*** (0.005)	-0.007** (0.002)
Quantile 0.75: same corner	-0.027* (0.014)	-0.007* (0.003)	-0.002 (0.001)
N	4,183	4,183	8,312

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The "same corner" variable is a dummy variable that takes value one for pairs of gas stations separated by a distance of less than 500 meters, namely those for which consumers are expected to be relatively well informed about prices at each retailer.

Overall, results seem to support the connection between price dispersion and consumer information. Pairs separated by a short distance exhibit less rank reversals, while the average price difference tends to be smaller. Though significant, the effect is weaker for differentiated gas stations. This might partly be due to errors introduced by the cleaning of prices. However, for pairs exhibiting little or no differentiation, results do not differ significantly whether raw prices or residual prices are taken into account.

6 Market price dispersion

While the emphasis has so far been laid on examining empirical evidence about the link between price dispersion and consumer information, the following section seeks to compare predictions from a model à la Varian (1980) with French retail gasoline market data.

The methodology employed by Chandra and Tappata (2011) consists in considering each gas station successively as the center of a market delimited by a circle of a given radius (robustness of results to variations in radius size is checked). Various statistics accounting for price dispersion can then be computed for each market and period. If one observes N stations close to each other (e.g. with a limit radius of 2km: if no two stations are separated by a distance of more than 2 km) and with prices posted over T periods, this will result in $N * T$ observations as each station is successively considered and, for each station, a statistic representing price dispersion is computed at each period. An expected consequence is that it tends to over represent markets with higher gas station density (typically cities).

Table 3: Overview of market price dispersion

	Radius 5km	Radius 3km	No overlap
Nb of markets	6794	5614	386
Avg nb competitors	11.2	6.7	3.9
Raw prices			
Market price range	0.116	0.105	0.089
Market gain from search	0.046	0.044	0.036
Market price std	0.046	0.045	0.045
Residual prices			
Market price range	0.033	0.028	0.020
Market gain from search	0.017	0.014	0.010
Market price std	0.011	0.011	0.010

Each statistic accounting for price dispersion is first computed for each local market across time. The table reports averages across markets. At the market level, price dispersion has been computed only for days when more than two thirds of competitors' prices were available. Only markets with more than 50 days of data are then taken into account to compute reported statistics.

Regressions are then run using raw prices and cleaned prices. Also, in order to address the issue of overlapping markets, regression are performed on dispersion data generated in a way that prevents market overlap. Descriptive statistics performed on raw prices are provided.

Standard market dispersion statistics such as range and standard deviation are strongly influenced by static price dispersion. As a consequence, contrarily to the previous section where descriptives statistics performed on raw prices arguably provided a most conservative first idea of price dispersion, clean prices are now to be preferred despite their shortcomings. Results are reported with clean prices, raw prices for markets obtained with radiuses of respectively 3 and 5km. Additionnally, results are reported for non overlapping markets built in a way that is detailed in appendix.

Estimation results show that dispersion is increasing in the number of firm and decreasing in price.

7 Conclusion

This paper discusses the methodology of Chandra and Tappata (2011) to measure price dispersion in the field and see how it compares with potential theoretical explanations. The presence of differentiation in the French retail gasoline market yet requires to work both with raw and residual prices.

Rank reversals are generally found to be less frequent for pairs gas stations separated by a short distance i.e. competitors whose prices are easy to compare for consumers, hence supporting the connection between consumer information and price dispersion. When differentiation is significant, regressions performed on residual prices show a weaker effect of distance. This suggests that one should be cautious when there is differentiation despite the high number of rank reversals that one obtains with residual prices.

At the market level, price dispersion is found to increase with the number of competitors, and decrease with cost. This is consistent with predictions of a model à la Varian (1980). Importantly, in the presence of significant differentiation, working with raw prices typically leads to overestimate price dispersion reflecting randomized pricing. The average gain from search across markets and time over the period computed with residual prices varies between 1.7 and 1.0 Euro cents per litre meaning that knowing all prices in the market allows to spare on average between 50 and 85 Euro cents for a gas tank of 50L.

Further research remains to be undertaken in two directions: one is the investigation of the heterogeneity in local market price patterns and the other is the combined analysis of price dispersion together with data about consumer search (e.g. price comparison website frequentation over time if data are released).

References

- 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.
- Chandra, A. and M. Tappata (2011). Consumer search and dynamic price dispersion: an application to gasoline markets. *RAND Journal of Economics 42*(4), 681–704.
- Stigler, G. (1961). The economics of information. *The Journal of Political Economy 69*(3), 213–25.
- Varian, H. (1980). A model of sales. *American Economic review 70*(4), 651–659.

A Descriptive statistics

Figure 1: Station-level summary statistics

	All	Supermarkets	Others
Nb stations			
All periods	9,217	5,055	4,162
Nb of daily observations	7,895	4,399	3,495
Price and Markup (euro / litre)			
Price after tax	1.37	1.33	1.40
Price excl. tax	0.71	0.68	0.74
Markup over wholesale cost	0.10	0.08	0.13
Retail price variations (euro cents / litre)			
Daily price change probability	0.17	0.18	0.15
Average price increase	1.30	1.29	1.31
Average price decrease	-1.47	-1.48	-1.46
Competition			
Distance to closest rival	2.25	2.64	1.79
Nb of rivals within 3 km	3.22	2.51	4.08
Nb of rivals within 5 km	6.57	4.89	8.61

Notes: 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, hence a net margin generally estimated at c. 1 euro cent / litre.

Figure 2: Daily Brent and French diesel retail prices excluding taxes)

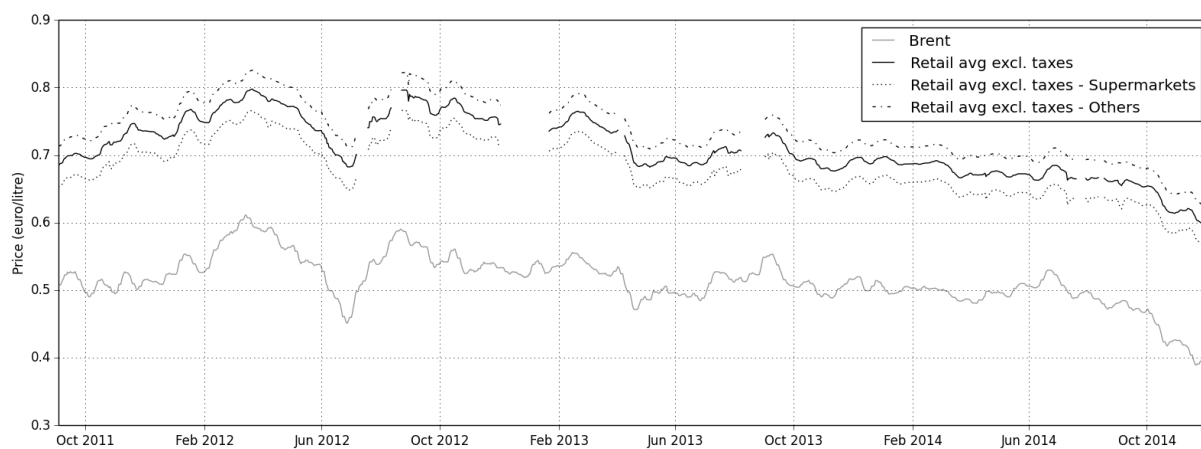
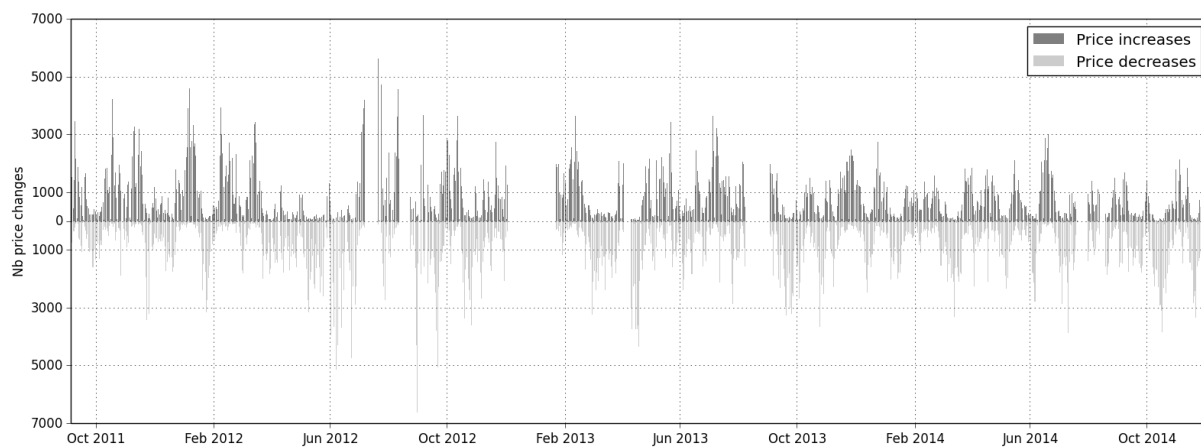
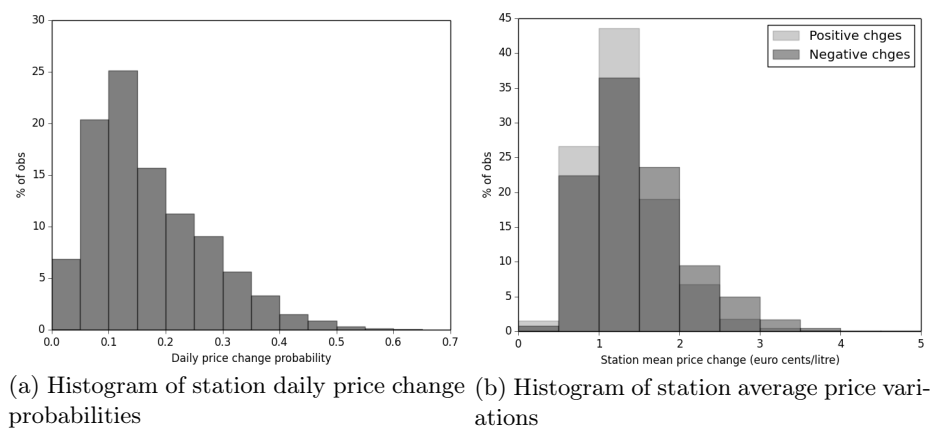


Figure 3: Daily number of retail price changes



The daily collection of prices has occasionally been delayed by a few hours resulting in some imprecision regarding the exact number of price changes occurring within one day

Figure 4: Price changes by station)



B Rank reversals

Figure 5: Empirical distribution functions of rank reversals (raw prices)

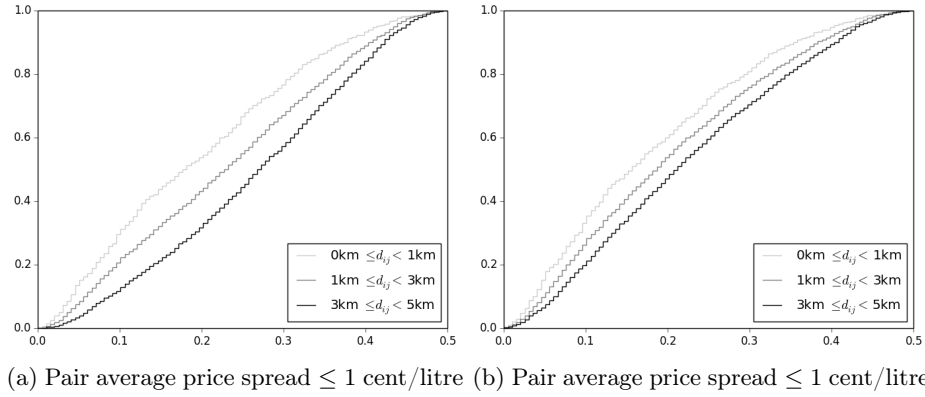
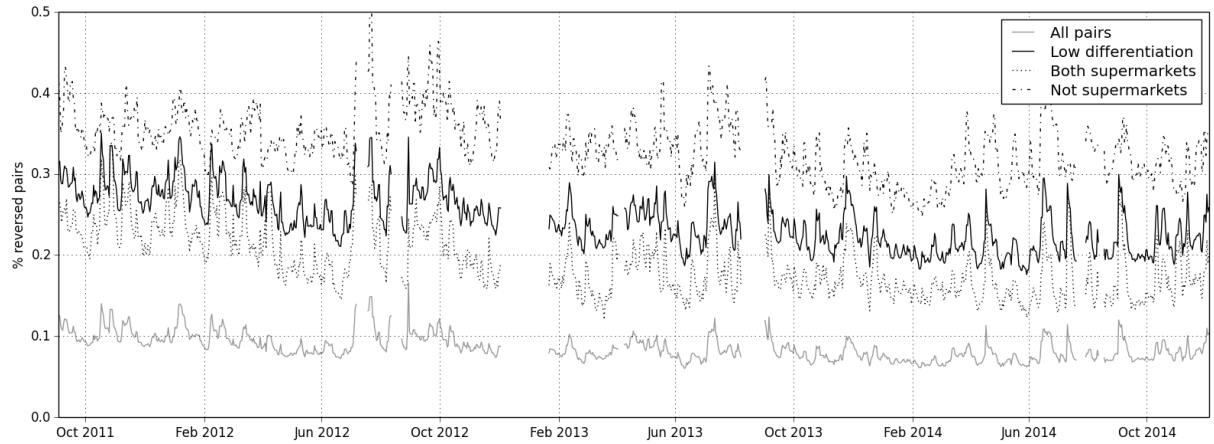


Figure 6: Percentage of rank reversals among pairs



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.

C In progress

C.1 Spurious rank reversals

French data raise the issue of the nature of rank reversals measured by the statistic previously described. Indeed, the transformation by the firm "Total" of a significant number of its gas stations to low cost gas stations generates rank reversals which are clearly unrelated to the use of randomized pricing (cf. appendix). In order to avoid including "spurious rank reversals" in the analysis, the frequency of changes in price ranking was also considered in the analysis. Conservatively, pairs of gas stations involving such a brand and price policy change have been excluded from the analysis. It has also been checked that price policies implemented by competitors of rebranded gas stations were not significantly affected.

The following figure displays prices posted at a gas station which is rebranded "Total Access" (low cost brand of Total) in the middle of the period studied and at two nearby gas stations. It is easily seen how naively considering the percentage of rank reversals of the rebranded gas station vs. each of its competitors over the whole period would lead to record spurious price dispersion. The price level indeed changes abruptly in a way that cannot go unnoticed by consumers. It thus clearly does not reflect any use of randomized pricing and such rank reversals are thus excluded from the analysis. More generally, this suggests the need to control the frequency of changes in price ranking so as to limit the risk of capturing dispersion reflecting such deterministic changes in pricing policies.

C.2 Number of digits after decimal point

There is heterogeneity between the number of digits displayed after decimal point among stations. Some stations indeed appear to restrict themselves to two digits after decimal point, hence giving up the possibility to undercut or match a competitor's price. Though accounting for this phenomenon is out of the scope of this paper, its connection with price dispersion is (to be) investigated.

C.3 Leadership

Maskin and Tirole (1988) [TODO: add ref] offer an alternative explanation for rank reversals. They indeed show that in a context where sellers have to commit on prices for at least one period, cycles can occur which are characterized by periods of undercutting punctuated by one firm accepting to set a high price when margin becomes too small. XXX [TODO: add ref] find evidence of such cycles with XXX data. Though standard test rule out the presence such cycles in French data, evidence of other specific patterns of prices has been found.

C.4 Monopoly, information, inventory and rule of thumb

Finally, a natural explanation for rank reversals could be that gas stations actually do not really compete and that their prices simply reflect their flow of information and/or inventory through more or less complex pricing rules. Such explanations are (to be) examined.

Table 4: Regressions of market dispersion

	All Raw prices	All Residual prices	No overlap Residual prices	Stable markets Residual prices
Gains from search				
Nb competitors	0.003*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Cost	0.007* (0.003)	-0.006* (0.003)	-0.005 (0.003)	-0.005 (0.003)
Intercept	0.015*** (0.004)	0.015*** (0.004)	0.013*** (0.004)	0.012** (0.004)
Standard deviation				
Nb competitors	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000 (0.000)
Cost	-0.001 (0.002)	-0.009*** (0.002)	-0.009*** (0.002)	-0.008*** (0.003)
Intercept	0.044*** (0.003)	0.022*** (0.003)	0.021*** (0.003)	0.021*** (0.004)
Coefficient of variation				
Nb competitors	0.000*** (0.000)			
Cost	-0.024*** (0.001)			
Intercept	0.064*** (0.002)			
Range				
Nb competitors	0.006*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.000)
Cost	-0.004 (0.005)	-0.025*** (0.005)	-0.021*** (0.005)	-0.019*** (0.006)
Intercept	0.074*** (0.007)	0.048*** (0.007)	0.041*** (0.007)	0.038*** (0.008)
N	2,756,096	2,756,096	493,356	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Regression of standard deviation in spread

	No differentiation raw prices	No differentiation residual prices	Differentiation residual prices
OLS: distance	0.001 *** (0.000)	0.001 *** (0.000)	0.000 *** (0.000)
OLS: dummy	-0.001 *** (0.000)	-0.001 *** (0.000)	-0.001 *** (0.000)
Quantile 0.25: dummy	-0.001 ** (0.000)	-0.001 ** (0.000)	-0.000 (0.000)
Quantile 0.50: dummy	-0.001 ** (0.000)	-0.001 * (0.000)	-0.000 (0.000)
Quantile 0.75: dummy	-0.001 ** (0.001)	-0.001 ** (0.001)	-0.000 (0.000)
N	4,183	4,183	8,312

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$