Price dispersion and informational frictions: Evidence from Supermarkets Purchases*

Pierre Dubois† Helena Perrone‡

This Version: August 2015 - Incomplete

Abstract

Traditional demand models assume consumers are perfectly informed about product characteristics, including price. However, sales are a common supermarket practice. Using French data we show that retailers frequently change position in the price rankings, making it unlikely that consumers know all deals offered at each period. Further empirical evidence relating transaction prices, consumers' opportunity cost of time, and store visits is consistent with a model of imperfectly informed consumers that have to pay a cost to find out about prices. We develop such a model for horizontally differentiated products and structurally estimate the search cost distribution. Results show that in equilibrium consumers observe a very limited number of prices before making a purchase decision, implying that informational frictions are important and that local market power is potentially high. We also show that a demand model that assumes perfect information underestimates price elasticities.

Key Words: price dispersion, sales, imperfect information, search costs, product differentiation, consumer behavior, demand estimation, price-elasticites.

JEL Classification: D4, D83, L11, L66

*Previously circulated as "Price Dispersion and Search Costs: The Roles of Imperfect Information and Product Differentiation"

†Toulouse School of Economics (GREMAQ, IDEI). pierre.dubois@tse-fr.eu

‡Universitat Pompeu Fabra and Barcelona GSE. helena.perrone@upf.edu

1 Introduction

Utility parameters in demand for differentiated products are usually identified assuming consumers are perfectly informed about all product characteristics. However, unannounced short term reductions on price of certain products (sales) is a common and frequent pricing strategy of supermarkets. If consumers cannot perfectly predict the timing of sales at every store, they cannot know every period which deals are offered by the different stores. This means they have to incur a cost to find out about the prices being practised before making a purchase decision.

In this paper we investigate the importance of price informational frictions in explaining consumer food product choice. We first test wether observed behavior and prices are consistent with a model of imperfect price information. We find compelling evidence that price information is not freely and readily available without search. We then develop and estimate a demand model with imperfect information about prices where consumers sequentially search for the best prices.

The existence of informational frictions in the context of supermarket products has relevant implications for competition policy. In the presence of search costs, firm entry does not necessarily improve welfare. Stahl (1989) shows that an increase in the number of firms in the market may actually decrease welfare depending on how search costs are distributed among the population of consumers. Also, if search costs are important, firms may retain considerable market power even in seemingly competitive situations, which in some markets may justify price regulations. Rhodes (2014) for instance shows that multiproduct retailers set high prices because once consumers are at the store they are hold up due to high search costs.

From the supermarkets' point of view, search costs are relevant for its assortment decision (Cachon, Terwiesch and Xu, 2008). Search costs also play a role in Monetary Economics. Head, Liu, Menzio, and Wright (2010) use search frictions in the product market to explain price stickiness even in the absence of menu costs or other technological constraints to price adjustment.

Using food purchase data from France, we perform a number of empirical exercises to assess the relevance of price informational frictions. First, we look at price dispersion. When stores and consumers are identical and perfectly informed and there are no capacity constraints, the unique Nash equilibrium is the Bertrand outcome, with perfectly competitive prices. For price dispersion to arise, there must be some heterogeneity between consumers or stores and positive search costs.[1] Consistent with a large body of literature,[2] we show that price dispersion is

[1]See Stigler (1961), Salop and Stiglitz (1982), Burdett and Judd (1983), Stahl (1989), among others.

[2]See, for instance, Sorensen (2000), Lach (2002), Lewis (2008), and Giulietti, Otero and Waterson (2009). For a review of empirical studies of price dispersion online, see Baye, Morgan, and Scholten (2006).

prevalent in the French food market, even after controlling for observed and unobserved product characteristics. Second, we check for intertemporal price dispersion. If price information is not perfect, then price dispersion should persist over time. Otherwise, consumers would learn the identity of the cheapest store, and all other stores would have zero demand. We observe stores frequently changing rank position in the pricing distribution. Hence, consumers have to incur shopping or search costs to find out which deals are offered at the market each period.

Inspired by Aguiar and Hurst (2007), we study the correlation between prices paid and household's opportunity cost of time. If consumers are imperfectly informed about prices, then families with a high opportunity cost of time – and thus higher search costs – should pay higher prices, once we control for store and time period. Controlling for unobservable household and store characteristics, we find that households with a higher opportunity cost of time pay higher prices, and also have a lower probability of paying the lowest price. We also find that households with higher opportunity cost of time visit less stores per period, and that prices decrease with number of per period store visits.

It is hard to explain the empirical evidence we find considering a model where consumers freely observe all practised prices at every period. The evidence is consistent with a model of choice where price information is costly. At best, consumers need to pay at least the transportation cost up to the store to find out the price of the products that week. We model consumer choice with sequential search costs and develop an empirical strategy to identify the parameters of the search costs distribution, which we estimate along with the parameters of the utility function. Products are considered to be heterogeneous with both vertical and horizontal attributes. Hence, consumers search for the product with the highest indirect utility instead of the lowest price. Notice however that consumers are perfectly informed about the existence and all characteristics of the product other than price.

An important contribution of our paper is to identify search costs for food products sold in brick and mortar stores when products are horizontally differentiated and search is unobservable. The horizontal dimension is particularly important when dealing with physical (not online) stores. In this context, the relative geographical location of the store, which is consumer specific, is clearly an important characteristic affecting choices, and ignoring this differentiation dimension will bias estimated parameters.

There is a growing literature on estimation of consumer choice with search costs. Hong and Shum (2006), using only price data, identify search costs for books sold online. Also requiring only price data, Moraga-González and Wildenbeest (2006) identify search costs for personal computer memory chips, obtained from a web-based search machine. Honka (2014) develops a discrete choice model of demand with non-sequential consumer search applied to auto insurance contracts. Koulayev (2014) studies hotel web search. Moraga-González, Sandor and Wildenbeest (2011) study consumer choice of cars with search costs where consumers' consumption sets are endogenous.

Similar to our application, Wildenbeest (2011) studies search for grocery items in brick and mortar UK supermarkets. Although he allows for vertically differentiated products, there is no consumer taste heterogeneity. This is an important restriction when dealing with traditional stores since location is clearly an important horizontal attribute of stores which affects consumers preferences.

Another paper that studies search in supermarkets is Seiler (2013). His approach is quite different from our however: consumers visit a certain store for exogenous reasons and, once at the store decide whether to search for the price of a product or not. Search is therefore a binary decision of whether to walk down the product aisle.

Our model is based on Hortaçsu and Syverson (2004), who study the mutual fund industry. In their model, consumers have identical tastes, but have different search costs. Consumers sequentially search to find out about prices in the market each period. For the case of vertically differentiated products, identification is only possible when the sampling probabilities are equal across sellers. When these probabilities change according to observable characteristics of the seller, identification is possible only if products are considered homogenous. That is, it is not possible to identify search costs when both products characteristics and sampling probabilities are heterogeneous. We extend Hortaçsu and Syverson (2004)'s model to allow for heterogeneous consumer taste. Hence, products are both vertically and horizontally differentiated. We can identify search costs considering both product differentiation and unequal sampling probabilities.

Our approach is very flexible with respective to drawing probabilities. Most empirical models restrict the sampling probabilities to be the same across different households, periods, and stores Here, the probability of each price draw is household, store, and period specific. We extract these probabilities from observed shopping behavior. Since we observe all store visits made by households in a certain week, we know every time a household goes to a store and purchases something (any food product). Hence we can recover, for each household and period, the probability of visiting a certain chain store. This is another contribution of our work.

When shopping at supermarkets, consumers frequently purchase not an individual item but a basket of goods. Purchasing a basket of goods is an equilibrium response for the existence of search costs, which also include transportation costs. If search costs were zero, consumers would buy each good at the store offering it at the lowest price. It is true however that even perfectly informed about prices, consumers could decide to buy all products in only one store to minimize travelling. Arguably, modeling search for a food basket is a more realistic description of consumers' groceries shopping behavior than search for an single product. However, considering search for a basket of goods can be very complex especially in what concerns one stop versus multiple stop shopping. If a consumer is searching for a basket of goods, he may buy a part of the basket in the first visit, another part in the second visit and so on until he purchases all items in his shopping list. The theoretical problem allowing for multiple search and multiple stop shopping can get extremely complex and empirical identification when search behavior is unobservable may be very difficult to obtain. We are not aware of any paper in the literature that identifies the parameters of the search model in a context of multiple stop shopping. Wildenbeest (2011) considers search for a basket of food products. However, he does not allow for consumers to purchase different items of the basket in different stores. In this case, consumers are concerned only with the total price or the total utility of the bundle and the multiproduct problem degenerates into a single product model. Using data only on those purchases made under the same roof introduces selection bias and probably an overestimation of search costs because the estimation process will exclude the segment of consumers with low transportation costs (or low value of time) that are willing to shop around.

Notice that search costs in our setting are basically the transportation cost needed for the consumer to get to a new store and find out what prices are being offered that period. In which situation would transportation and search costs be different? If there is no imperfect information about prices, then search costs and transportation costs are not the same. That is, if before leaving home consumers know exactly which store is offering which deal at every week, then search costs are null, though transportation costs are not (consumers still have to travel to the store to make the purchase). In this case, though informed about it, consumers may not purchase the best deal available if they are unwilling to pay the transportation cost to get there.[3] We are not able to separately identify search costs and transportation costs. But we show compelling evidence that stores are frequently changing ranking positions in the distribution of prices, so it is very unlikely that consumers are perfectly informed about the identity of the cheapest store at each period.

Sticking to the empirical literature on search costs, we assume there is perfect recall, that is, consumers can go back to previously searched stores with no additional cost. This is an

[3] If consumers opted for home delievery, then we could abstract from transportation costs. However, there was very little home delievery in France at the time studied here.

important assumption when the number of stores is finite because it guarantees that the optimal stopping search rule is stationary. Without perfect recall, the marginal cost of search would depend on the search history and the number of non-sampled stores (Jansen and Parakhonyak, 2014), making identification extremely hard. When the number of stores is infinite, stationarity does not require perfect recall. In our application, there is a limited number of stores consumers can search from. But we could assume that consumers see the number of potential price draws as large since in principle they could continue searching the same stores over time.

In our model consumers search sequentially for the best deal. This means, at every price draw, they decide whether to continue searching or not by comparing the cost of an additional search with the expected benefit of an additional search. An alternative to sequential search is to assume non-sequential or fixed-sample search. In this case, before leaving home, consumers decide on how many stores to visit before making a purchase decision. Hence, the number of price draws is independent of the realizations of each draw.

De Los Santos, Hortaçsu and Wildenbeest (2012) test between a sequential and a non-sequential search model using data on web browsing and purchase behavior. They find evidence that non-sequential search, or fixed sample search, provides a better description of how consumers search for books online. We believe this is a less credible search protocol for the case of brick and mortar supermarkets. It would imply for instance that even when consumers find a very good deal for an item in an early draw, they stick to the initial plan of visiting a fixed number of stores. It seems more realistic to assume that consumers leave home with their shopping list and once they find one of the items at a price lower than a certain reservation price, they buy the item at that visit and stop searching, which is the behavior implied by the sequential search assumption.

The empirical investigation is performed on a comprehensive individual level dataset which includes every food product purchased by a representative survey of French households during 3 years, 1999, 2000, and 2001. We have information on product and store characteristics, as well as household demographics. We focus on 4 product categories: beer, coffee, cola, and whisky. Results from the demand model show that most consumers (around 90%) obtain at most three utility quotes before purchasing a product. A large proportion of consumers observe only the price of the product they actually purchase (that is, they do not draw any other prices). Notice that this is indicative of important information frictions. Indeed, if price information is freely available (in other words, consumers are perfectly informed), we should find that consumers observe all prices -pay the cost of getting informed about all products' prices/search all products– before making a purchase choice. Also, relatively high local market power of firms. We also calculate the own-price elasticities implied by the demand model with imperfect information and compare them to the own-price elasticities on a multinomial logit model. We find the perfect information model consistently overestimates own-price elasticities.

The paper is organized as follows. In Section 2, describes the data and the products used in the analysis, while Section 3 presents the reduced-form tests. In Section 4 we describe the model of consumer choice behavior with sequential search and the empirical identification strategy. Results of estimation of the search cost distribution and price-elasticities are in Section 5. Finally, the last section concludes and discusses extensions.

2 Data and Product Choice

The data set is a representative survey of households distributed across all regions of France. We have information on three years: 1999, 2000, and 2001. Household register every food product purchased using a scanner. For each product purchased, we have information on its brand and characteristics, including price and pack size, label, the date of the purchase and the brand and square foot store size where it was purchased. We also have comprehensive information on household demographics.

We study 4 product categories, namely beer, cola, coffee, and whisky. We consider relatively cheap and relatively expensive categories, e.g., cola and whisky, to check if price dispersion decreases with average price. We also compare categories that are "necessities" and frequently consumed (e.g., coffee) against "luxury" items (e.g., cola, beer, whisky).

Household characteristics used in the analysis include number of store visits per household per week, age of the household head, a dummy variable indicating the presence of a baby (a child of less than 4 years old) in the household, the education level of the household head, the household size, a dummy indicating whether the household is rich (in the upper half of the income distribution) and a dummy indicating whether the household head is professionally inactive. The variable indicating whether the household head is inactive is equal to one if the household head is either a student, retired, in long term unemployment, or has no professional activity. The variable *rural* indicates whether the household lives in a rural area. The education level variable is organized in six levels, depending on diploma of the household head, starting with no diploma (level 0). this information is missing for some of the households in the sample, thus the lower number of observations.

Table 1 brings some descriptive statistics on the purchased quantity per purchase occasion and product category, as well as on household characteristics. Quantities are measured in milliliters, except for coffee which is measured in grams. The product less frequently purchased is whisky. All other products are very frequently purchased. Households visit on average 1.5 stores per week. But this number can go up to 10 visits a week.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Stores visits per week	1.571	0.783	1	10	531,687
Age of household head (years)	47.90	13.75	18	94	531,687
Baby	0.136	0.343	0	1	531,687
Inactive	0.402	0.49	0	1	531,687
Rural	0.537	0.499	0	1	531,687
Rich	0.353	0.478	0	1	531,687
Household size	3.21	1.39	1	9	531,687
Education level	2.35	1.32	0	5	445,038
Quantity per purchase					
Beer (liters)	3.255	2.695	0.500	54	45,023
Coffee (kg)	0.473	0.301	0.250	6	122,362
Cola (liters)	3.367	3.067	0.330	72	86,127
Whisky (liters)	0.785	0.236	0.700	5	7,642

3 A Map of Price Dispersion in French Supermarkets

We look at price dispersion by focusing on the prices of two tightly defined products within each category. Our product definition allows for only one source of differentiation, i.e., the store where they are purchased. So, for example, within the Cola category, a product is defined by its brand, whether it comes in a bottle or in a can, the pack size, whether it is light cola, and the size of the bottle or can. We select products with a high market share, sold at a large number of stores. Hence, among each category, we choose the product most frequently purchased. The second product chosen in a category is a product that, among those most frequently purchased, has an average price clearly higher or lower than the first product chosen.[4]

[4] We consider the following products (we do not name brands but use A and B to signal they are different brands): (i) beer brand A, bottle size: 250 ml , pack: 24 bottles; and beer brand B, bottle size: 250 ml , pack: 10 bottles; (ii) coffee brand A, no "gamme", arabica, caffeinated, 1 package per pack, package size 250g; and coffee brand B, degustation, arabica, caffeinated, 1 package per pack, package size 250g; (iii) cola brand A, plastic bottle, non light, bottle size: 1500 ml, pack: 1 bottle; and cola brand B, plastic bottle, non light, bottle size: 1500 ml, pack: 4 bottle; and (iv) whisky brand A, 1 liter bottle, no age, blended; and whisky brand B, 1 liter bottle, 5 years of age, blended.

Table 2 brings some descriptive statistics on the price distributions. The first four columns bring respectively the average price per liter in the case of liquids or per kilo in the case of coffee, the coefficient of variation, and the ratio of the third to the first quartile, as well as the ratio of the 95% to the 5% quantile. Here there is no control for store heterogeneity (or period). Hence, even if we are dealing with tightly defined products, they may still be differentiated because they embed potentially differentiated characteristics of the store where they were purchased.

Part of the price dispersion observed in prices may be explained by product differentiation and time variation. Although we compare goods with exactly the same physical attributes, they are sold at different stores and different time periods, which means that the products are not homogenous. To clear prices from the heterogeneity due to the store and period, we run product by product regressions of prices, measured as log deviations from the weekly mean, on month and year fixed effects, supermarket chain dummies, store type, and regional dummies. The residuals of these regressions are considered to represent the price of a homogeneous product. This method, which allows to obtain measures of the price of the common attributes of the good (all attributes excluding store and period), has become standard in the literature (see Lach, 2002, Zhao, 2006, and Sorensen, 2000, among others). However these "residual prices" implicitly assume that final log-prices are a linear combinations of the prices of individual attributes (the sum of the price of the homogenous product and the price of the differentiated services offered by the retailer).

The four last columns of Table 2 show some descriptive statistics for the dispersion of the residuals of the fixed effects regressions described above. It includes weekly averages of the first and fourth quartiles, and differences between the first and fourth quartile and the 95% and 5% quantile (mean value of the residuals are zero by construction).

The statistics show that the price dispersion is important in all categories considered. Indeed, looking at the interquartile ratio, we see that 50% of prices in the middle of the distribution differ up to 34%. This difference is less important for whisky (1%), which is the most expensive product under study. This is consistent with the idea that search costs have a fixed cost component. In the case of expensive products, the search cost is low relatively to the price of the good and consumers have more incentives to search more intensively for the best price. Since more search is undertaken, consumers are better informed about shelf prices, forcing stores towards the Bertrand equilibrium, in which price dispersion is minimal.

Controlling for observed and unobserved fixed product characteristics decreases price dispersion for all products but whisky A, for which price dispersion goes up, probably due to mispecification. Price differences remain on average high for many products. In terms of interquartile differences, the difference in prices in the middle of the price distribution goes up to 29%.

Table 2: Descriptive Statistics of Pricing Patterns

		Uncontrolled Price Dispersion					Residual Price Dispersion			
Product		Avg Price	Standard Deviation	Coef of Variation	Q75/Q25	Q95/Q5	Q25	Q75	Q75-Q25	Q95-Q5
Beer	A	0.98	0.10	0.10	1.21	1.27	-0.06	0.06	0.12	0.23
	B	1.30	0.10	0.07	1.04	1.26	-0.01	0.02	0.03	0.14
Cola	A	0.705	0.05	0.07	1.08	1.18	-0.10	0.11	0.21	0.48
	B	0.494	0.11	0.23	1.28	1.81	-0.10	0.11	0.21	0.48
Coffee	A	8.73	0.64	0.07	1.10	1.22	-0.03	0.02	0.06	0.16
	B	4.49	0.77	0.17	1.34	1.70	-0.17	0.12	0.29	0.45
Whisky	A	15.5	1.65	0.11	1.01	1.44	-0.04	0.03	0.06	0.21
	B	14.9	0.63	0.04	1.02	1.05	-0.01	0.01	0.02	0.06

Notes: Prices are in levels in the left panel, and in logs in the right panel (that is why we calculate the difference between the third and first quartile, not the ratio, as in the left.)

3.1 Temporal Price Dispersion

If stores position in the price distribution remain constant, then consumers can learn the identity of the cheapest store and there is no imperfect information about prices. That is, before leaving home, consumers know where to find the best deal. They do not have to incur in the cost of visiting a store to find out which price is being offered. But if stores periodically change position in the ranking of prices, i.e., there is intertemporal price dispersion, then consumers cannot learn before hand which deals stores are offering at a given period. Before start shopping, which has costs, they do not know which store is offering which price. The existence of temporal price dispersion is therefore direct evidence of the importance of informational frictions, and it does not depend on any restriction on consumer search behavior.[5]

To study the temporal price dispersion, we look at the position of stores in the cross sectional price distribution, and check whether stores frequently change position in this ranking over time. Notice that if store services are an important part of the price of the otherwise homogenous product, we could observe low transition probabilities even if the price of the homogeneous

[5] Standard tests of the importance of search costs consist in regressing measures of price dispersion on proxies of search costs (or search benefits) and on the number of firms in the market. However, Chandra and Tappata (2009) show that the relationship between price dispersion and search costs or the number of firms on the supply side is not necessarily monotone (and not always positive), depending crucially on restrictive assumptions on the consumers' search strategy.

product is constantly changing positions in the price ranking. The transition probabilities for the uncontrolled price should therefore be interpreted as a lower bound for the movements of the homogeneous product in the cross sectional ranking.

Table 3: Stores Transition Probabilities in the Price Ranking

Product	Beer A				Beer B			
Rank t \ Rank at $t+1$	1	2	3	4	1	2	3	4
1	0.472	0.204	0.163	0.120	0.484	0.230	0.138	0.103
2	0.288	0.293	0.234	0.142	0.332	0.361	0.214	0.077
3	0.195	0.214	0.320	0.235	0.15	0.199	0.423	0.198
4	0.189	0.118	0.281	0.341	0.129	0.097	0.200	0.523
Obs	2,642				4,489			
Product	Coffee A				Coffee B			
Rank t \ Rank at $t+1$	1	2	3	4	1	2	3	4
1	0.583	0.071	0.161	0.146	0.408	0.198	0.200	0.157
2	0.415	0.171	0.220	0.111	0.335	0.264	0.255	0.134
3	0.403	0.098	0.273	0.201	0.235	0.265	0.287	0.191
4	0.351	0.059	0.208	0.337	0.192	0.240	0.304	0.192
Obs	4,558				1,392			
Product	Cola A				Cola B			
Rank t \ Rank at $t+1$	1	2	3	4	1	2	3	4
1	0.641	0.143	0.119	0.080	0.497	0.197	0.161	0.067
2	0.264	0.467	0.132	0.113	0.353	0.242	0.170	0.104
3	0.235	0.193	0.375	0.184	0.303	0.250	0.230	0.143
4	0.121	0.106	0.164	0.587	0.278	0.185	0.166	0.232
Obs	12,740				1,237			
Product	Whisky A				Whisky B			
Rank t \ Rank at $t+1$	1	2	3	4	1	2	3	4
1	0.438	0.200	0.182	0.117	0.522	0.159	0.147	0.083
2	0.286	0.289	0.231	0.131	0.408	0.185	0.223	0.076
3	0.233	0.185	0.313	0.202	0.202	0.174	0.261	0.206
4	0.181	0.177	0.247	0.301	0.142	0.162	0.311	0.284
Obs	1,586				862			

Table 3 shows the average (across periods and stores) of the Markovian transition probabilities for each product. At each week, we assign stores to one of the four price intervals limited by the quartiles of the price distribution.[6] The transition probabilities in the table are the empirical probabilities of changing from position $j$to position k, with $j = 1, 2, 3, 4$and $k = 1, 2, 3, 4$. When the price for a certain store in a certain period is not observed, the transition probability (for that store and that period) is considered missing.

For all products, the probability of remaining in the first position is higher than the other transition probabilities, averaging between 1/3 and 1/2. The probability of remaining in the same position from one week to the next varies between around 25 to 40 percent. There is lot of movement in the price ranking, with stores frequently changing position over a very short period of time. The evidence shows that it is hard for consumers to keep track of which stores are offering the best deals each week.

3.2 Prices, Store Visits, and the Opportunity Cost of Time

If consumers have to incur a cost to find out about prices, households with a higher opportunity cost of time will shop around less and will, on average, pay higher prices for otherwise identical products controlling for store and period of purchase. A positive correlation between the cost of time and prices paid is therefore evidence that imperfect information about prices importantly affects the demand behavior of consumers.

Notice that this test does not rely on any particular assumptions on consumer search protocol or stores pricing strategies. If information on the best deal is not readily and freely available, then consumers with a high opportunity cost of time have a lower probability of finding the best deals (because they will search less) and will on average pay higher prices. The same reasoning applies for the probability of paying the lowest price available in the market that period. Consumers that search less due to a higher opportunity cost of time will have a lower probability of finding the best deal.

However, the test does rely on store and time fixed effects correctly controlling for potential quality differences between products purchased at different locations and periods. If there are remaining differences between the products, then a correlation between price and opportunity cost of time could be revealing that households with a high opportunity cost of time prefer purchasing at stores which offer more expensive services but not that search costs are relevant.

[6] That is, if a store is in position 1 at t, it means that its price at t is lower or equal to the first quartile of the price distribution at that period. The store is in position 2 if its price is between the first and second quartiles, and in position 3 if its price is between the second and third quartiles. Finally, the store is in position 4 if its price is greater than the third quartile.

In order to have comparable prices across products and product categories, we use a price index where product prices are weighted by total quantity purchased that period. The price index captures how much more or less than the average the household is paying for a given product. We also regress the probability of paying the lowest price for a product on opportunity cost of time. In this case, the dependent variable is a binary variable equal to 1 if the price paid by household $i$at period $t$for product $j$is equal to the lowest price (with a 2% margin) observed in the market that period.

Table 4: Opportunity Cost of Time and Prices

	Transaction price		Paying lowest price	
	(OLS)		(Probit)	
Frequency	0.0017***	(5.03)	-0.0008***	(-12.74)
Age	-0.0273***	(-4.17)	0.0103***	(8.07)
Age square	0.0003***	(4.04)	-0.0001***	(-7.85)
Baby	0.1056***	(2.98)	-0.0199***	(-2.90)
Inactive	-0.0635**	(-2.51)	0.0359***	(7.33)
Rural	-0.0470**	(-2.10)	-0.0143***	(-3.33)
Rich	-0.0376	(-1.58)	0.0089*	(1.91)
Hh size = 2	-0.3648***	(-3.77)	0.0819***	(4.29)
Hh size = 3	-0.3585***	(-3.66)	0.0682***	(3.51)
Hh size = 4	-0.6324***	(-6.47)	0.0865***	(4.47)
Hh size = 5	-0.5567***	(-5.59)	0.0658***	(3.33)
Hh size = 6	-0.5408***	(-4.93)	0.0724***	(3.32)
Hh size = 7	-1.1012***	(-7.03)	0.1299***	(4.44)
Hh size = 8	-1.3021***	(-6.56)	0.2591***	(5.78)
Hh size = 9	-2.7904***	(-7.22)	0.3103***	(4.49)
Education level 1	-0.1227**	(-2.57)	0.0236**	(2.54)
Education level 2	0.0010	(0.02)	0.0026	(0.32)
Education level 3	-0.0698	(-1.46)	0.0405***	(4.38)
Education level 4	-0.1755***	(-2.99)	0.0767***	(6.82)
Education level 5	-0.1211**	(-2.33)	0.0180*	(1.81)
Store Fixed Effects	Yes		Yes	
Time Fixed Effects	Yes		Yes	
Constant	101.7492***	(216.10)	0.8956***	(11.21)
N	445,038		415,405	

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The household characteristics used to capture the opportunity cost of time are age and age square, education level and professional activity status of the household head (the variable *inactive* is equal to 1if the household head is a student, retired or unemployed), socioeconomic class (the variable *rich* is equal to 1if the household is in the upper half of the income distribution), presence of a baby of less than 48months, and household size. We also include controls for region of residence, frequency of purchases, and number of stores visited. Furthermore, we include controls for the store brand where the purchase was made, the size of the store, and the period of purchase.

Table 4 shows the empirical results. The first column shows the coefficient estimates of the OLS regression where the dependent variable is the price index of the product (by household and period). The second column shows probit model estimates where the dependent variable is a discrete variable equal to 1 if the household paid the lowest price practised in the regional market at that period.

Table 5: Stores Visits and Prices

	Transaction price	Paying the lowest price
number of store visits	-0.0007***	0.0096***
	(-4.60)	(5.37)
more than 1 visit		0.0021
		(0.77)
Time Fixed Effects	Yes	Yes
Product Fixed Effects	Yes	Yes
Constant	1.0010***	0.8305***
	(1079.96)	(126.57)
N	531,687	531,687

Notes: t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

There is strong empirical evidence supporting the hypothesis that households with a higher opportunity cost of time pay higher prices, even after controlling for store heterogeneity. The number of stores visited in a week reduces prices paid and increases the probability of paying the lowest price in the market that week, implying that there are gains from shopping around. The relationship between age and prices paid is U-shaped. Prices paid increase on average with having a baby. Accordingly, the probability of paying the lowest price decreases with having a baby. Being inactive, living in a rural area, and being rich reduce prices paid (and increase the probability of paying the lowest price). Finally, prices paid decrease with household size.

We also run regressions of prices and the probability of paying the lowest price in a certain week on the number of stores visited (by household $i$in week t) and a dummy variable indicating wether the household visited more than one store that week. Table 5 shows that visiting more than one store is negatively correlated with prices paid. Prices are also negatively correlated with the number of store visits.

Table 6: Determinants of the number of stores visited

	Store Visits		Store Visits		Store Visits>1	
	(OLS)		(Poisson)		(Probit)	
Age	0.0174***	(-22.39)	0.0285***	(-21.86)	0.0243***	(17.78)
Age square	-0.0002***	(-21.02)	-0.0003***	(-20.34)	-0.0002***	(-17.09)
Baby	-0.0068	(-1.52)	-0.004	(-0.57)	-0.0077	(-1.03)
Inactive	0.0694***	(-20.61)	0.1014***	(-21.12)	0.1203***	(22.50)
Rural	-0.0830***	(-29.69)	-0.1244***	(-29.81)	-0.1080***	(-24.14)
Rich	-0.0201***	(-6.50)	-0.0307***	(-6.50)	-0.0163***	(-3.23)
Hh size = 2	0.0265**	(-2.05)	0.0522**	(-2.17)	0.1110***	(5.43)
Hh size = 3	0.0660***	(-5.09)	0.1147***	(-4.76)	0.1565***	(7.53)
Hh size = 4	0.1123***	(-8.74)	0.1900***	(-7.95)	0.2483***	(12.01)
Hh size = 5	0.1961***	(-14.85)	0.3065***	(-12.69)	0.3333***	(15.81)
Hh size = 6	0.1941***	(-12.98)	0.3008***	(-11.67)	0.3181***	(13.53)
Hh size = 7	0.3831***	(-17.93)	0.4996***	(-17.2)	0.6224***	(19.50)
Hh size = 8	0.2323***	(-6.85)	0.3410***	(-7.93)	0.3338***	(6.85)
Hh size = 9	1.0660***	(-13.73)	1.0038***	(-20.86)	(0.8971***	(11.58)
Education level 1	-0.0475***	(-7.26)	-0.0684***	(-7.43)	-0.0748***	(-7.38)
Education level 2	-0.0562***	(-9.80)	-0.0801***	(-10.14)	-0.0858***	(-9.78)
Education level 3	-0.0639***	(-9.92)	-0.0921***	(-10.04)	-0.0832***	(-8.26)
Education level 4	-0.1171***	(-15.65)	-0.1812***	(-15.63)	-0.1728***	(-14.04)
Education level 5	-0.0722***	(-10.67)	-0.1065***	(-10.95)	-0.0792***	(-7.38)
Time Fixed Effects	Yes		Yes		Yes	
Product Fixed Effects	Yes		Yes		Yes	
Constant	1.1682***	(46.34)	-0.7165***	(-16.83)	-0.8386***	(-19.86)
N	338,961		338,961		338,961	

Notes: t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first two columns of Table 6 show coefficient estimates of the regression of the number of stores visited on opportunity cost of time. In column 1, coefficients were estimated using OLS, whereas in column 2, we used a Poisson model. The third column brings results from a regression of the probability of visiting more than 1 store per week on opportunity cost of time. Age and store visits have an inverted U- shape relationship, implying that younger and older consumers visit more stores in a week. Having a baby, living in a rural area, and being rich decrease the number of store visits and the probability of visiting more than one store. Being inactive and household size increases the number of store visits and the probability of visiting more than one store. The greater the number of stores the consumer visits, more information he gathers on transaction prices and sales at that period.

4 Consumer Behavior with Search Costs

We have shown evidence consistent with imperfect information about prices, which means consumers have to incur a cost to find out which prices are being practised at the market at a certain period of time. We now consider a model with choice with costly search and develop an empirical strategy to estimate the search cost distribution and the parameters of the choice model when consumers are imperfectly informed.

4.1 Model of Consumer Behavior

Our demand model with search costs is close to Hortaçsu and Syverson (2004)' s extension of the framework developed by Carlson and McAfee (1983). However, in their model consumers are identical except for search costs, whereas we allow for observable heterogeneity in preferences. This means that in our model products are both vertically and horizontally differentiated (i.e., consumers do not all agree on the value of each product attribute). The horizontal dimension will be related to some observable characteristics of consumers. We allow for a number $I$of consumer types in terms of their valuation of the product.

Consumers purchase at most one unit of the product. Before purchasing, consumers sequentially search for the product with the highest indirect utility. Search is costly and its cost is heterogeneously distributed across the population of consumers. The cost of the first quote is normalized to zero. This ensures that everyone willing to purchase a product will do so. The indirect utility of a consumer of type i from purchasing product j at period t is denoted u_{ijt}. Notice that within each i, search costs may vary, though valuations may not. We assume consumers search with replacement.

Let $F_i(\cdot)$ be the belief distribution of indirect utilities u_{ijt} of a type i consumer. Then, the optimal search rule for consumer i with search cost c_i who already found a highest (among past searches) indirect utility u^*_{it} is to search once more if the sunk cost of searching is lower than the expected utility gain conditional on finding a better alternative:

$$c_i \leqslant \int_{u^*_{it}}^{\bar{u}_{it}} (u - u^*_{it})\, dF_{it}(u) \tag{1}$$

where $\bar{u}_{it}$ is the upper bound of the support of $F_{it}(\cdot)$. The above condition means that the marginal cost of searching one more time is smaller than or equal to the expected gain of searching one more time given u_{it}^*.

We assume consumers know $F_{it}(\cdot)$, which means that they know the support of the distribution of indirect utilities so that they can label the N_i available products in ascending order with respect to the indirect utility: $u_{i1t} < u_{i2t} < ... < u_{iN_it}$. For simplicity, we assume that there are no two products (stores) that provide the same indirect utility. Notice that we index the number of available products N_i by the consumer type i. This is to make clear that consumers do not necessarily have access to the same products since they shop at different markets. Notice also that we index the ranking of indirect utilities ($j = 1, .., N_i$) by t, the period of the purchase, since the ranking may change from one period to the next.

As all indirect utilities of stores are strictly different, we get that:

$$F_i(u) \equiv P(u_{ijt} \leq u) = \sum\nolimits_{k=1}^{N_i} \phi_{ik} \mathbf{1}_{\{u_{ikt} \leq u\}} \tag{2}$$

where ϕ_{ik} is the probability that the store ranked k is sampled by consumer i (this probability belief is known by consumers and common to all consumers of type i) and $\mathbf{1}_{\{u_{ikt} \leq u\}} = 1$ if and only if $u_{ikt} \leq u$.

Using (1) and (2), this yields the following cut-off points on the search cost distribution:

$$c_{ij}^t \equiv \sum\nolimits_{k=j+1}^{N_i} \phi_{ik} (u_{ikt} - u_{ijt}) \tag{3}$$

where c_{ij}^t is the search cost level that makes any consumer of type i indifferent between purchasing at store j and searching once more (i.e., it is the lowest possible search cost of any type i consumer who purchases product j). Given that j is already quoted, products with indirect utility lower than u_{ijt} do not enter the calculation of the expected gain of searching once more (right-hand side of the above equation).

Remark that although search costs are assumed to be time invariant, cut-off points will depend on the period of purchase. Notice also that $c_{iN_i}^t = 0$ and that the expected gain of an additional search decreases with the index of the product, so that $c_{i1}^t > .. > c_{iN-1}^t > c_{iN}^t = 0$.

Then, a consumer will purchase the worst product if he searches only once and finds the lowest utility product in his first and only draw. The probability of sampling the product $j = 1$ on the first draw is equal to ϕ_{i1}. Therefore, the demand for the lowest indirect utility product at period t, from type-i consumers is equal to:

$$q_{i1}^t = \phi_{i1} \left[1 - G\left(c_{i1}^t\right)\right] \tag{4}$$

where G is the cumulative distribution function of search costs. Remark that no type i consumer whose search cost is below c_{i1}^t will end up purchasing product 1 since it is always better for him (in expected utility) to search once more if he has only u_{i1t}. Following the same kind of reasoning, we get (Hortaçsu and Syverson (2004)):

$$q_{i2}^t = \phi_{i2}\left[1 + \frac{\phi_{i1} G\left(c_{i1}^t\right)}{1-\phi_{i1}} - \frac{G\left(c_{i2}^t\right)}{1-\phi_{i1}}\right] \tag{5}$$

and for $j = 3, .., N_i$:

$$q_{ij}^t = \phi_{ij}\left[\sum\nolimits_{k=1}^{j} \frac{G\left(c_{ik-1}^t\right) - G\left(c_{ik}^t\right)}{1-\sum\nolimits_{l=0}^{k-1}\phi_{il}}\right] \tag{6}$$

which can be re-written

$$q_{ij}^t = \phi_{ij}\left[1 + \sum\nolimits_{k=1}^{j-1} \frac{\phi_{ik} G\left(c_{ik}^t\right)}{\left(1-\sum\nolimits_{l=0}^{k}\phi_{il}\right)\left(1-\sum\nolimits_{l=0}^{k-1}\phi_{il}\right)} - \frac{G\left(c_{ij}^t\right)}{1-\sum\nolimits_{l=0}^{j-1}\phi_{il}}\right] \tag{7}$$

where by convention $G\left(c_{i0}^t\right) = 1$ and $\phi_{i0} = 0$.

4.2 Identification of the Search Costs

Let's first assume that we know the probabilities of finding a store ϕ_{ij}. If we observed indirect utility rankings, then it is straightforward to get $G(c_{ij}^t)$, for $j = 1, .., N_i$, by solving the system of equations (4) to (6) above, using observed purchases in store j and the probabilities ϕ_{ij}. The problem is that the support of the indirect utilities is unknown to the econometrician. In the case where there is only price differentiation, these consumer specific rankings are simply price ranks (Hortaçsu and Syverson (2004)) which are observed. Otherwise, we cannot use price ranks to rank the indirect utilities of each alternative that depend on consumer preferences for horizontally differentiated goods.

Although, we observe the set of demands q_{ij}^t and choice probabilities ϕ_{ij} for every store, but do not observe the identity of the j^{th} store because we don't know the position of the store in the type i specific utility ranking at period t, we can however notice that (7) implies

$$\frac{q_{ij}^t}{\phi_{ij}} - \frac{q_{ij-1}^t}{\phi_{ij-1}} = \frac{G(c_{ij-1}^t) - G(c_{ij}^t)}{1-\sum\nolimits_{l=0}^{j-1}\phi_{il}} > 0$$

which means that

$$0 < \frac{q_{i1}^t}{\phi_{i1}} < \frac{q_{i2}^t}{\phi_{i2}} < ... < \frac{q_{iN_i}^t}{\phi_{iN_i}}$$

Hence, if we know quantities q_{ij}^t and the probabilities of finding a store ϕ_{ij}, we know the elements of the vector of ratios $\left\{\frac{q_{ij}^t}{\phi_{ij}}\right\}_{j=1,..,N_i}$ and can then recover the ranking of indirect utilities of the different stores for each type i.

Knowing $\frac{q_{ij}^t}{\phi_{ij}}$ and ϕ_{ij}, we can solve the following triangular system in the unknowns $G(c_{ij}^t)$ provided all choice probabilities $\phi_{ij} \in (0,1)$:

$$\begin{cases} \frac{q_{i1}^t}{\phi_{i1}} = 1 - G\left(c_{i1}^t\right) \\ \frac{q_{i2}^t}{\phi_{i2}} = 1 + \frac{\phi_{i1}}{1-\phi_{i1}} G\left(c_{i1}^t\right) - \frac{1}{1-\phi_{i1}} G\left(c_{i2}^t\right) \\ \frac{q_{i3}^t}{\phi_{i3}} = 1 + \frac{\phi_{i1}}{1-\phi_{i1}} G\left(c_{i1}^t\right) + \frac{\phi_{i2}}{(1-\phi_{i1})(1-\phi_{i1}-\phi_{i2})} G\left(c_{i2}^t\right) - \frac{1}{(1-\phi_{i1}-\phi_{i2})} G\left(c_{i3}^t\right) \\ .. \\ \frac{q_{ij}^t}{\phi_{ij}} = 1 + \sum_{k=1}^{j-1} \frac{\phi_{ik} G\left(c_{ik}^t\right)}{\left(1-\sum_{l=0}^{k} \phi_{il}\right)\left(1-\sum_{l=0}^{k-1} \phi_{il}\right)} - \frac{G\left(c_{ij}^t\right)}{1-\sum_{l=0}^{j-1} \phi_{il}} \end{cases}$$

which gives:

$$\begin{cases} G\left(c_{i1}^t\right) = 1 - \frac{q_{i1}^t}{\phi_{i1}} \\ G\left(c_{i2}^t\right) = 1 - q_{i1}^t - (1-\phi_{i1}) \frac{q_{i2}^t}{\phi_{i2}} \\ \ldots \\ G\left(c_{ij}^t\right) = 1 - \sum_{k=1}^{j} \left(\frac{q_{ik}^t}{\phi_{ik}} - \frac{q_{ik-1}^t}{\phi_{ik-1}}\right)\left(1 - \sum_{l=0}^{k-1} \phi_{il}\right) \end{cases} \tag{8}$$

The above system enables identification of the value of the search cost cumulative distribution function evaluated at the cutoffs points, that is $G(c_{ij}^t)$ for $j = 1, .., N_i$.

If we know the cumulative distribution function G and if it has non zero density on its support, we can invert G and identify c_{ij}^t by

$$c_{ij}^t = G^{-1}\left(G(c_{ij}^t)\right)$$

and also identify the indirect utilities up to a constant by solving the system of equations given by (3). Let's then assume that $G(.)$ belongs to a known family of c.d.f. parameterized by θ and denoted $G(., \theta)$. The lowest utility at each period is normalized to zero. Hence, there are $N_i - 1$ equations and $N_i - 1$ unknown values u_{ik}^t for each type i consumer:

$$u_{ikt} = \frac{c_{i1}^t}{\sum_{j=2}^{N} \phi_{ij}} + \sum_{k'=2}^{k-1} \frac{\phi_{ik'}^t c_{ik'}}{\left(\sum_{j=k'+1}^{N} \phi_{ij}\right)\left(\sum_{j=k'}^{N} \phi_{ij}\right)} - \frac{c_{ik}^t}{\sum_{j=k}^{N} \phi_{ij}} \tag{9}$$

for $k = 2, .., N_i$ and $u_{i1}^t = 0$.

Indirect utilities u_{ij}^t depend on joint characteristics of the consumer and store denoted x_{ijt}, common parameters γ_t, price p_{jt} and a consumer-store specific random deviation to mean utility v_{ijt} such that:

$$u_{ijt} = x_{ijt}\beta + \gamma_t - \alpha p_{jt} + v_{ijt} \tag{10}$$

In practice, x_{ijt} are observable characteristics of the store (that may vary with the identity of the consumer like distance to the store for example), γ_t are time-period fixed effects, p_{jt} is the price paid for product j at period t, α and β are parameters. Consumers' valuation of products characteristics has both horizontal and vertical dimensions.

The probability ϕ_{ij} that a consumer of type i finds store j can be identified empirically from the observation of all store visits made by the different households. Using data on all searches that generated a positive purchase of at least one product category (not only the categories we focus on) allows such identification.[7]

Then, (8) gives

$$c_{ij}^{t}(\theta) = G^{-1}\left(1-\sum\nolimits_{k=1}^{j}\left(\frac{q_{ik}^{t}}{\phi_{ik}}-\frac{q_{ik-1}^{t}}{\phi_{ik-1}}\right)\left(1-\sum\nolimits_{l=0}^{k-1}\phi_{il}\right),\theta\right)$$

Using (9), we have

$$u_{i2t}(\theta) = \frac{c_{i1}^{t}(\theta)-c_{i2}^{t}(\theta)}{\sum_{j=2}^{N}\phi_{ij}}$$

and for $k \geq 3$,

$$u_{ikt}(\theta) = \frac{c_{i1}^{t}(\theta)}{\sum_{j=2}^{N}\phi_{ij}} + \sum\nolimits_{k'=2}^{k-1}\frac{\phi_{ik'}c_{ik'}^{t}(\theta)}{\left(\sum_{j=k'+1}^{N}\phi_{ij}\right)\left(\sum_{j=k'}^{N}\phi_{ij}\right)} - \frac{c_{ik}^{t}(\theta)}{\sum_{j=k}^{N}\phi_{ij}}$$

Remark that by construction $u_{ikt}(\theta) > u_{ik-1t}(\theta)$ because

$$u_{ikt}(\theta) - u_{ik-1t}(\theta) = \frac{c_{ik-1}^{t}(\theta)-c_{ik}^{t}(\theta)}{\sum_{j=k}^{N}\phi_{ij}} > 0$$

As some endogeneity problem may generate a correlation between v_{ijt} and the prices p_{jt}, we cannot use an orthogonality condition between prices and unobserved shocks v_{ijt}. Instead, we assume that we observe some instrumental variables z_{ijt} that are uncorrelated with v_{ijt}, such that (10) gives the following moment condition

$$E\left[\left(u_{ijt}(\theta) - x_{ijt}\beta - \gamma_t + \alpha p_{jt}\right) z_{ijt}\right] = 0$$

Using this moment condition allows us to identify parameters θ, β, γ_t, α provided the usual rank condition for GMM.

5 Estimation and Empirical Results

5.1 Search Cost and Preferences Estimates

We define consumer types using the region of residence, whether they live in a urban or rural area, and wether their income is above or below the median income ("rich" or "poor", respectively). Hence, there are four types of consumers per region: poor and urban, poor and rural, rich and urban, and rich and rural.

[7] Another route would be estimate the probability of sampling a certain store by assuming a parametric distribution as in Hortaçsu and Syverson (2004).

Table 7: C.d.f. value estimates at search cutoffs points

Product	Household Type (i)		G1	G2	G3	G4	G5
Beer							
	urban	poor	0.569	0.151	0.090	0.053	0.031
		rich	0.485	0.099	0.053	0.030	0.016
	rural	poor	0.551	0.128	0.074	0.0423	0.023
		rich	0.456	0.087	0.043	0.021	0.010
Coffee							
	urban	poor	0.640	0.218	0.151	0.101	0.066
		rich	0.560	0.131	0.073	0.041	0.022
	rural	poor	0.631	0.216	0.147	0.102	0.071
		rich	0.532	0.106	0.057	0.030	0.015
Cola							
	urban	poor	0.551	0.135	0.081	0.045	0.024
		rich	0.451	0.054	0.024	0.012	0.005
	rural	poor	0.599	0.174	0.117	0.079	0.052
		rich	0.487	0.073	0.032	0.016	0.009
Whiskey							
	urban	poor	0.263	0.021	0.009	0.004	0.002
		rich	0.236	0.020	0.010	0.005	0.001
	rural	poor	0.381	0.050	0.021	0.010	0.005
		rich	0.189	0.022	0.010	0.002	0.001

We assume search costs are log-normally distributed over the population of consumers. The search cost model is estimated for beer, cola, coffee, and whisky. Different products within the category differ with respect to the store where they were bought but also with respect to other observable characteristics of the products. The observed characteristics of the products that enter the utility function are product brand, material of the container, flavor ("color" for beer, and light or normal for cola), and pack size (or bottle size, in the case of whisky). The characteristics of the store are the store chain identity, store surface, and region. The effect of store and product characteristics on utility are allowed to vary according to the type of the consumer. We also include time period fixed effects (year and month).

For our instrumental variables, we use the weekly average price paid by other household types living in the same region, and the total number of stores per region, week, and type. The price of beer is also instrumented by the average weekly price paid by other household types for the same color of beer ("blonde" or not) in the same region, and the weekly average price paid by the same household type for the same color of beer in other regions. Actually, these instruments are supposed to be uncorrelated with consumer specific taste shocks v_{ijt}, but are likely correlated with the store j price p_{jt} such that the rank condition is satisfied.

The probabilities of visiting a given store are recovered from the data of all purchases made by the consumer. We observe all households' store visits that generated a purchase. We then identify the probability a type i household visits a store chain in a week t as the ratio of the number of visits to stores of chain j during t by all households of type i over their total number of store visits during that week.

Table 7 displays the average (across regions and periods) of the cumulative distribution function values $G(c_{ij}^t)$ evaluated at the search cost cutoffs c_{ij}^t for each type of consumer and each of the four products studied using (8). We denote by G_j these averages for $j = 1, 2, 3, 4, 5$. In general, poor people (either urban or rural) search more on average. In the case of beer, coffee and cola, the proportion of consumers that do not search at all $(1 - G1)$ is around 40% or less among the poor, whereas about half of the richer consumers do not search. The proportion of people that do not search is highest for whisky: around 70% of the poorer, and 80% of the richer do not search for whisky. There is on average more people that do not search in rural areas than in urban areas but the main difference on searching behavior seems to be across income levels.

The proportion of people searching one, two, three, or four times before making a purchase decreases fast. Around 40% of consumers search once on average (this proportion corresponds to $G1 - G2$), except in the case of whisky for which only around $18\% - 38\%$ of consumers search once (depending on the type). The proportion of people searching twice $(G2 - G3)$ drops to $6\% - 7\%$ among the poor (for whisky, less than 3%), and $4\% - 5\%$ among the richer consumers (for whisky, less than 1%). Less than 5% of the poor and less than 3% of the rich search 3 times $(G3 - G4)$ on average before making a purchase. These numbers are again lower for whisky, for which less than 1% search 3 times, independently of the income level. The number of people searching more than 3 times is almost zero, irrespective of the product and the consumer type.

Table 8 shows the results of the estimation of utility parameters. To be completed.

Table 8: GMM estimation of utility parameters - Cola

	coefficient	S.E.
price X type 1	-0.2607*	0.0166
price X type 2	-0.1433*	0.0129
price X type 3	-0.2027*	0.0176
price X type 4	-0.1695	0.0153
rich	0.0083*	(0.0021)
rural	0.0628*	(0.0072)
other product charac	yes	
brand FE	yes	
store FE	yes	
time FE	yes	
region FE	yes	
N	20181	

*(p<0.01)

5.2 Price Elasticities

Once we have estimated the search cost distribution and the consumers preferences, we can estimate the own and cross price elasticities of demand in this consumer search model.

For any of the store $s \in \{1,..,N\}$ and any consumer type i, the store s demand $q_{i,s}^t$ by type i is

$$q_{i,s}^t(\mathbf{p_t}) = \sum\nolimits_{j=1}^{N} P(R(s)=j)\, q_{ij}^t(\mathbf{p_t}) \tag{11}$$

where $R(s)$ is the utility rank of store s among all stores available for consumer i and q_{ij}^t is the demand to store ranked j by consumer i at t that depends on the vectors of all prices $\mathbf{p_t}$.

Then, assuming that v_{ijt} is iid type I extreme value in (10), we have can write $P(R(s)=j)$ as a function of all characteristics and prices. Actually, denoting Θ_j^s the set of subsets of size $j-1$ of the set $\{1,...,N\}\setminus\{s\}$, we have

$$\begin{aligned} P(R(s)=1) &= P(u_{ist} \leq u_{irt}\ ; \forall\ r \in \{1,...,N\}) \\ &= \frac{1}{1+\sum_{r=1}^{N}\exp\left((x_{ist}-x_{irt})\beta - \alpha(p_{st}-p_{rt})\right)} \end{aligned} \tag{12}$$

and for $j = 2,..,N$

$$P(R(s)=j) = \sum_{\Omega\in\Theta_j^s} P\left(u_{irt} < u_{ist} \leq u_{ir't}\ ; \forall\ r \in \Omega, \forall r' \notin \Omega\right) \tag{13}$$

because the probability that store s is of rank j is equal to the probability that $j-1$ stores have lower utility and $N-j+1$ have higher utility, and this happen for all possible combinations in two groups of the other $N-1$ stores (the group of stores with lower utility being denoted Ω and the group of stores not belonging to Ω that will have higher utility than s).

For a given store s, the set Θ_j^s is thus

$$\Theta_j^s = \bigcup_{\Omega} \underbrace{\Omega \subset \{1, 2, .., N\}}_{s.t.\ card(\Omega)=j-1 \text{ and } s \notin \Omega}$$

Then, for any set $\Omega \in \Theta_j^s$, the probability that u_{ist} is the larger in $\Omega \cup \{s\}$ and the smaller in $\{1, ..., N\} \setminus \Omega$ is

$$\begin{aligned} & P\left(u_{irt} < u_{ist} \leq u_{ir't}; \forall r \in \Omega, r' \notin \Omega\right) \\ = \;& P\left(u_{irt} < u_{ist} \; ; \forall \; r \in \Omega\right) P\left(u_{ist} \leq u_{ir't} \; ; \forall \; r' \notin \Omega\right) \\ = \;& \pi_{s,\Omega}^{\max}\left(\mathbf{p}\right) \pi_{s,\Omega}^{\min}\left(\mathbf{p}\right) \end{aligned} \tag{14}$$

where

$$\pi_{s,\Omega}^{\max}\left(\mathbf{p}\right) \equiv P\left(u_{irt} < u_{ist} \; ; \forall \; r \in \Omega\right) = \frac{1}{1 + \sum_{r \in \Omega} \exp\left(\left(x_{irt} - x_{ist}\right)\beta - \alpha\left(p_{rt} - p_{st}\right)\right)} \tag{15}$$

and

$$\begin{aligned} \pi_{s,\Omega}^{\min}\left(\mathbf{p}\right) \;&\equiv\; P\left(u_{ist} \leq u_{ir't} \; ; \forall \; r' \notin \Omega\right) = \prod_{r' \notin \Omega} P\left(u_{ist} \leq u_{ir't}\right) \\ &=\; \prod_{r' \notin \Omega} \frac{1}{1 + \exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)} \end{aligned} \tag{16}$$

We can now write

$$\begin{aligned} \frac{\partial P\left(R(s) = j\right)}{\partial p_{lt}} \;&=\; \sum_{\Omega \in \Theta_j^s} \frac{\partial}{\partial p_{lt}} P\left(u_{irt} < u_{ist} < u_{ir't} \; ; \forall \; r \in \Omega_j^s, \forall r' \notin \Omega_j^s\right) \\ &=\; \sum_{\Omega \in \Theta_j^s} \pi_{s,\Omega}^{\max}\left(\mathbf{p}\right) \frac{\partial \pi_{s,\Omega}^{\min}\left(\mathbf{p}\right)}{\partial p_{lt}} + \pi_{s,\Omega}^{\min}\left(\mathbf{p}\right) \frac{\partial \pi_{s,\Omega}^{\max}\left(\mathbf{p}\right)}{\partial p_{lt}} \end{aligned} \tag{17}$$

where

$$\begin{aligned} \frac{\partial \pi_{s,\Omega}^{\min}\left(\mathbf{p}\right)}{\partial p_{lt}} \;&=\; \alpha \sum_{r' \notin \Omega} \Bigg[\frac{\exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)}{\left[1 + \exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)\right]^2} \prod_{r \notin \Omega \setminus \{r'\}} \frac{1}{1 + \exp((x_{irt} - x_{ist})\beta - \alpha(p_{rt} - p_{st}} \\ &=\; \alpha \sum_{r' \notin \Omega} \left[\frac{\exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)}{\left[1 + \exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)\right]} \prod_{r \notin \Omega} \frac{1}{1 + \exp\left(\left(x_{irt} - x_{ist}\right)\beta - \alpha\left(p_{rt} - p_{st}\right)\right)} \right] \text{ if} \\ &= \alpha \pi_{s,\Omega}^{\min}\left(\mathbf{p}\right) \sum_{r' \notin \Omega} \frac{\exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)}{\left[1 + \exp\left(\left(x_{ir't} - x_{ist}\right)\beta - \alpha\left(p_{r't} - p_{st}\right)\right)\right]} \text{ if } l = s \end{aligned} \tag{18}$$

and

$$\frac{\partial \pi^{\min}_{s,\Omega}(\mathbf{p})}{\partial p_{lt}} = \prod_{r\notin\Omega, r\neq l} \frac{1}{1+\exp\left(\left(x_{irt}-x_{ist}\right)\beta-\alpha\left(p_{rt}-p_{st}\right)\right)} \frac{\alpha\exp\left(\left(x_{ilt}-x_{ist}\right)\beta-\alpha\left(p_{lt}-p_{st}\right)\right)}{\left[1+\exp\left(\left(x_{ilt}-x_{ist}\right)\beta-\alpha\left(p_{lt}-p_{st}\right)\right)\right]^2} \text{ if } l\neq s \text{ and}$$

$$= \alpha\pi^{\min}_{s,\Omega}(\mathbf{p}) \frac{\exp\left(\left(x_{ilt}-x_{ist}\right)\beta-\alpha\left(p_{lt}-p_{st}\right)\right)}{\left[1+\exp\left(\left(x_{ilt}-x_{ist}\right)\beta-\alpha\left(p_{lt}-p_{st}\right)\right)\right]} \text{ if } l\neq s \text{ and } l\notin\Omega \tag{19}$$

and

$$\frac{\partial \pi^{\min}_{s,\Omega}(\mathbf{p})}{\partial p_{lt}} = 0 \text{ if } l\neq s \text{ and } l\in\Omega \tag{20}$$

and

$$\begin{aligned}
\frac{\partial \pi^{\max}_{s,\Omega}(\mathbf{p})}{\partial p_{lt}} &= \frac{\partial}{\partial p_{lt}}\left(\frac{1}{1+\sum_{r\in\Omega}\exp\left(\left(x_{irt}-x_{ist}\right)\beta-\alpha\left(p_{rt}-p_{st}\right)\right)}\right) \\
&= -1_{\{s=l,l\notin\Omega\}}\frac{\alpha\sum_{r\in\Omega}\exp\left(\left(x_{irt}-x_{ilt}\right)\beta-\alpha\left(p_{rt}-p_{lt}\right)\right)}{\left(1+\sum_{r\in\Omega}\exp\left(\left(x_{irt}-x_{ilt}\right)\beta-\alpha\left(p_{rt}-p_{lt}\right)\right)\right)^2} \\
&\quad +1_{\{l\in\Omega,l\neq s\}}\frac{\alpha\exp\left(\left(x_{ilt}-x_{ist}\right)\beta-\alpha\left(p_{lt}-p_{st}\right)\right)}{\left[1+\sum_{r\in\Omega}\exp\left(\left(x_{irt}-x_{ist}\right)\beta-\alpha\left(p_{rt}-p_{st}\right)\right)\right]^2} \\
&= -1_{\{s=l,l\notin\Omega\}}\alpha\pi^{\max}_{l,\Omega}(\mathbf{p})\left(1-\pi^{\max}_{l,\Omega}(\mathbf{p})\right) \\
&\quad +1_{\{l\in\Omega,l\neq s\}}\alpha\exp\left(\left(x_{ilt}-x_{ist}\right)\beta-\alpha\left(p_{lt}-p_{st}\right)\right)\pi^{\max}_{s,\Omega}(\mathbf{p})^2
\end{aligned} \tag{21}$$

which allows to obtain $\frac{\partial q^t_{i,s}(\mathbf{p_t})}{\partial p_{lt}}$ using

$$\frac{\partial q^t_{is}(\mathbf{p})}{\partial p_{lt}} = \sum_{j=1}^{J}\left[\frac{\partial P\left(R(s)=j\right)}{\partial p_{lt}}q^t_{ij}(\mathbf{p}) + P\left(R(s)=j\right)\frac{\partial q^t_{ij}(\mathbf{p})}{\partial p_{lt}}\right] \tag{22}$$

where

$$\frac{\partial q^t_{ij}(\mathbf{p})}{\partial p_{lt}} = \phi_{ij}\sum_{k=1}^{j-1}\frac{\phi_{ik}g\left(c^t_{ik}\right)\frac{\partial c^t_{ik}}{\partial p_{lt}}}{\left(1-\sum_{k'=0}^{k}\phi_{ik'}\right)\left(1-\sum_{k'=0}^{k-1}\phi_{ik'}\right)} - \phi_{ij}g\left(c^t_{ij}\right)\frac{\partial c^t_{ij}}{\partial p_{lt}}\sum_{k=1}^{j-1}\frac{1}{1-\sum_{k'=0}^{j-1}\phi_{ik'}} \tag{23}$$

that is

$$\begin{aligned}
\frac{\partial q^t_{ij}(\mathbf{p})}{\partial p_{lt}} &= -\alpha\phi_{ij}\sum_{k=1}^{j-1}\frac{\phi_{ik}g\left(c^t_{ik}\right)\sum_{k'=k+1}^{N_i}\phi_{ik'}\left(1_{\{R(l)=k'\}}-1_{\{R(l)=k\}}\right)}{\left(1-\sum_{k'=0}^{k}\phi_{ik'}\right)\left(1-\sum_{k'=0}^{k-1}\phi_{ik'}\right)} \\
&\quad +\alpha\phi_{ij}g\left(c^t_{ij}\right)\left(\sum_{k=1}^{j-1}\frac{1}{1-\sum_{k'=0}^{j-1}\phi_{ik'}}\right)\left(\sum_{k'=j+1}^{N_i}\phi_{ik'}\left(1_{\{R(l)=k'\}}-1_{\{R(l)=j\}}\right)\right)
\end{aligned} \tag{24}$$

because using (3), we have

$$\frac{\partial c^t_{ik}}{\partial p_{lt}} = \sum_{k'=k+1}^{N_i}\phi_{ik'}\frac{\partial}{\partial p_{lt}}\left(u_{ik't}-u_{ikt}\right) = -\alpha\sum_{k'=k+1}^{N_i}\phi_{ik'}\left(1_{\{R(l)=k'\}}-1_{\{R(l)=k\}}\right) \tag{25}$$

and

$$q^t_{ij} = \phi_{ij}\left[1+\sum_{k=1}^{j-1}\frac{\phi_{ik}G\left(c^t_{ik}\right)}{\left(1-\sum_{k'=0}^{k}\phi_{ik'}\right)\left(1-\sum_{k'=0}^{k-1}\phi_{ik'}\right)} - \frac{G\left(c^t_{ij}\right)}{1-\sum_{k'=0}^{j-1}\phi_{ik'}}\right] \tag{26}$$

where by convention $G\left(c_{i0}^t\right) = 1$ and $\phi_{i0} = 0$.

The demand elasticity will depend on (22) which shows that the demand of a product in a store will depend on the rank probability of that store-product in the preferences ranking of the consumer as well as the conditional demand elasticity once the consumer has ranked all stores mean utilities.

This shows that information frictions on the side of the consumer (the search cost distribution), as well as the drawing probabilities of each store available to the consumer play an important role in the own price elasticity of demand at each store in addition to the usual effect of consumers marginal utility of income α.

For the sake of comparison, one could also use the specification of (10) and estimate consumers preferences under the assumption that there are no informational frictions. In such case we simply need to estimate a logit model (Berry, 1994) using the same instrumental variables to obtain consumers preference estimates and thus own price elasticities.

In Table 9 we report the estimated own-price elasticities for cola under our search model and the own-price elasticities of a logit model without information frictions. To be completed.

6 Conclusion

Price dispersion is an important characteristic of the french food market. Our empirical results show that only a part of the observed price differentials can be explained by store heterogeneity. We find evidence that the differences that remain are likely due to imperfectly informed consumers who need to engage in costly search in order to find the best available deal. As a result, consumers with a high opportunity cost of time search less and pay higher prices on average.

We develop an empirical strategy to estimate the magnitude and distribution of sequential search costs. We allow for products to be both vertically and horizontally differentiated. We identify the search cost distribution without having to make any restriction on the drawing probabilities of stores. The drawing probabilities are recovered from the data and are heterogeneous across time, store chain, and household type.

Results of the structural estimation of model parameters show that search costs for the products considered (beer, coffee, cola, and whisky) are high and that consumers do not search much. There is also indication of differential behavior with respect to search depending on the consumers' income level and whether the household lives in a rural or urban area. Lower income consumers tend to search more than higher income consumers, and urban consumers tend to search more than rural consumers.

To be completed.

7 References

Aguiar, Mark and Erik Hurst (2007). "Lifecycle Prices and Production", *American Economic Review*, Vol. 97, No. 5, 1533-1559.

Baye, Michael R., John Morgan, and Patrick Scholten (2006) "Information, Search, and Price Dispersion", in: Terrence Hendershot (Editor), Handbook on Economics and Information Systems, Elsevier.

Burdett, Kenneth and Kenneth L. Judd (1983) "Equilibrium Price Dispersion", *Econometrica*, 51: 955-969.

Cachon, Gérard P., Christian Terwiesch, and Yi Xu (2008). "On the Effects of Consumer Search and Firm Entry in a Multiproduct Competitive Market", *Marketing Science*, Vol. 27, No. 3, 461-473.

Carlson, John A. and R. Preston McAfee (1983) "Discrete Equilibrium Price Dispersion", *Journal of Political Economy*, 91: 480-93.

Chandra, Ambarish and Mariano Tappata (2011). "Consumer Search and Dynamic Price Dispersion. An Application to Gasoline Markets", RAND Journal of Economics, Vol. 42, No. 4, 681-704.

De los Santos Babur, Ali Hortaçsu and Wildenbeest, Matthijs (2012) "Testing models of consumer search using data on web browsing and purchasing behavior" *American Economic Review* 102, 2955-80

Giulietti, Monica, Jesus Otero and Michael Waterson (2009) "Pricing Behavior under Competition in the UK Electricity Supply Industry", *Oxford Economic Papers.*

Head, Allen, Lucy Q. Liu, Guido Menzio, and Randall Wright, R. (2012). "Sticky Prices: A New Monetarist Approach". *Journal of the European Economic Association* (10), 939-973.

Hong, Han and Matthew Shum (2006) "Using Price Distributions to Estimate Search Costs", *Rand Journal of Economics*, 37 (2), 275-257.

Honka, Elizabeth (2014) "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry", *RAND Journal of Economics*, 45(4), 847-884

Hortaçsu, Ali and Chad Syverson (2004) "Product differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index funds", *Quarterly Journal of Economics*, 119: 403-456.

Jansen, Maarten C. W. and Alexei Parakhonyak (2014) "Consumer search markets with costly revisits", *Economic Theory*, 55: 481-514.

Koulayev, Sergei (2014) "Estimating demand in online search markets, with application to hotel bookings", *RAND Journal of Economics*, 45(3), 553-575

Lach, Saul (2002) “Existence and Persistence of Price Dispersion: An Empirical Analysis”, *Review of Economics and Statistics*, 84 (3): 433-444.

Lewis, Matthew (2008) “Price Dispersion and Competition with Differentiated Sellers”, *Journal of Industrial Economics*, 56 (3), 654-678.

Moraga-González José-Luis, Zsolt Sándor and Matthijs E. Wildenbeest (2011) "Consumer Search and Prices in the Automobile Market", mimeo.

Rhodes, Andrew (2014) "Multiproduct Retailing", *Review of Economic Studies*, forthcoming.

Sorensen, Alan T. “Equilibrium Price Dispersion in Retail Markets for Prescription Drugs”, *Journal of Political Economy*, 108 (4): 833-850.

Salop, Steven C. and Joseph E. Stiglitz (1982) “The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents”, *American Economic Review*, 72: 1121-1130.

Stahl, Dale O., II. (1989) “Oligopolistic Pricing with Sequential Consumer Search”, *American Economic Review*, 79: 700-712.

Stigler, George J. (1961) “The Economics of Information”, *Journal of Political Economy*, 69:213-225.

Wildenbeest, Matthijs (2011) “An Empirical Model of Search with Vertically Differentiated Products”, *RAND Journal of Economics*, 42, 729-757