Search intensity, search time and prices: evidence from retail diesel markets in France

Jeisson Cárdenas[1] · Jesús Otero[2] · Luis H. Gutiérrez[2]

Abstract

We study the effect of price variations over time and across space on search intensity and search time by consumers in retail markets for diesel in France. The main contribution of the paper is that existing work in industrial organisation in this area has already studied the effect of such variability on the first measure of search but not on the second one. Using novel data from daily consultation measures and price information from a government-run website, we find that price dispersion across space increases search activity and the amount of time allocated to search. Furthermore, while contemporaneous price changes do not appear to influence the number of visits and time per visit to the website, several coefficients on past price changes are positive and statistically different from zero. Our results thus suggest that price dispersion and price variability play a role in inducing search by current and potential customers.

1 Introduction

Ever since Stigler (1961), a growing number of papers have analysed empirically the association between consumer information and price dispersion. Studying gasoline demand, Marvel (1976) finds that both price dispersion at a point in time and price variability over time depend on a set of proxy variables that represent the benefits and costs to consumers of acquiring information. In a review of the literature, Baye et al. (2006) conclude that

✉ Jeisson Cárdenas
   jeisson.cardenas-rubio@warwick.ac.uk

   Jesús Otero
   jesus.otero@urosario.edu.co

   Luis H. Gutiérrez
   luis.gutierrez@urosario.edu.co

1   Institute for Employment Research, University of Warwick, Coventry, UK

2   Facultad de Economía, Universidad del Rosario, Bogotá, Colombia

🍁 Springer

much of the observed price dispersion in online and offline markets stems from consumers' costs of acquiring information about firms, and firms' costs of transmitting information to consumers. The question that arises is: How do consumers acquire information? The theoretical literature offers two models: clearinghouse and sequential search models. In the former, (some) consumers enter an information clearinghouse and become fully informed through, for instance, newspaper and all-or-nothing non-sequential searches. In the latter, consumers assemble information by sequentially visiting the stores in the market and stopping at some point without becoming perfectly informed. The main feature in all-or-nothing non-sequential search models is that they do not allow for searching more or less, but for a discrete decision only: consumers do not search and remain uninformed, or do search and become fully informed.

During the last decade or so, the introduction of transparency initiatives in the form of mandatory price posting and the advent of vendor websites have led to the study of the effect of information disclosure mechanisms on market outcomes such as price levels and price dispersion; see, inter alia, Hong (2014), Jang (2015), Rossi and Chintagunta (2016) and Luco (2019). At the same time, a number of authors have started to use the number of daily visits to these websites as a proxy for consumer search, with the purpose of studying the factors that guide consumers to search for better prices, reversing, to some extent, the price dispersion-consumer search link studied in (almost all) the previous research.

In this paper, we examine the effect of price changes and price dispersion on consumer search in retail diesel markets in France. We take advantage of a piece of French legislation requiring that all stations that sell more than 500 m^3 of fuel per year must report all fuel price changes to the Ministry of Economy, which are then made freely available to the public through a governmental website. The motivation behind the website is to provide motorists with accurate and up to date price information to help guide their purchase decisions. From theoretical models of price dispersion, the French government website serves as an information clearinghouse that fully informs consumers who visit the site. It means sequential search models are less likely to apply to this situation.[1] We use these data to obtain an accurate measure of market-level cross-sectional price dispersion and their changes from day to day. In addition, the main feature of this research is that for measuring consumer search we use novel data on daily consultation measures from this government-run site.

While Lewis and Marvel (2011), Byrne et al. (2015), Noel (2018) and Noel and Qiang (2019) have already used retail price data from websites, i.e. *GasBuddy.com* for the US and Canada, the prices in such websites are provided by voluntary observers, and so there is the possibility of sample selection bias; see e.g. Atkinson (2008). In contrast, the fact that French legislation mandates that stations must report their prices implies that geographical coverage is maximised, while sample selection biases, data inaccuracies and report delays are minimised. In this respect, we follow closely Byrne and de Roos (2017) inasmuch as we also have a measure of search intensity given by the number of unique daily visits to the website. Unlike Byrne and de Roos (2017), the distinctive aspect of our analysis is that we use for the first time a measure of time spent per search. As in all models of search (see e.g. Varian 1980, Stahl 1989, Yang and Ye 2008, Lewis and Marvel 2011), consumers make decisions on whether to search to become informed about prices or not to search and stay uninformed. In these models, informed consumers are those that observe the prices

---

[1] See Baye et al. (2006) for a critical survey on information clearinghouse models and Pennerstorfer et al. (2020), who use this model in their work.

charged by all firms, and then purchase from the firm that offers the lowest price, while uninformed consumers shop randomly and only observe the price at which they purchase. A consumer's decision to search depends on whether the expected benefit of searching outweighs individual search cost. Thus, the information disclosure policy enforced by the French government affects consumer search decisions by reducing the costs of obtaining accurate and reliable price information.

In this paper, we provide two consumer search intensity measures: the number of visitors to the respective website and the time spent there.[2] The availability of a free site where fuel prices are posted raises a number of questions of interest: How can the disclosure of daily prices affect consumer behaviour? What does the time spent on this government website tell us? Who is expected to spend more time searching the market? Answering these questions involves the consideration of first- and second-order effects. Indeed, one potential first-order effect is that the government website pushes some part of the mass of uninformed consumers to become informed, since it reduces search costs and implicitly increases search benefits. Then, assuming consumers behave rationally, as in Tappata (2009), one would expect that the number of daily visits to the website are positively associated with both price dispersion and price variability over time. However, increased search benefits occur in models that only deal with the consumer side. A firm would find it optimal to compete for the larger share of informed consumers with a low price. Hence, the mass in the price distribution would shift downwards, giving less dispersion. The benefits to searching would go down because a consumer is more likely to already observe a low-price firm at random; see e.g., Petrikaitė (2016) for a theoretical analysis of increased transparency and the pricing decision by firms.

The second-order effects comprise the easiness to search the web relative to the gains from searching.[3] Some people who are efficient at using computers and surfing the Internet can retrieve information rapidly from the website and therefore have low or zero search costs. These individuals will use the website even if the gains from search are small (for example, because price dispersion is small). Other consumers are less efficient and only search if they perceive significant benefits. The time spent on the website reflects the efficiency of the search process. Therefore, to the extent that less and less efficient consumers start searching the market, the gains from search increase. Two effects emerge: First, the number of visits increases and, second, the average time spent on the website increases. Thus, the second-order effect arises from the successive entrance of less proficient consumers in using computers and surfing the Internet. Conceptually and to some extent, one can think of the customers with a high ability to navigate the government website in the spirit of the heterogeneous consumers in Tappata's model. To say, the portion of consumers he calls shoppers with negative or low search costs.

Another salient aspect that distinguishes our work from that in Byrne and de Roos (2017) is that these authors use data from a single city where, for reasons not specified,

---

[2] It is also possible to think of an alternative measure of search intensity following Kaplan and Menzio (2015). In that paper, the idea is that families with fewer employed members are likely to pay lower prices because they have more time to visit a more significant number of stores. Although there is no doubt that it is an interesting insight, data availability prevents us from using this related measure of search in our econometric analysis.

[3] Early literature on the cost of search in gasoline markets suggests two components. The first one comprises the costs of automotive depreciation, consuming gasoline, and the like, while the second one involves the cost of time spent in searching (Marvel 1976). Mandatory gasoline price retail reporting initiatives in government websites considerably reduce the first type of search costs.

price changes always occur on Thursdays, allowing consumers to anticipate variations and increase their searches in the prior days. In contrast, in the French database at our disposal price changes do not exhibit any regular behaviour that can be anticipated.

The paper is organised as follows. Section 2 describes the data. Section 3 presents the results of the econometric analysis on the determinants of search activity, as measured by the number of website visits and time spent per visit. Section 4 concludes.

2 Data description

We use two sources of data. The first comes from the website https://www.prix-carburants.gouv.fr/, which contains information on price changes in all gas stations that sell more than 500 m^3 of fuel per year in France. Specifically, we downloaded daily prices from 1 January 2016 to 30 June 2018 ($T = 912$ time observations). The choice of the sample period is dictated by the need to match the one available on the measures of daily search activity. The data contain an identification number for each gas station, the price of a litre of fuel (which we express in cents of Euro), and the date of the price report. In addition, it is also possible to identify the type of retailer (supermarket, major oil company, or small independent retailer), and if the station is located on a motorway. We focus on stations located in continental France; stations in the overseas territories of Corse-Du-Sud, Guadeloupe, Martinique, French Guiana and Réunion are excluded because they are not geographically contiguous with any other station in continental France, and so some of the arbitrage mechanisms that explain consumer search across stations do not operate in these instances. Initial inspection of the data reveals that the number of gas stations varies between 9266 and 9499 over the study period. Bearing in mind that the number of stations in France was 11,200 at the end of 2016 (FuelsEurope 2017, Figure 51), the database used in the empirical analysis offers a coverage of approximately between 83 and 85%. Lastly, we focus on the price of diesel (*gazole*), denoted p_t, because of the relative importance of this type of fuel in France over the sample period; for instance, in 2016 the share of diesel demand amounted to 83% compared to only 17% for gasoline (FuelsEurope 2017, Figure 10).

The second source of data is the open platform of French public data available at http://www.data.gouv.fr/. This website, which contains certified traffic statistics of the main government websites in the country, provides daily information on the number of visits, visitors, page views, and time spent on each site. In particular, we retrieve information of two variables for the website https://www.prix-carburants.gouv.fr/, namely total number of visits and time spent per visit. Very recently, Byrne and de Roos (2017) analysed data on the first variable for Perth (Australia), but not for the second one. The search activity data that we use are available from 1 January 2016 to 30 June 2018, refer to diesel only, and include searches performed using both desktop and mobile devices. At this point, it should be noted that one of the difficulties that we face has to do with the mismatch between the price data and the search data. In other words, we have very detailed local price data at the station level, while the search data is only at the national level. This issue forces us, like Lewis and Marvel (2011), to examine aggregate search behaviour with price levels and price dispersion aggregated across the country. Ideally, we would have preferred search data at a more local market-level (like a city or town) as in Byrne and de Roos (2017), which would make the local market definition clearer. Unfortunately, we were not able to work with the French price platform directly to gain access to such search data.
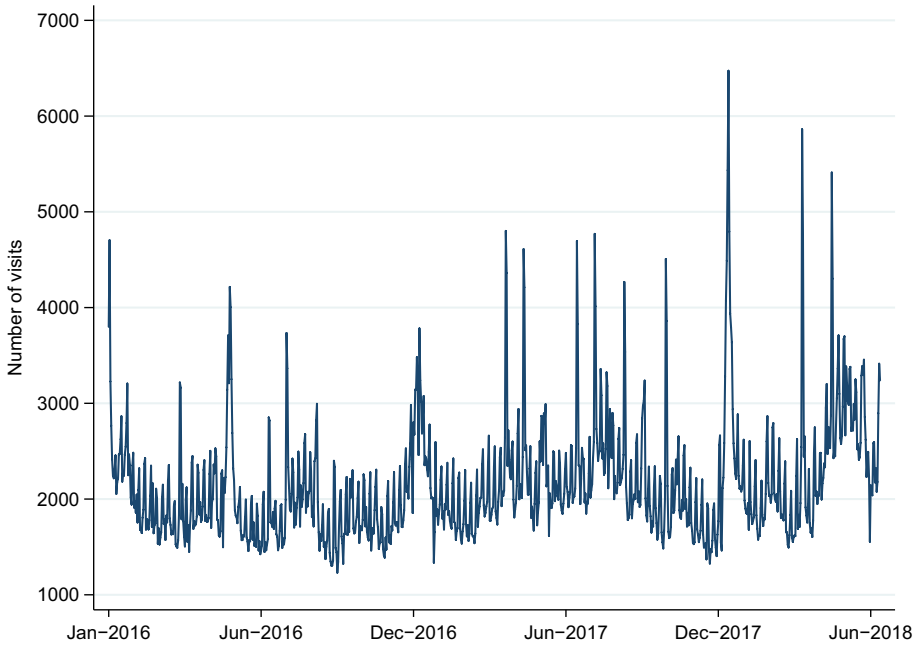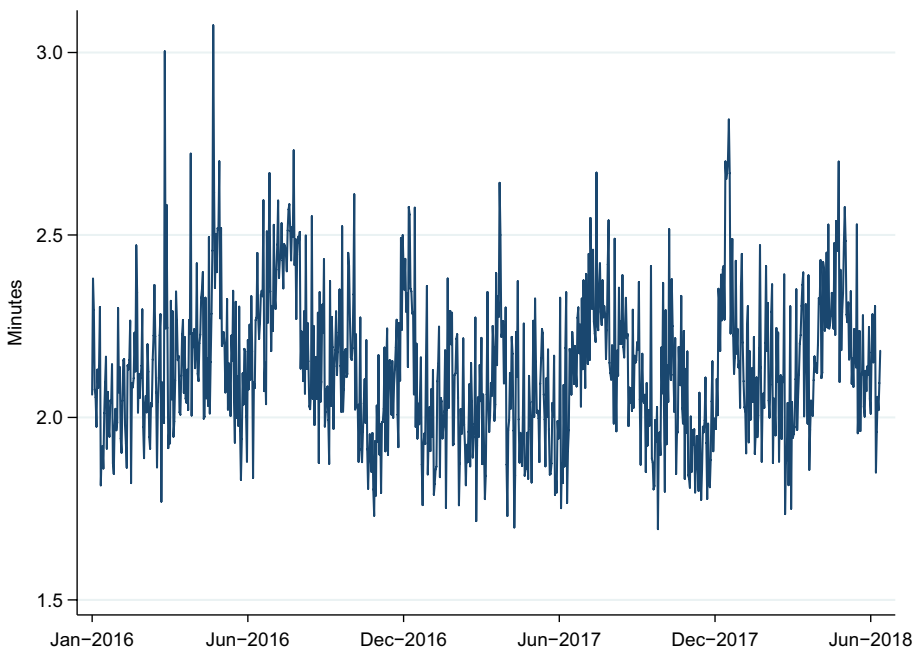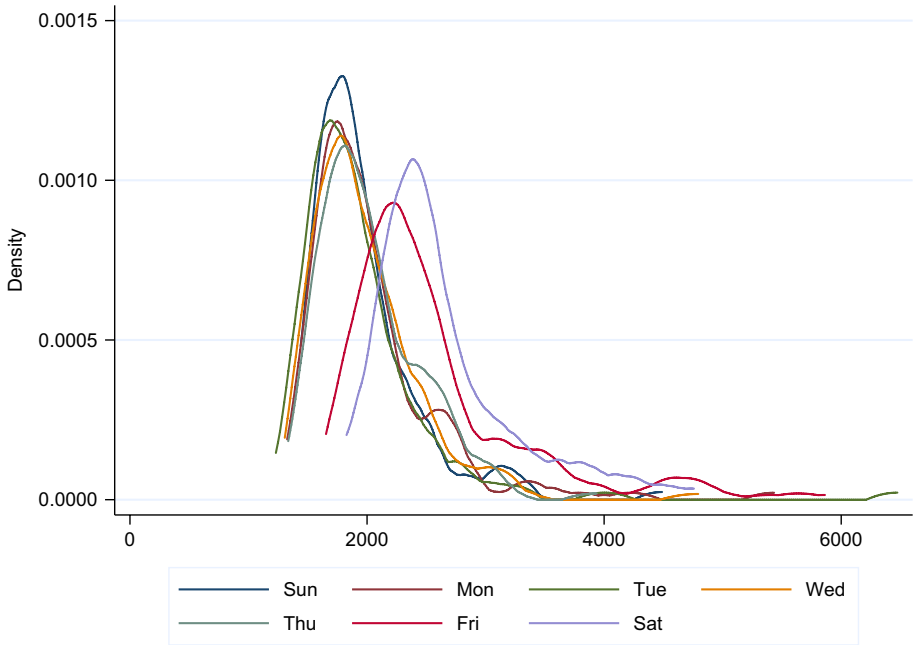
Fig. 1 Website visits
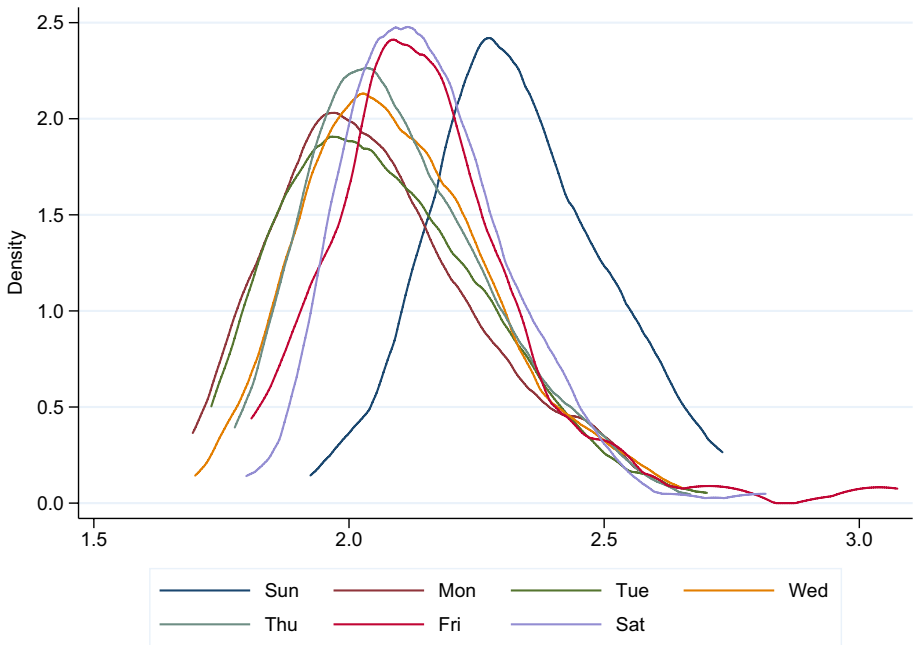

Fig. 2 Time spent per website visit

Fig. 3 Empirical density of website visits

Figures 1 and 2 present plots of the time path of the total number of visits and time spent per visit, respectively. The number of daily visits has fluctuated around 2000, although there are occasional days when it doubles and triples in size. As for the time spent per visit, the range of variation goes from 1.7 to 3.1 min, with an average of 2.1 min. The correlation coefficient between these two measures of search activity is approximately equal to 0.35, and is statistically different from zero. The fact that the correlation coefficient is moderate suggests that the two variables seem to be capturing different dimensions of the consumer search process (or, to put it another way, that one variable should not be used in place of the other one). Figures 3 and 4 offer kernel density plots of these measures of search activity by day of the week. As can be seen, the two plots are noticeable different. Indeed, while consumers tend to visit the site mostly on Saturdays and Fridays (Fig. 3), they tend to allocate more time to search for better deals on Sundays (Fig. 4). As for consumer search by season of the year, for both variables there appears to be more activity during the summer months of June, July and August, compared to the other months (these plots are not reported here to save space though).

3 Econometric analysis

Our empirical analysis of search activity examines the following variant of the model estimated by Byrne and de Roos (2017):

$$\ln y_t = \beta_1 + \sum_{i=0}^{k} \beta_{2,i} \Delta p_{t-i} + \beta_3 \sigma_{p_t} + \mathbf{x}_t^\mathsf{T} \boldsymbol{\beta}_4 + \varepsilon_t, \tag{1}$$

Fig. 4 Empirical density of time spent per website visit

where y_t is either the number of visits to the website, or the amount of time spent per visit. The set of regressors includes contemporaneous and past variations in diesel prices, $\Delta p_t = p_t - p_{t-1}$. It also includes the vector $\boldsymbol{x}_t^{\mathsf{T}}$ which consists of indicator variables for day-of-week, season and holiday effects, and the dispersion of diesel prices as measured by their cross-sectional standard deviation σ_{p_t}. Looking closely at the measure of price dispersion, it is worthy of notice that search data are observed on the national level, while local-ised search data would be better because competition in retail fuel is localised. Put another way, price dispersion is an indicator of the gains derived from search, and local (not national) price dispersion determines search benefits. If, for example, costs decline, sta-tions in some local markets might collude on the "old" price (as a focal point), while sta-tions in other local markets might not collude and lower their prices rapidly. In this case, we would have local markets with high prices and little price dispersion, and markets with low prices and little price dispersion. It means that the national measure on price disper-sion, σ_{p_t}, would deceptively suggest a high level of price dispersion in the market. Thus, as part of the robustness analysis, we further calculate measures of price dispersion at the local market (postcode) level and take the average of these localised dispersion measures rather than price dispersion calculated at the national level. The resulting measurement of price dispersion calculated in this alternative way is $\sigma_{p_t}^{\star}$. The model in Eq. (1) differs from that in Byrne and de Roos (2017) in as much as we do not include price-cycle related dummy variables, since price cycles do not appear to exist in France's fuel prices. The dis-turbance term is ε_t.

The presence of lags of Δp_t allows for the possibility that consumers respond to recent prices when deciding how much to search, e.g. if the press has been discussing rising oil/gas prices for several days. We estimate Eq. (1) for up to $k = 30$ and then select the optimal value of k using the Schwarz information criterion (SIC). Performing all estimations over

the same sample period, so that the resulting regressions are comparable for the different choices of k, and using σ_{p_t} we find an optimal value of $k = 7$ when the number of visits is the dependent variable, and $k = 8$ when the dependent variable is time spent per visit. Using $\sigma_{p_t}^{\star}$, we find optimal values of $k = 5$ and $k = 8$ when the dependent variable is the number of visits and the time spent per visit, respectively.[4]

An important estimation aspect that one must take into account is that consumers' search and firms' price-setting behaviour are determined simultaneously, as shown by Varian ([1980]) and, subsequently, Chandra and Tappata ([2011]). Thus, it is relevant to account for the possibility that price movements may cause search and vice versa.[5] To ensure that the estimates of the coefficients in Eq. ([1]) are not subject to simultaneity bias, we apply instrumental variable (IV) estimation. The idea behind this estimation method is to isolate and identify the effect of Δp_t on consumer search using a variable that does not appear in the equation, is correlated with Δp_t, and is not correlated with the error term ε_t. More specifically, we use Brent crude oil price changes (in Euro) as an instrument for Δp_t. The rationale is that the price of Brent crude oil is likely to influence retail fuel prices in France but is not influenced by consumer search activity in that country. In other words, the price of Brent crude oil is treated as an exogenous supply shifter. This choice of instrument is similar to that in Lewis and Marvel ([2011]), who use a measure of the wholesale price of gasoline.[6]

Table [1] presents IV estimation results of Eq. ([1]) for the number of website visits and time spent per visit. For both dependent variables, we estimate two models, one using the cross-sectional standard deviation σ_{p_t} as a measure of price dispersion, and the other one using the average of the localised (postcode) dispersion measures $\sigma_{p_t}^{\star}$. Focusing first on the number of visits, the positive sign on σ_{p_t} (and $\sigma_{p_t}^{\star}$) suggests that as price dispersion increases the number of visits to the website increases. Byrne and de Roos ([2017]) also find a qualitatively similar result, but only after allowing for the presence of price-cycle related dummy variables. Furthermore, while contemporaneous price changes do not appear to influence the number of visits to the website, there are several coefficients on past price changes which are positive and statistically different from zero. We obtain qualitatively similar results when using time per visit as the dependent variable. Lastly, further instrumenting σ_{p_t} and $\sigma_{p_t}^{\star}$ with their past values yield qualitatively similar estimates (these results are not reported here but are available upon request).

The previous findings are in line with Marvel ([1976]) observation that when firms decide to set a new price, they take into account the unanticipated responses of their current and potential clientele. Providers in a given area play a game of attractions and defections of customers. As indicated above, information costs of searching can be relatively low, i.e., the second-order effect, for current clients who are very skilful in

---

[4] We use SIC as it keeps the number of parameters to be estimated at a reasonable level. We also experimented with longer lags as chosen by the Akaike information criterion (AIC) and the results were largely the same, except that we observed (perhaps not surprisingly) larger standard errors due to multicollinearity among past prices.

[5] A more natural approach to finding causal effects would be to estimate a difference-in-differences model before and after introducing legislation mandating that all stations of minimum size must publicly report all price changes. However, data availability before the exogenous policy shock of transparency in the market prevents us from following this modelling approach.

[6] In contrast, empirical papers, inter alia Marvel ([1976]), Kaplan and Menzio ([2015]), investigating the influence from search to prices typically do not draw on the information of consumer search behaviour, but rather use variables influencing costs or benefits of search that are exogenous to firms' price decisions.

Table 1 IV estimates of the determinants of search activity

	Number of visits		Time per visit	
	Model 1	Model 2	Model 1	Model 2
Δp_t	− 0.230	− 0.149	0.035	0.043
	(0.124)	(0.156)	(0.053)	(0.052)
Δp_{t-1}	0.205	0.178	0.038	0.034
	(0.109)	(0.125)	(0.029)	(0.029)
Δp_{t-2}	0.110**	0.097**	0.032**	0.033**
	(0.028)	(0.024)	(0.010)	(0.010)
Δp_{t-3}	0.070**	0.057*	0.011	0.011
	(0.022)	(0.023)	(0.011)	(0.011)
Δp_{t-4}	0.045	0.045	0.041**	0.043**
	(0.024)	(0.025)	(0.011)	(0.011)
Δp_{t-5}	0.066**	0.129**	− 0.005	− 0.002
	(0.024)	(0.027)	(0.010)	(0.010)
Δp_{t-6}	0.076**		0.035**	0.036**
	(0.026)		(0.012)	(0.012)
Δp_{t-7}	0.081*		− 0.000	0.000
	(0.032)		(0.009)	(0.009)
Δp_{t-8}			0.037**	0.039**
			(0.010)	(0.010)
σ_{p_t}	0.372**		0.038*	
	(0.079)		(0.018)	
$\sigma_{p_t}^{\star}$		0.283*		0.067**
		(0.116)		(0.025)
N	904	906	903	903
R^2	0.153	0.090	0.255	0.268
F-test	9.100	8.280	21.580	19.890
p-Value	[0.000]	[0.000]	[0.000]	[0.000]

The coefficients on the intercept and indicator variables that account for day-of-week, season and holiday effects are not reported to save space. Newey and West (1987) HAC standard errors in parentheses. F-test tests the hypothesis that all estimated coefficients (except the intercept) are jointly equal to zero

* and ** denote statistical significance at the 5 and 1% significance levels, respectively

browsing the web. In these instances, a change in the price of their provider will almost instantaneously increase searching. The same may happen to other firms' customers who are also good at searching for prices on the government website. Both price dispersion and price variability play a role in inducing search by current and potential customers. Indeed, the effect on the search process will probably be more important when current clients have observed substantial price dispersion over a day in the past from their usual providers and vice versa. Our finding of positive and statistically significant coefficients on σ_{p_t} and $\sigma_{p_t}^{\star}$ provide evidence in favour of this view. In addition, searching by the current clientele will very likely increase when the provider has changed its prices over time most frequently, i.e., high price variability. In our results, we find support for

this behaviour through the positive and statistically significant effect of at least three of the coefficients associated with the Δp_{t-i}.[7]

4 Concluding remarks

This paper investigates the determinants of consumer search using two novel measures of search intensity, namely the number of visitors to the respective website and the time spent there. Other work in the area of consumer search has already studied the first measure but not the second one. We find that lagged price changes and price dispersion have a positive effect on both types of search. Improved transparency and reduced search costs have been the factors behind the French government mandating the posting of all station-level fuel price changes on a government website. This policy puts in place an important effect. Indeed, to the extent that market power is a feature of fuel retail markets, government efforts in increasing price transparency help reduce price coordination, so that fuel suppliers tend to keep their competitive pressure over rivals.

We hope our findings stimulate further work in different directions. For instance, Chandra and Tappata (2011) report that price dispersion is smaller among stations that are closer to each other because consumers face lower search costs when comparing prices among nearby stations. However, in our empirical analysis the measures of consumer search are nation-wide as they refer to the whole of France, and as such they do not permit us to identify features associated with specific local markets. Therefore, an interesting avenue of future research might involve the use of more refined measures of consumer search based on local-level data.

Data availability The data used in the paper will be made available for replication purposes.

Code availability Stata ado-files will be made available for replication purposes.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

---

[7] In Tappata's (2009) static model, a fraction of consumers search, observe all prices in the market, then become fully informed and buy from the cheapest producer. The remaining fraction of consumers keep uninformed and buy randomly from a given producer. This model assumes that the decision of consumers to search or not is affected by the expected price dispersion in the market and their search costs, i.e., consumers form rational expectations on producers' prices. Our interpretation assumes that consumers form their expectations rationally.

References

Atkinson, B.: On retail gasoline pricing websites: Potential sample selection biases and their implications for empirical research. Rev. Ind. Org. **33**(2), 161–175 (2008)

Baye, M.R., Morgan, J., Scholten, P.: Information, search, and price dispersion. In: Hendershott, T. (ed.) Economics and Information Systems (Handbooks in Information Systems, Volume 1). Bingley, Emerald (2006)

Byrne, D.P., de Roos, N.: Consumer search in retail gasoline markets. J. Ind. Econ. **65**(1), 183–193 (2017)

Byrne, D.P., Leslie, G.W., Ware, R.: How do consumers respond to gasoline price cycles? Energy J. **36**(1), 115–147 (2015)

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

FuelsEurope. Statistical Report. Belgium (2017)

Hong, W.-H.: Do smartphones spur competition? evidence from the Korean retail gasoline market. (2014) https://www.researchgate.net/publication/260676753

Jang, Y.: The effects of the internet and mobile search technologies on retail markets: Evidence from the Korean gasoline market. (2015) http://www.economics.mit.edu/files/10134

Kaplan, G., Menzio, G.: The morphology of price dispersion. Int. Econ. Rev. **56**(4), 1165–1205 (2015)

Lewis, M.S., Marvel, H.P.: When do consumers search? J. Ind. Econ. **59**(3), 457–483 (2011)

Luco, F.: Who benefits from information disclosure? the case of retail gasoline. Am. Econ. J. Microecon. **11**(2), 277–305 (2019)

Marvel, H.P.: The economics of information and retail gasoline price behavior: an empirical analysis. J. Polit. Econ. **84**(5), 1033–1060 (1976)

Newey, W.K., West, K.D.: A simple, positive semi-definite, heteroskedastic and autocorrelation consistent covariance matrix. Econometrica **55**(3), 703–708 (1987)

Noel, M.D.: Gasoline price dispersion and consumer search: evidence from a natural experiment. J. Ind. Econ. **66**(3), 701–738 (2018)

Noel, M.D., Qiang, H.: The role of information in retail gasoline price dispersion. Energy Economics **80**(C), 173–187 (2019)

Pennerstorfer, D., SchmidtâDengler, P., Schutz, N., Weiss, C., Yontcheva, B.: Information and price dispersion: theory and evidence. Int. Econ. Rev. **61**(2), 871–899 (2020)

Petrikaitė, V.: Collusion with costly consumer search. International Journal of Industrial Organization **44**(C), 1–10 (2016)

Rossi, F., Chintagunta, P.K.: Price transparency and retail prices: evidence from fuel price signs in the Italian highway system. J. Market. Res. **53**(3), 407–423 (2016)

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

Stigler, G.J.: The economics of information. J. Polit. Econ. **64**(3), 213–225 (1961)

Tappata, M.: Rockets and feathers: understanding asymmetric pricing. RAND J. Econ. **40**(4), 673–687 (2009)

Varian, H.R.: A model of sales. Am. Econ. Rev. **70**(4), 651–659 (1980)

Yang, H., Ye, L.: Search with learning: understanding asymmetric price adjustments. RAND J. Econ. **39**(2), 547–564 (2008)