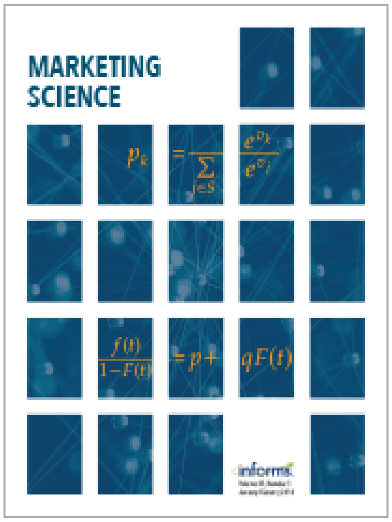
This article was downloaded by: [154.59.124.38] On: 30 May 2021, At: 08:59
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA



Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Price Uncertainty and Market Power in Retail Gasoline: The Case of an Italian Highway

Federico Rossi, Pradeep K. Chintagunta

To cite this article:
Federico Rossi, Pradeep K. Chintagunta (2018) Price Uncertainty and Market Power in Retail Gasoline: The Case of an Italian Highway. Marketing Science 37(5):753-770. https://doi.org/10.1287/mksc.2018.1105

Full terms and conditions of use: https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2018, INFORMS

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

informs
http://pubsonline.informs.org/journal/mksc/
MARKETING SCIENCE
Vol. 37, No. 5, September–October 2018, pp. 753–770
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Price Uncertainty and Market Power in Retail Gasoline: The Case of an Italian Highway

Federico Rossi,[a] **Pradeep K. Chintagunta**[b]

[a] Krannert School of Management, Purdue University, West Lafayette, Indiana 47907; [b] Booth School of Business, University of Chicago, Chicago, Illinois 60637

Contact: rossif@purdue.edu, http://orcid.org/0000-0003-4492-9375 (FR); pradeep.chintagunta@chicagobooth.edu, https://orcid.org/0000-0003-2854-5216 (PKC)

Received: July 24, 2016
Revised: June 13, 2017; February 21, 2018
Accepted: March 22, 2018
Published Online in Articles in Advance: October 23, 2018

https://doi.org/10.1287/mksc.2018.1105

Copyright: © 2018 INFORMS

Abstract. We quantify the effect of consumers' price uncertainty on gasoline prices and margins on an Italian highway. We observe the change in prices triggered by a longitudinal policy-based change in consumers' price information from one in which drivers on the highway have no information on the prices of stations they encounter to one that allows consumers to observe the prices of four upcoming stations on a single price sign by the side of the highway. Using these data, we estimate a model of consumer search and purchase behavior and a corresponding model of gas station pricing. We then measure the impact of varying degrees of price information on equilibrium prices, including (i) no price information, (ii) the current policy, and (iii) full price information. We also compare the current policy with an alternative policy in which stations' prices are advertised with individual price signs. We find that when consumers do not have price information, gas stations are able to charge 31% more, in terms of higher price-cost margins, than when prices are known. Our welfare analysis suggests that price information is worth €0.57 to consumers every time they take the highway. Relative to the current mandatory policy, advertising price on individual signs is worth €0.19 more to consumers.

History: Avi Goldfarb served as the senior editor and Tat Chan served as associate editor for this article.
Funding: The authors thank the Center for Research on Marketing and Services at Bocconi University and the Initiative on Global Markets and the Kilts Center for Marketing at the University of Chicago Booth School of Business for financial support.
Supplemental Material: The e-companion is available at https://doi.org/10.1287/mksc.2018.1105.

Keywords: consumer search • price uncertainty • retail competition • retail gasoline

1. Introduction

A driver interested in purchasing gasoline is usually not a price-informed consumer, for at least two reasons. First, the price of gasoline changes frequently, sometimes even on a daily basis. Such a variation is due, in large measure, to the high volatility of crude oil prices, which represents the largest cost driver in this market.[1] For example, in the second half of 2014, the national average price of regular gasoline in the United States fell by almost 40%.[2] Because most of this variation is hard to predict, consumers cannot rely on prior price knowledge to evaluate current prices. Second, each oil company distributes most of its gasoline exclusively through a network of affiliated stations. This implies that, in contrast to markets such as consumer packaged goods, consumers cannot compare the prices of competing brands at the point of sale. Because gasoline is purchased periodically and can be stockpiled only in limited amounts, it is inconvenient for consumers to visit nearby stations and search for price information at each purchase.

Price uncertainty is not unique to retail gasoline markets. The marketing literature has documented evidence that consumers in general have scarce information regarding prices (Dickson and Sawyer 1990) and search to acquire more information (Mehta et al. 2003, Gauri et al. 2008, Seiler 2013, Honka 2014). Retail gasoline, however, has attracted attention because of the relative lack of differentiation and homogeneity in marginal costs. These characteristics would suggest the existence of the law of one price, which is not confirmed empirically. Instead, researchers have shown the role of price information on price dispersion. For example, Chandra and Tappata (2011) have found that consumer search explains a significant part of the variation in the price difference between gas stations. Pennerstorfer et al. (2016) have shown that the relationship between the level of consumers' price information and its dispersion exhibits an inverted-U shape, as predicted by theory (Varian 1980, Stahl 1989).

To reduce the level of price uncertainty, lawmakers regulate gas stations' price advertising with the mandatory use of price signs next to each station. A number of U.S. states as well as several other countries, have similar mandatory price disclosure policies. This practice can help provide consumers with price information before they make a purchase. However, there are instances where price advertising is not mandatory,[3] or simply ineffective. For example, consumers driving on a highway may be unable to observe the prices advertised by the gas stations located along the road, as that requires exiting the roadway to observe the prices. Even in situations where they can see the prices from the road, it could be difficult to turn back to avail oneself of a lower price at a previous station.

In this paper, we investigate the effect of consumers' price uncertainty on gas stations' prices and margins. While a station's ability to increase its prices relative to its competitors might be due to its own characteristics, such as its location or the presence of a nearby convenience store, it is also in part due to consumers' cost of acquiring price information. Using a longitudinal change in consumers' price information, triggered by a mandatory policy that allows consumers to observe on the same sign the prices of four nearby stations, we estimate a model of consumers' search and purchase behavior and a corresponding model of gas station pricing. The model estimates help us quantify the impact of consumers' price knowledge on the market price of gasoline. In particular, we measure the impact of varying degrees of price information on equilibrium prices, including a hypothetical scenario in which consumers have full price knowledge. In addition, we evaluate the effect on prices and welfare of the current mandatory price policy, which allows consumers to observe the prices of four nearby stations on the same sign, and compare it with an alternative policy where each station's price is advertised on a separate sign.

A key empirical challenge we face when estimating our parameters is that we have access to price data but not demand data. This challenge is common for policy makers, who typically do not have access to demand data of competing firms. A number of papers have shown how price data alone can help identify key demand primitives (Thomadsen 2005, Hong and Shum 2006). In particular, identification requires us to make assumptions on (i) the parametric structure of demand and costs and (ii) the equilibrium conduct of stations in setting their prices. Different from previous work, however, there are several aspects of our empirical application that aid in identification of search costs and other parameters. First, we observe a change in the informational environment, which shifts equilibrium prices; this shock allows us to identify additional factors such as search costs or preexisting sources of market power. Second, we have access to data on traffic patterns that reveal the choice set of consumers and whether they have access to the price information disclosed. Third, we observe the purchases of a panel of consumers with loyalty cards for one of the oil companies competing on the highway. These data give us the empirical distribution of how far consumers travel on the highway before they purchase gas. We use these data to construct an empirical measure of consumers' *ideal refill points*—that is, their likelihoods of purchasing gas—as a function of distance driven on the highway. The differences in ideal refill points across consumers allow us to capture heterogeneous preferences of consumers for stations' locations.

Our analysis reveals that in the absence of mandatory price disclosure, gas stations are able to charge consumers about 31% higher margins than in the case of perfect price information. With the current price disclosure policy, this informational rent is reduced to 9%. Therefore, the current disclosure reduces about two-thirds of the stations' market power due to information.

Looking at the consumers' perspective, we find that price information is worth €0.58 to a consumer on each driving occasion on the highway. Next, we compare the welfare effect of the current (four prices per sign) policy with that of a price disclosure policy where each individual station's price is advertised by a sign installed before the station. In this case, the market reaches an outcome similar to the full-information one, and consumers are able to obtain most of the €0.57 value. In contrast, with the current mandatory policy, they lose part of this value (€0.19), as in this case the information does not reach many of the consumers traveling on the highway.

We believe our results contribute to the literature in two important ways. First, although the relationship between price information and market prices has been documented, we are not aware of any study measuring the value of information in terms of retail market power. Here, we quantify the consumers' cost of visiting a gas station and its implication for the price of gasoline: price uncertainty is responsible for a large part of gas stations' market power. Second, we show that the effect of price disclosure could be significantly improved by adopting a more effective policy. These results provide insights to regulators who are considering the benefits of changing the current disclosure policy in the retail gasoline market or even expanding it to other markets (e.g., healthcare).

1.1. Related Research

Our work relates to a growing empirical literature on imperfect information, which quantifies search frictions and its implications for the market. Many of these studies focus on demand-side implications. For example, using scanner data set for liquid detergents, Mehta et al. (2003) propose a model where consumers form a consideration set before making their final purchase. Seiler (2013) estimates a search model using laundry detergent data and shows the relevance of consumer search costs for the effectiveness marketing promotions. Honka (2014) quantifies the search and switching costs for shoppers of auto insurance and shows that such costs rationalize the high level of retention in this industry. Differently from these papers, our work studies the effect of price uncertainty on the supply side. Other recent research has also adopted this perspective. For example, Moraga-González et al. (2015) use automobile data to show the effect of a change in consumer search costs on prices. Our research differs from this work mainly in the identification strategy adopted. Instead of relying on variation of observed markets' demographics, such as distance from dealership or income, which may affect the cost of acquiring price information, in our research, we identify search costs using a shift in the price information available to consumers. The observed longitudinal variation in market prices is more likely driven by the causal effect of price information on firms' pricing decisions.

Our research is also related to other work in retail gasoline. For example, Chan et al. (2007) use consumers' demographics and pricing data to estimate a model of location and pricing decisions in the retail gasoline market. They find that consumers are willing to travel up to one mile to save $0.03 per liter. Iyer and Seetharaman (2008) show how consumer heterogeneity and competitive pressure can help explain both product and pricing decisions made by the gas stations. Using variation from a pseudonatural experiment, Sen et al. (2011) show that the introduction of a gas station in the parking lot of a supermarket generates positive externalities for the grocery business, increasing its sales by almost 10%.

The longitudinal informational change we observe in our data relates our work to recent research on information disclosure. For example, Gaynor et al. (2016) exploit a mandatory policy that increases the consideration set of patients in the English national health service. In a recent paper, Rossi and Chintagunta (2016) explored the effect of price transparency on gasoline retail prices. The goal of their paper was to describe the effect of price transparency, via the installation of price signs, on retail prices. Our paper uses the same price data. At the same time, there are notable differences between the two studies. First, we augment the data in that study with data on consumers' traffic patterns. More importantly, in this paper, we estimate a behavioral model for consumers and firms. This allows us to measure, via counterfactuals, the impact of varying degrees of price information on market prices. Our analysis also allows us to compare the effect of alternative disclosure policies on fuel prices. In other words, we extend the descriptive nature of the study by Rossi and Chintagunta (2016) to the more prescriptive domain by providing insights for policy makers considering measures to enhance the price transparency in the market.

2. Background, Data, and Descriptive Analysis

2.1. Background: Highway System and Price Signs

The Italian highway system is a network of tollways. Inside the highway, approximately every 27 km (about 17 miles), there are service areas specific to each direction of travel. In each service area, there is a gas station and a convenience store/restaurant.

Each gas station is affiliated with an oil company; therefore, it carries its logo and purchases fuel exclusively from that company, but has independence in pricing decisions. The oil company provides marketing support such as national advertising and administering a reward program, but does not enforce a retail price.[4]

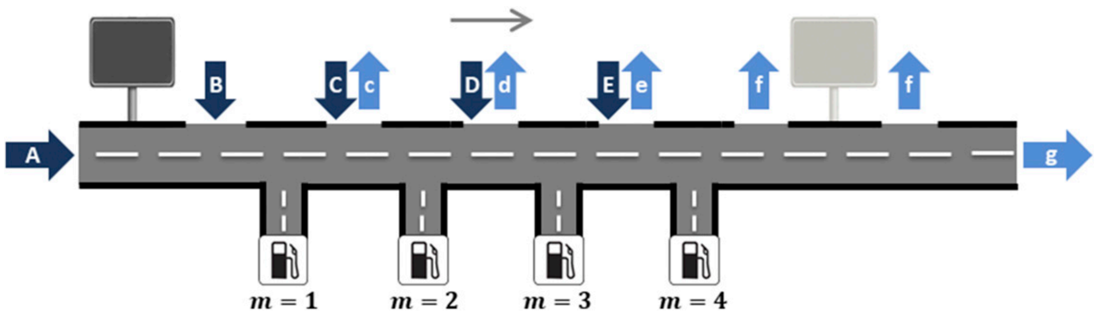
Because of the similar distances between two consecutive stations, consumers driving on the highway have a good idea of their distance from the stations. However, before the year 2007, they had no information regarding fuel prices charged by the stations. A driver would have to visit a station to observe its price. Between 2007 and 2009, a number of electronic price signs were installed across the network. In most cases, each sign posts the fuel prices (both gasoline and diesel) of four gas stations, starting from the first station in each route, suggesting that the choice of the location can be considered orthogonal to the characteristics of the stations. For example, on the east side of the highway route A14, which is one of the longest routes and stretches for about 750 km north to south, a total of seven signs were installed, one every four stations. Each sign posts the price of the next four stations, covering all the 28 gas stations located on this route. Each sign also provides information about the brand affiliation of each station and its distance from the sign, although information on brand affiliation was also available through regular signage prior to the introduction of the signs.

In Figure 1 we represent a typical segment of the highway that includes a sign with the four consecutive stations whose prices are posted on the sign. Between any two stations, there are toll gates that allow vehicles to enter or exit the highway. Because of their particular positions, after price signs are introduced, the price knowledge of consumers depends on the driving path followed. For example, a consumer driving along the path A–d will transit in front of the sign, so the consumer will know the price of station 1 and station 2 even before visiting either of these stations. In contrast, a consumer driving along the path B–d will not have such information. The level of price uncertainty ultimately depends on the number of consumers taking each of the available paths.

2.2. Data

We use several data sets in our analysis: the two primary ones are gasoline prices and traffic data. The first

Figure 1. (Color online) Highway Segment Including a Sign and Four Consecutive Stations Whose Prices Are Posted on the Sign



Note. Arrows with uppercase (lowercase) letters denote entry (exit) points.

data set, which is also used in Rossi and Chintagunta (2016), consists of daily gasoline prices charged by gas stations from January 2008 to September 2010. Table 1 provides summary statistics of these data.

The second data set, traffic data, is unique to this study and was purchased from Autostrade. It provides the gates of origin and destination of each vehicle that entered the highway in the year 2010. Using these data, we identify the driving path of each vehicle. This allows us to determine which consumers transit in front of signs and observe prices, and which do not. Also, it allows us to establish which stations are in competition for the same consumer.

For any given highway segment shown in Figure 1, we identify 19 unique driving paths: A–c, A–d, A–e, A–f, A–g, B–c, B–d, B–e, B–f, B–g, C–d, C–e, C–f, C–g, D–e, D–f, D–g, E–f, and E–g. Each path goes past a unique subset of signs and stations. We transform the number of transiting vehicles found for each driving path into a percentage; that is, we divide the number of consumers in each driving path by the total number of consumers on the 19 paths. In Table 2 we report for each driving path the average size (in percentage) across different highways segments. In parentheses we report the standard deviation across segments. Entry and exit gates are denoted by letters, as in Figure 1. The last row of the table shows that about half of the vehicles transiting in the segment exit from the highway (exit c to f), whereas the other half continue driving on the highway (exit g); the largest group of consumers is the one in which drivers traverse the segment and continue on the highway (A–g). One in five consumers follows this path. We see from the table that the standard deviations are relatively high, which means that the sizes of the paths change considerably from segment to segment. This variation is helpful in identifying the effect of the percentage of transiting drivers on stations' prices.

Although the traffic data are from 2010, we do not expect significant differences in driving paths from one year to another. It is possible that consumers might temporarily change the length of their trips because of fluctuations in fuel prices; however, these variations should average out over the course of a year. Driving

Table 1. Descriptive Statistics of Price Data

	Mean	Standard deviation	Minimum	1Q	Median	3Q	Maximum
Price levels							
Stations in position 1	1.390	0.140	1.040	1.296	1.381	1.499	1.743
Stations in position 2	1.393	0.139	1.044	1.298	1.385	1.501	1.748
Stations in position 3	1.391	0.138	1.040	1.299	1.383	1.501	1.748
Stations in position 4	1.392	0.140	1.035	1.296	1.383	1.502	1.809
Stations outside highway	1.405	0.134	1.105	1.319	1.383	1.500	1.726
Price dispersion on sign							
Range	0.059	0.052	0.000	0.022	0.046	0.087	0.451
Standard deviation	0.021	0.012	0.000	0.012	0.020	0.029	0.123
Price change frequency							
No. of consecutive days no change	5.617	1.737	1.556	4.422	5.387	6.485	15.088

Notes. We report statistics on the price levels, price dispersion, and frequency of price changes. For price levels, we consider the position of the station with respect to the price sign. For price dispersion, we report both price variance and difference between highest and lowest prices of stations that appear or will appear on the same price sign. 1Q, first quartile; 3Q, third quartile.

Table 2. Driving Paths as Percentages of Total Number of Drivers

	Exit					
Entry	c	d	e	f	g	Total
A	10.7 (5.0)	3.9 (3.6)	5.0 (3.8)	1.5 (1.2)	21.5 (10.6)	42.6 (8.2)
B	1.1 (1.6)	0.2 (0.4)	0.5 (1.0)	0.0 (0.1)	0.4 (0.8)	2.2 (3.3)
C		4.9 (2.6)	4.2 (1.9)	4.4 (6.8)	7.5 (6.4)	21.0 (14.3)
D			7.7 (6.1)	1.4 (1.3)	4.8 (3.4)	13.9 (8.9)
E				5.3 (3.2)	15.0 (6.7)	20.3 (5.5)
Total	11.8 (6.1)	9.0 (5.2)	17.4 (9.6)	12.6 (11.1)	49.2 (13.3)	100.0

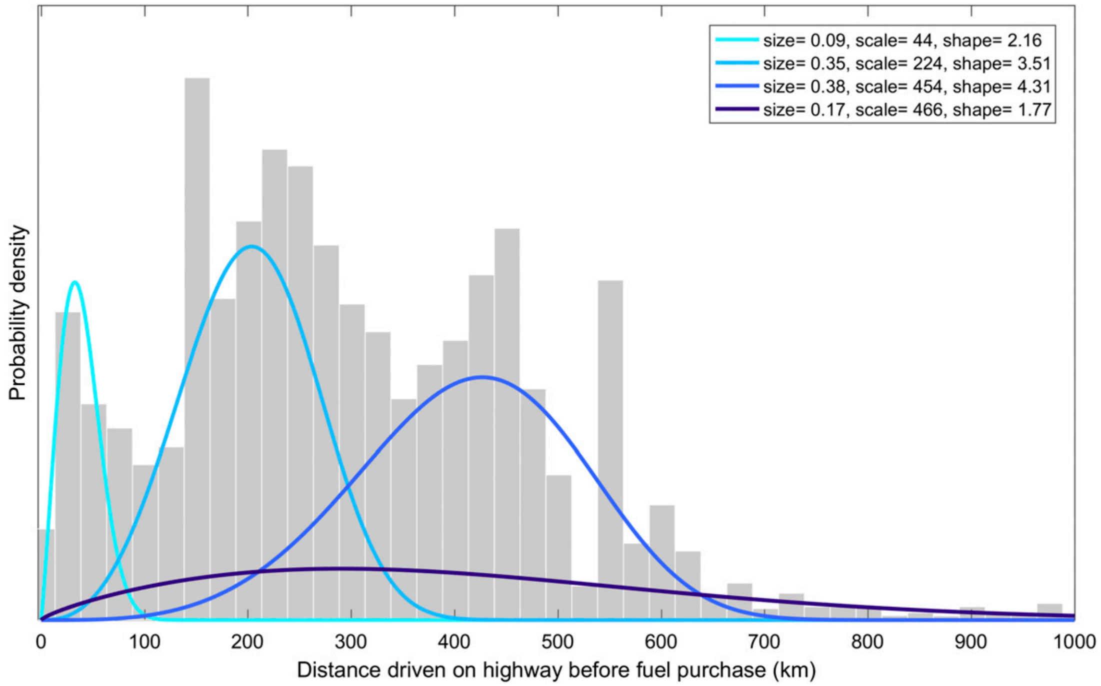
Notes. We report means across markets. Standard deviations are reported in parentheses.

paths might also change because of seasonality. Our price data reveal that stations change prices on average every six days, which suggests that stations' pricing is not affected by seasonality between days of the week. We will not be able to capture other types of seasonality, such as monthly variations, which might affect the length of traffic paths. A reduction in the total number of vehicles, instead, will not affect our analysis, because our interest is in the relative sizes of driving paths.

We also use two additional pieces of information in our analysis. The first comes from a sample of about 35,000 cardholders enrolled in the reward program of a major oil company, who made at least one purchase at any affiliated station located on the highways in the period 2008 to 2009. For each observed purchase, we consider the last purchase made by the cardholder outside the highway, find the shortest path between the two stations, and determine the number of kilometers along the path that are driven on the highway.[5] This is the number of kilometers the consumer is driving on the highway before purchasing fuel. The empirical frequency distribution of this variable is plotted in Figure 2. We use a mixture of four Weibull distributions to capture the heterogeneity within and across consumers and smooth the empirical frequencies.[6] This purchase probability is conditional on the number of kilometers driven; with the use of a large sample of consumers and stations, this measure is presumed independent of stations' characteristics, such as price or location. It suggests that 1 in 10 consumers decide to refill soon after they access the highway, whereas most others postpone their purchases till after they have driven for some distance. One-third of consumers drive for about 200 km before they stop to refill their tank, whereas another third drive for about 450 km without a fuel stop. The variable might include some measurement error due to the specific locations of the stations,[7] and because purchases at competing stations are not observed. Despite this potential noise, each distribution can be considered as a measure for consumers' ideal refill points because it informs us about the likelihood of making a purchase depending on how far the consumer has traveled on the highway. In the estimation, we use this information to characterize the heterogeneity across consumers in their likelihoods of stopping at various points along the highway to refill their tanks with gasoline. For example, consider a consumer following the path C–e in Figure 1. This consumer makes a purchase at either one of the two highway stations encountered along her driving path or at a station outside the highway. The consumer has to drive, say, 10, 35, and 60 km to reach the stations, respectively. We first find the segment covered by each station, which ranges from the midpoint between the location of the station and that of the previous station to the midpoint between the location of the station and that of the subsequent station. In our example, the segments are $(0, 22.5)$ km, $[22.5, 47.5)$ km, and $[47.5, +\infty)$ km, respectively. Then, we calculate the probability associated with each segment using the distributions in Figure 2. If we consider the first distribution, the likelihoods of refilling at the different stations become 20%, 50%, and 30%, respectively.

The second additional piece of information we use is pricing data from about 3,000 off-highway gas stations. These stations are affiliated with one leading brand and refer to a period prior to our analysis (year 2009). Using these prices, we compute (i) aggregated national prices and (ii) aggregated prices by province. We then compute the deviation of each province's average price from the national average. Finally, we apply these differences to the average daily national prices of gasoline provided by *Quotidiano Energia*. In this way, we obtain local prices of off-highway stations. As validation for these prices, we note that they are in fact closer to the price levels of the nearby highway stations than the national average. One limitation of these data is that because they reflect prices of one brand of gasoline, the mean price is representative of the stations outside the highway but the variance is not; that is, it is smaller than the spread across brands. To address this issue, we obtain a data set from the Italian Ministry of Economic Development that tracks the daily prices of most gas stations in Italy but starting at a later date than our data, in 2015. We use these data to obtain the standard deviation of prices for off-highway stations. We compute this at the level of each province. Although the data are from a different year, we expect similar price variance across local stations. Additionally, we assess the sensitivity of our results to this assumption.[8]

Figure 2. (Color online) Distribution of Distance Driven Before Purchase Occurrence



Note. The empirical frequency distribution is fitted via maximum likelihood with a mixture of four Weibull distributions.

2.3. Descriptive Analysis: Price Uncertainty After Policy Change

Rossi and Chintagunta (2016) show how the introduction of price signs reduces the price of fuel. Their results confirm that before the introduction of signs, consumers are uncertain about fuel prices, and that the price disclosure increases competitive pressure between gas stations. But are consumers informed *after* the disclosure? In this section, we provide evidence that even after price signs are installed, some consumers are still uncertain about fuel prices, and stations charge higher prices as a result.

The driving path data allow us to quantify the number of consumers left out by the policy. In Table 2, the last column shows the size of driving paths conditional on the entry location. It indicates that, on average, only 43% of consumers enter from A and transit in front of a sign (see Figure 1); most consumers enter the highway at a later point, so they cannot access the information disclosed.

Using the driving path data, we can also compute the percentage of price-informed consumers transiting in front of each station. According to Table 2, this number decreases significantly as we move from the first station (96%) to the second (59%), third (46%), and last stations (37%). This variation is helpful for testing whether stations charge higher prices because of the policy's lack of reach. We run such an analysis using a difference-in-differences approach and compare the pricing behavior of stations whose prices have been posted (treatment group) with that of stations whose prices have not been posted during the same period (control group). The analysis is similar to that by Rossi and Chintagunta (2016), but the key difference here is that we use driving path data, which allows us to test whether the percentage of price-uninformed consumers affects prices.

For robustness, we use two different types of control groups: other highway stations previously treated or yet to be treated, and a large national sample of gas stations located outside the highway (which have not been treated). In the first case, the dependent variable is stations' daily prices. To test for the effect of consumers' knowledge, we interact the variable indicating the presence of a sign with the percentage of consumers who passed in front of the sign. We control for the heterogeneity of responses across stations by including a sign-specific response effect. We also control for the station's position, as we want to distinguish the effect of percentage of price-informed consumers (which decreases for stations positioned far from the sign) from the effect of a station's position. The latter effect might capture additional factors such as consumers forgetting price information acquired previously, etc.

The regression model is specified as follows:

$$p_{it} = \delta_0 + \psi_i + \tau_t + PostOwn_{it}(\delta_1 + \tilde{\delta}_k + \delta_2 PrcSign_i) + \delta_3 PostPrec_{it} + \delta_4 PostFoll_{it} + \nu_{it}, \quad (1)$$

Table 3. Estimates of Price Responses to Price Posting

	Highway stations treated in other period		Stations outside the highway	
Control group	(i)	(ii)	(iii)	(iv)
PostOwn * *PrcSign* (δ_2)	−1.532* (0.676)	−2.847* (1.413)	−1.850* (0.708)	−3.228* (1.429)
PostOwn * *Position*2		−0.785 (0.462)		−0.743 (0.500)
PostOwn * *Position*3		−0.529 (0.570)		−0.455 (0.600)
PostOwn * *Position*4		−0.698 (0.658)		−0.774 (0.709)
PostPrec (δ_3)	−0.186 (0.400)	−0.161 (0.400)	−0.452 (0.414)	−0.431 (0.414)
PostFoll (δ_4)	−0.602* (0.296)	−0.621* (0.298)	−1.208* (0.318)	−1.224** (0.319)
Station fixed effects (δ_0, ψ_i)	✓	✓	✓	✓
Time fixed effects (τ_t)	✓	✓		
Sign fixed effects for *PostOwn* ($\delta_1, \tilde{\delta}_k$)	✓	✓	✓	✓
Adjusted R^2	0.980	0.980	0.592	0.593
Observations	183,058	183,058	132,753	132,753

Notes. The table shows OLS estimates of the effect of price signs on gasoline price levels. Specifications (ii) and (iv) control for stations' positions. Clustered standard errors are in parenthesis. Note that *PrcSign* varies only across stations, so its main effect is included through the stations' fixed effects. The number of observations when the control group is stations outside the highway is lower, because the data for the control group have fewer observations.

where p_{it} is the retail price[9] (expressed in euro cents) for a liter of regular gasoline at station i on date t, δ_0 is a constant, ψ_i is the fixed effect specific to station i, τ_t is the fixed effect specific to time t, $PostOwn_{it}$ is an indicator for if a price sign is activated and station i's price is posted on it, δ_1 is a constant of the effect of $PostOwn_{it}$ across stations, $\tilde{\delta}_k$ is its effect specific to stations of sign k, $PrcSign_i$ is the percentage of consumers passing in front of station i who also passed in front of the point where the sign is installed,[10] and ν_{it} is an independent and identically distributed (i.i.d.) error term. Finally, we include the indicator variables associated with signs for stations not on the focal sign to control for any "cross-sign" effects on prices, as in Rossi and Chintagunta (2016): $PostPrec_{it}$ ($PostFoll_{it}$) indicates whether another station preceding (following) the focal station i within 100 km on the same highway route and direction of travel has its prices posted on a different sign. The key parameter of Equation (1) is δ_2, because it captures the correlation between the percentage of consumers who passed in front of the sign and the effect of the sign on the stations whose prices are posted.

Ordinary least squares (OLS) estimates of the model above are reported in Table 3. The estimates reported in column (i), where the control group consists of the stations treated in other periods, show that the interaction between *PostOwn* and *PrcSign* is negative and significant: prices are reduced more when the percentage of consumers passing in front of the sign is higher. In addition, we find that stations respond also to the introduction of the signs posting the prices of other stations following the focal station on the same route. When we consider a different control group (column (iii) of Table 3), we obtain similar results. In particular, the interaction between *PostOwn* and *PrcSign* is again negative and significant.

Note that because *PrcSign* has only cross-sectional variation, which depends mainly on the station's position with respect to the sign, the identification of δ_2 is in part due to the different positions of stations from the sign. We would like to distinguish the effect of *PrcSign* from the effect of a station's position, as this second effect might capture additional factors, for example, consumers' forgetting price information acquired previously. In columns (ii) and (iv) of Table 3, we estimate the same model proposed in (1) after controlling for the position of stations. The interaction between *PostOwn* and *PrcSign* is negative and significant even after controlling for the position of the station. The station's position, on the other hand, does not produce any significant effect. These results suggest that the effect of *PrcSign* is not driven by the position of the station. Two stations in the same position show a different response to signs depending on the percentage of their potential consumers who drove in front of the sign.[11] Overall, this evidence seems to confirm that a significant number of consumers are uncertain about prices even after the introduction of signs, and that stations take advantage of the policy's lack of reach by charging higher prices.

3. Model

In this section, we propose a model of consumers' price search and purchase decisions. We divide the highway into several segments, one for each sign installed, as in Figure 1. Each segment is a "market" and consists of $M = 4$ consecutive gas stations along the highway following the location where the price sign is installed. Depending on their driving path, consumers can purchase from any of these stations or from an outside option. For consumers leaving the highway (i.e., at exit c, d, e, or f in Figure 1), the outside option is a local station outside the highway. If they keep driving on the highway past the fourth station and into the next segment (i.e., exit g in Figure 1), the outside option is the first station of the next segment.

We do this for two reasons. First, because of the large number of toll gates on the highway, the total number of unique driving paths (discussed below) becomes very large. Accounting for all such paths makes the problem numerically challenging. Second, and more importantly, the price signs post only the prices of the next four stations, so the prices of the following stations remain unknown even after the introduction of signs. It is also worth pointing out that the four stations after the sign cover, on average, a highway segment of about $27 \times 3 = 81$ km. Consumers interested in purchasing gasoline are likely to consider stations mainly within this distance.

Driving Paths

Each consumer enters a market from one of several locations along the highway (i.e., entry A, B, C, D, or E in Figure 1) and follows one of the J unique driving paths. Depending on the driving path j, the consumer encounters a subset of $M_j \leq 4$ consecutive stations. The driving path also identifies the distance traveled and the consumer's knowledge about prices—when price signs are installed, only consumers driving in front of a sign can observe the prices of the four stations in the market. The driving path is considered exogenous to the consumer's purchase problem, because a change in driving path implies a higher toll and significant increase in traveling time.

Decision Flow

Consumers solve a sequential search problem.[12] Each consumer considers purchasing gasoline from any of the stations encountered. Every time the consumer reaches the point on the highway near a gas station, she decides whether to visit the station, that is, enter the service area and drive up to the gasoline pumps, or to proceed on the highway to the next station. If she decides to continue on, she cannot purchase gasoline from that station. If she decides to visit, the price charged by the station and other characteristics of the station not observed from the highway are revealed, and the consumer decides whether to purchase (and leave the market) or exit the service area and proceed to the next station. The problem is modeled as a dynamic model with finite periods, where each period corresponds to driving up to the next station in the market.[13]

Consumer Uncertainty

Before signs are introduced, consumers do not know what prices are charged by the highway stations. Visiting a station (i.e., search) is costly, but allows the consumer to drive to the pumps and observe the price charged by the station. It also allows the consumer to observe other factors affecting purchase such as waiting lines, temporarily inoperational pumps, etc. (that are not observed by the researcher). After signs are introduced, a consumer driving in front of a sign observes the prices of the four stations in the market.[14] This means that the consumer knows all stations' prices before ever visiting the first station.

Quantity

We assume quantity decisions as exogenous. In principle, after a consumer visits the station and observes a high price, she might decide to reduce the quantity of gasoline purchased. However, this decision implies a higher future search cost, as the next visit to a station is now sooner (because of the lower quantity purchased today). Therefore, even if quantity is exogenous, the model can capture, in terms of aggregate behavior, the effect of price uncertainty and the consumer's trade-off between higher price and higher search costs.

3.1. State Variables and Flow Utility

Each consumer i drives along the path j following the direction of travel and sequentially encounters stations $m = m_{j1}, \ldots, m_{j,M_j+1} \in \mathcal{M}_j$, where $\mathcal{M}_j$ is the subset of stations located along path j. For each station encountered, the consumer has to make two decisions. The first decision is whether to visit the station or not; this decision is made outside the service area where the station is located. The second decision is whether to purchase or not; this decision is made inside the service area, that is, only upon visiting the station. The location of the consumer is given by the identity of the station reached (m), by a dummy variable s indicating whether the consumer is outside ($s = 0$) or inside the service area ($s = 1$), and by the path j taken.

We begin defining the utility a consumer receives during the purchase decision. From inside the service area where station m is located, the consumer can observe both price and quality (including the factors unobserved by the analyst noted previously) of the station, so the utility is known. The consumer decides whether to purchase ($d_{pm} = 1$) or leave the station and proceed on the highway to the next station ($d_{pm} = 0$). Her utility is

$$u(d_{pm}, m, s = 1, j, p_m, \phi_h, \epsilon_m; \theta) = \begin{cases} X_{jm}\beta + \mu\phi_h(\kappa_{jm}) + \alpha p_m + \epsilon_{m1} & \text{if } d_{pm} = 1, \\ \epsilon_{m0} & \text{if } d_{pm} = 0, \end{cases} \quad (2)$$

where X_{jm} is a matrix that includes time-invariant characteristics of the service area where station m is located, such as its size (in square meters), the number of parking spots, whether the food court has a restaurant, the presence of an ATM, public showers, a park for kids, and the number of competing stations that can be reached within 10 km off the highway; κ_{jm} is the number of kilometers driven on the highway by the consumer in path j before reaching station m; and ϕ_h represents the consumer's ideal refill point and depends on κ_{jm}. It captures the likelihood of making a purchase as a function of the distance driven. It is one of the four Weibull functions plotted in Figure 2, which are derived from individual cardholders' data, as discussed previously. The introduction of this variable allows the model to capture consumer heterogeneity in their preferences on where to fill up on the highway. This preference for location is key in geographical competition between firms (e.g., see Chan et al. 2007, Thomadsen 2005).[15] The variable p_m is the price for gasoline charged by station m, $\theta = [\alpha; \mu; \beta]$ is a vector of consumer primitives to be estimated, and ϵ_m is a two-by-one vector of shocks to the purchase utility. These values are revealed to the consumer during her visit and represent factors such as a short versus long line at the pump, inoperational restrooms, and other events affecting the quality of the station that are unobserved to the researcher.

From outside the service area, the consumer does not know the price p_m charged at the station, unless she encountered a price sign at the beginning of her path. Also, the consumer does not know the shock ϵ_m. The purchase utility is therefore unknown when the consumer is outside the service area. Hence, at this stage, consumers form expectations about this utility:

$$E[u(d_{pm}, m, s = 1, j, p_m, \phi_h, \epsilon_m; \theta) \mid \omega_j] = \begin{cases} \int_\epsilon \int_p u(d_{pm}, m, s = 1, j, p_m, \phi_h, \epsilon_m; \theta)\, dF_p dF_\epsilon & \text{if } \omega_j = 0, \\ \int_\epsilon u(d_{pm}, m, s = 1, j, p_m, \phi_h, \epsilon_m; \theta)\, dF_\epsilon & \text{if } \omega_j = 1, \end{cases} \quad (3)$$

where $\omega_j = 1$ denotes a price sign at the beginning of path j; F_p is the price distribution representing the consumer's beliefs, which we discuss below; and F_ϵ is the distribution of the utility shocks.

Next, we define the utility that the consumer receives outside the service area, when making the decision about whether to visit the station ($d_{sm} = 1$) or proceed to the next station ($d_{sm} = 0$). Her utility is

$$u(d_{sm}, m, s = 0, j; \theta) = \begin{cases} -c & \text{if } d_{sm} = 1, \\ 0 & \text{if } d_{sm} = 0, \end{cases} \quad (4)$$

where c is the cost of visiting the station. This cost represents the consumer's opportunity cost of the time spent driving to the MSA. Paying the cost c allows the consumer to move to state $\{m, s = 1, j\}$, observe the purchase utility, and decide whether to purchase. If the consumer decides not to visit station m, she moves to state $\{m + 1, s = 0, j\}$; that is, she drives on to the next station.

The last station encountered is the *outside option* (i.e., station $m = M_j + 1$). Depending on the path j, this is either a station outside the highway or another highway station further along on the same route. The price charged by the outside option, p_0^j, is not observed, so consumers build expectations about it. Consumers not purchasing from any of the M_j stations will purchase gasoline from this station and receive the following utility:

$$E[u(d_{pm} = 1, m = m_{j,(M_j+1)}, s = 1, j, p_m, \phi_h; \theta)] = \int_{p_m} [X_{jm}\beta + \mu\phi_h(\kappa_{jm}) + \alpha p_m]\, dF_p. \quad (5)$$

3.2. Price Expectations

The function F_p in Equations (3) and (5) represents the consumer's price beliefs in a given market. We assume this function to be a (multivariate) normal distribution where the mean is the current average price across stations in the market and the covariance matrix is the current price covariance matrix of stations across markets.

The choice of the model for consumers' beliefs could potentially have an impact on the estimation of search costs, because both beliefs and the price change after the introduction of signs help identify the cost to visit a station. To check the robustness of our results, we provide two alternative models. In the first model, we also use a normal distribution to represent beliefs. However, instead of pooling prices across stations, this time we use historic prices and derive station-specific beliefs. Because prices change significantly over time, we model beliefs of the price difference relative to the outside option (instead of modeling prices directly). In the second model, we assume a discrete uniform distribution where for every event, each station is assigned one of the 4 + 1 prices observed in the market. Results using these alternative price expectations models are reported in the e-companion.

3.3. Consumer Problem and Market Shares

Before reaching any station $m' \in \mathcal{M}_j$ along the path j, the consumer makes a sequence of search (d_{sm}) and purchase ($d_{pm} \mid d_{sm} = 1$) decisions that maximizes the expected sum of the finite number of utilities received along the driving path:

$$\begin{aligned} &V(m = m', s = 0, j \mid p, \phi_h, \omega_j, \epsilon, \theta) \\ &= \max_{\{d_{sm}, (d_{pm} \mid d_{sm}=1)\}_{m=m'}^{m_{M_j+1}}} E\Bigg[\sum_{m=m'}^{m_{M_j+1}} (u(d_{sm}, m, s = 0, j; \theta) \\ &\quad + d_{sm} \cdot u(d_{pm}, m, s = 1, j, p_m, \phi_h, \epsilon_m; \theta)) \mid j, p, \omega_j, \epsilon, \theta\Bigg] \\ &\text{s.t.} \sum_{m=m_{j1}}^{m_{M_j+1}} d_{pm} = 1, \end{aligned} \tag{6}$$

where p is a vector of market prices, and ϵ is a 2-by-m array of utility shocks. For each following station on the path j, the consumer receives the utility from the search decision and, if a visit is made (i.e., $d_{sm} = 1$), the utility from the purchase decision. The constraint implies that the consumer makes only one purchase across all stations encountered; that is, after a purchase, the continuation value is 0. Because all utilities are received within a short time frame, they are not discounted.

After the value function V is solved, the consumer model presented above can be used to derive the market share of each gas station. For each driving path j, the proportion of consumers purchasing at station m' is equal to their probability of being "alive" in state m' (i.e., not purchasing at any station r previously encountered) and both searching and purchasing in that state. This expression is derived as

$$\begin{aligned} &\sigma_{jm'}(j, p, \omega_j, \theta) \\ &= \int_\phi \int_p \prod_{\substack{r<m' \\ r\in\mathcal{M}_j}} [1 - \mathrm{Prob}(d_{pr} = 1 \mid r, s = 1, j, p, \phi, \omega_j; \theta) \\ &\quad \cdot \mathrm{Prob}(d_{sr} = 1 \mid r, j, p, \phi, \omega_j; \theta)] \\ &\quad \cdot \mathrm{Prob}(d_{pm} = 1 \mid m = m', s_m = 1, j, p, \phi, \omega_j; \theta) \\ &\quad \cdot \mathrm{Prob}(d_{sm} = 1 \mid m = m', j, p, \phi, \omega_j; \theta)\, dF_p dF_\phi, \end{aligned} \tag{7}$$

where ideal refill point (with distribution F_ϕ) and expected prices are integrated out. The probabilities of searching and purchasing that are needed to obtain market shares are derived in Section 4.3. The total market share of station m' will weight the market share of each driving path by the percentage of consumers taking that path, as follows:

$$\sigma_{m'}(p, \omega, \theta) = \sum_{j=1}^{J} \sigma_{jm'}(j, p, \omega_j, \theta)\rho_j, \tag{8}$$

where ρ_j is the percentage of consumers following the driving path j, and ω is a vector whose jth entry is ω_j. For a given set of demand parameters (θ), our model generates market shares that depend on price levels (p) and price sign policy (ω). These shares are used to obtain the optimal price equilibrium, which we derive next.

3.4. Supply Side

We now specify a supply-side interaction model between gas stations. Past research on U.S. markets (Hastings 2004) suggests that oil companies might exercise an influence on the price of their affiliated gas stations. In our data, we have not found systematic evidence of such an influence. For example, we tested whether the level of price reduction associated with the introduction of a sign is affected by the number of stations on the sign with the same affiliation; we did not find any significant result, which seems to confirm the relative independence of gas stations from oil companies discussed above.

Also, past research has used consumer search models to explain the price dispersion of markets (see Baye et al. 2006). According to this literature, firms mix their pricing strategies to extract rents from uninformed consumers, so a pure-strategy equilibrium is not guaranteed. In contrast with this literature, which mainly assumes symmetry among firms, in our model, gas stations have different costs and different sources of market power, such as location, brand affiliation, etc., which might contribute to the dispersed price equilibrium observed in the market.[16] Following recent structural work on retail gasoline (Houde 2012), which adopts pure-strategy equilibria and finds that the market is much closer to a competitive structure than a collusive one, we assume a competitive pure-strategy Bertrand–Nash equilibrium. We further recognize the need to make this assumption as a potential limitation of our analysis.

Let A be the size of the market. Also, let p_m and mc_m be the price and marginal cost of gas station m, and σ_m its market share. The profits of m are $\pi_m = (p_m - mc_m) A\sigma_m$, so the pricing behavior predicted for m is given by the standard first-order condition in p_m, which, after substituting for the market shares derived in (8), becomes

$$p_m = mc_m + \left[-\frac{\partial\sigma_m(p, \omega, \theta)}{\partial p_m}\right]^{-1} \sigma_m(p, \omega, \theta), \quad m = 1, \ldots, M. \tag{9}$$

The $M = 4$ equations above provide economic constraints for the estimation of the parameters, discussed next. These equations will be derived for each market and for each period in the data, to match with the observed price at each station.

4. Estimation

4.1. Overview

The typical approach for estimating the model proposed above would be to use data on purchased quantities, by comparing the stations' predicted and observed market shares. By making an assumption on stations' market conduct, the observation of market prices would also allow one to estimate the marginal costs of gas stations as the difference between prices and the predicted margins from the demand model.

In our data, we do not observe quantity information. Instead, we rely on prices. We estimate the model parameters by minimizing the distance between the observed prices and those predicted by the model. To do that, we need a parametric specification for marginal costs. Therefore, we compensate for our lack of demand data by adding more structure on the supply side (costs).

We next discuss the value function corresponding to the demand model and how we derive the probabilities necessary to predict market shares. Then, we set up the estimation problem and discuss the identification strategy.

4.2. Solving the Value Function

A key component of the demand-side specification is the value function V, which must be solved. We specify different value functions, to account for expected prices and heterogeneity across consumers. Overall, we consider $J \times H \times N$ value functions, where J is the number of total driving paths, H is the number of Weibull distributions representing the ideal refill point, and N is the number of nodes used for approximating the distribution representing expected prices. In total, $39 * 4 * 243 = 37{,}908$ unique functions have to be solved in each market period for a given vector of parameter values. More details are provided in the e-companion.

For each value function, we need to find the optimal search and purchase decisions at each station. Because the number of stations encountered is finite, we can solve each value function using backward induction, starting from the last station along the driving path. In addition, we assume the purchase shocks ϵ to be i.i.d. over consumers, stations, and time, according to the type I extreme value distribution. With this choice of distribution, the integration over the errors leads to a closed-form expression of the value function, which we denote by v. We can relax these assumptions and allow for a more flexible error structure at a higher cost in terms of computational burden.

At station m, the value received by the consumer visiting the station is the highest payoff between purchasing at the current station (and leaving the market) and not purchasing (i.e., proceeding back on the highway to the next station). Given the assumption on the purchase shocks, and conditional on expecting a vector of prices p^n in the market, the consumer receives the following value:

$$\begin{aligned} v(m, s=1, j \,|\, p^n, \phi_h, \omega_j; \theta) &= \log[\exp(X_m\beta + \mu\phi_h(\kappa_{jm}) + \alpha p_m) \\ &\quad + \exp(v(m+1, s=0, j \,|\, p^n, \phi_h, \omega_j; \theta))], \end{aligned} \tag{10}$$

where p^n are the consumer's expected prices of the gas stations in the market. The value received prior to visiting station m is the highest payoff between visiting the station (i.e., paying the search cost to observe the purchase utility and decide whether to purchase) and not visiting (i.e., continuing on the highway to the next station). Because there is no shock to the search utility, the value function is simply

$$\begin{aligned} v(m, s=0, j \,|\, p^n, \phi_h, \omega_j; \theta) &= \max\{-c + v(m, s=1, j \,|\, p^n, \phi_h, \omega_j; \theta), \\ &\quad v(m+1, s=0, j \,|\, p^n, \phi_h, \omega_j; \theta)\}. \end{aligned} \tag{11}$$

4.3. Computing Purchase and Search Probabilities

After each value function is computed, we can derive predictions regarding search and purchase decisions at each state of the function. These probabilities will then be plugged into Equation (7) to derive the market share of each station. From Equation (10) and from the assumption on the distribution of the unobserved state ϵ, the probability of purchasing at station m after visiting the station can be derived as

$$\begin{aligned} &\text{Prob}\,(d_{pm} = 1 \,|\, m, s=1, j, p^n, \phi_h, \omega_j; \theta) \\ &= \text{Prob}\,(u(d_{pm} = 1, m, s=1, j, p_m, \phi_h, \epsilon_m; \theta) \geq \epsilon_{m0} \\ &\quad + v(m+1, s=0, j, |p^n, \phi_h, \omega_j; \theta)) \\ &= \frac{\exp(X_m\beta + \mu\phi_h(\kappa_{jm}) + \alpha p_m^n - v(m+1, s, j, |p^n, \omega_j; \theta))}{1 + \exp(X_m\beta + \mu\phi_h(\kappa_{jm}) + \alpha p_m^n - v(m+1, s, j, |p^n, \omega_j; \theta))}. \end{aligned} \tag{12}$$

For the search decision, we do not have utility shocks, as consumers hold the same search cost across different stations and periods.[17] To avoid a step function with respect to the parameter c, which would generate discontinuity in the market share function and therefore complicate the estimation process, we represent the search probability using a smooth logit kernel (McFadden 1989); that is, we introduce small shocks to the consumer's benefit from search. Similar to previous literature (Honka 2014) this strategy allows us to introduce smoothness in the likelihood function. These shocks are logit distributed, with scale factor λ.[18] The search probability can therefore be written as follows:

$$\begin{aligned} &\text{Prob}(d_{sm} = 1 \,|\, m, s=0, j, p^n, \phi_h, \omega_j; \theta) \\ &= \text{Prob}(-c + v(m, s=0, j \,|\, p^n, \phi_h, \omega_j; \theta) \\ &\geq v(m+1, s=0, j \,|\, p^n, \phi_h, \omega_j; \theta)) \\ &= \frac{\exp(A/\lambda)}{1 + \exp(A/\lambda)}, \\ &\text{where } A = v(m, s=1, j \,|\, p^n, \phi_h, \omega_j; \theta) - c \\ &\qquad - v(m+1, s=0, j \,|\, p^n, \phi_h, \omega_j; \theta). \end{aligned} \tag{13}$$

This expression is twice differentiable in θ.

We can now plug the probability of search and purchase in Equation (7) and derive the predicted market share σ_m for each station m, conditional on a given vector of parameters θ. We then substitute the market share in Equation (9) to find the Bertrand–Nash prices predicted by the model.

4.4. Estimation Problem

To find the equilibrium prices predicted by the model, we first need to have a measure of the marginal costs of gas stations. Following the literature (Thomadsen 2005), we use a parametric specification for the stations' marginal costs. In particular, we use the price of crude oil, multiplied by station dummies (to obtain station-specific effect of crude oil). We also include market, period, and brand dummies to add flexibility to the specification.[19] The marginal cost for a given station, market, and period becomes

$$mc_{mkt} = \gamma_{1k} + \gamma_{2t} + \gamma_{3mk} + \gamma_{4mk} \cdot r_t + \xi_{mkt}, \tag{14}$$

where k and t denote market and period; γ_{1k}, γ_{2t}, and γ_{3mk}, are intercepts specific to market, period, and brand; γ_{4mk} is a station-specific coefficient; r_t is the price of crude oil in t; and ξ_{mkt} is an unobserved factor affecting marginal costs.

After plugging market shares and marginal costs in the first-order condition in Equation (9), we can derive the predicted prices and compare them with the observed prices, as follows:

$$P_{mkt} = \gamma_{1k} + \gamma_{2t} + \gamma_{3mk} + \gamma_{4mk} \cdot r_t + \xi_{mkt} - \frac{\sigma_{mkt}(\theta)}{\frac{\partial \sigma_{mkt}(\theta)}{\partial p_{mkt}}}, \tag{15}$$

where P_{mkt} is the observed price of station m. Note that the vector of supply-side parameters γ enters the equation linearly; therefore, it can be conveniently estimated via least squares within the nonlinear search for θ. To estimate θ, instead, we use nonlinear least squares (NLLS). Now, let R denote the array with the price of crude oil multiplied by station dummies, and with market, period, and brand dummies. After rearranging the terms in the equation above, the NLLS residual can be written in matrix notation as

$$\Xi(\theta) = P - R[(R'R)R'\Omega(\theta)] - \Omega(\theta), \tag{16}$$

where $\Omega(\theta)$ is a vector with ith element $[\partial\sigma_i(\theta)/\partial p_i]^{-1}\sigma_i(\theta)$. The NLLS estimator is the value of $\hat{\theta}$ that solves $\arg\min_{\hat{\theta}} \Xi(\theta)'\Xi(\theta)$.

4.5. Identification

We now discuss how the variation in the data allows us to identify the parameters of the model. For each gas station, we have two unknown sets of parameters: the demand primitives (θ) and the marginal cost parameters (γ). In our data, we observe price and longitudinal variation in the informational environment. Because we do not observe demand, instead of using the assumption on market conduct to back out marginal costs, as traditionally done in the literature (e.g., Bresnahan 1987), we use market conduct to identify demand parameters, whereas costs are recovered with a parametric assumption (i.e., Equation (14)).

It follows that the cost parameters are recovered through common variation of the price of crude oil with the marginal cost predicted for each station by the model, conditional on θ. The coefficients of the dummies pick up common variation within periods, within markets, and within brands, which is not explained by the observed variables. Together with prices, the estimation of γ allows us to obtain a measure of the price-cost margins charged by the station.

With price-cost margins on hand, we next discuss the identification of the demand parameters. The price coefficient (α) is identified with the variation of stations' margins due to a change in the price of the outside option, that is, station $M_j + 1$. For a given price of this station, if highway stations are able, on average, to extract a higher margin for a liter of fuel, the model will imply a lower sensitivity to price. The effects of the stations' time-invariant characteristics (β) are instead identified with the variation of these characteristics across stations and the variation of those stations' margins. The coefficient related to consumers' ideal refill points (μ) is identified with the variation across stations of the ideal refill points of consumers passing in front of stations and those stations' margins.

Finally, we discuss the identification of the search cost parameter. The variation that allows us to identify this primitive is the longitudinal variation in the informational environment, which occurs at different times across markets. In some markets, the introduction of price signs occurs early on, whereas in others it occurs later in the data. This idiosyncrasy across stations allows us to exploit longitudinal variation on price information, while still controlling for period-specific effects on cost with time dummies (Equation (14)). The change in the informational environment determines a shift to the stations' margins. Before the price disclosure, a consumer who visits a station does not know the exact price charged at the pump, and to avoid incurring additional search costs, the consumer is willing to accept a price higher than the expected price when driving to the station. The station can therefore mark up its price, up to this search cost. After signs are introduced, a fraction of consumers can observe the price at the pump before visiting the station, so the station can no longer extract the search cost from these consumers.

4.6. The Sample Used for Estimation

We estimate the model described above using data from 24 gas stations along the east side of the A14 route.

We make this selection for two important reasons. First, there are no main junctions on this segment of the highway, so price signs are located every four stations, with no exceptions. Moreover, each station has its price posted on one sign only, and the station is located on the same route where the sign is installed. In other parts of the highway, this is not true, as the price of a station can be posted on a sign on an intersecting route near the station. Second, by zooming in on one segment of the highway, we control for the large heterogeneity across gasoline retail markets in the highway system.

Overall, there are seven price signs installed on this route, covering 28 stations. One of the stations was closed during 2009 and 2010 for construction, and its price was not reported. We exclude this station and the three other stations whose prices were later posted on the same price sign from the analysis. Therefore, we are left with 24 gas stations, representing six markets. In 2008, some of the stations changed oil company affiliations and others underwent some construction. However, these stations did not stop or suspend their retail activities. To control for these changes, we focus our analysis on the 12-month period between July 1, 2009, and July 1, 2010. For three of the six markets, this period includes also the date of the price sign introduction (December 2009). For the other markets, the price sign was introduced before the period of the analysis.

5. Results

The results of the estimation are shown in Table 4. The mean coefficient of price is negative and significant. Its magnitude suggests that consumers are very sensitive to price changes. The overall median elasticity across markets before signs are introduced is −10.9. The search cost is positive and significant, confirming consumers' perceived inconvenience for visiting gas stations. Transforming this parameter from utils to euros, we find that the cost of a gas station visit is about €1.97.[20]

We next look at the coefficients of the time-invariant station characteristics. The size of the area occupied by the station, the presence of a restaurant, and a higher number of competitors within 10 km of drive are all characteristics linked with higher-share stations. In contrast, stations with ATM service, park/camp areas, and larger parking spaces have lower shares. Although some of these characteristics are of interest to consumers and should increase stations' shares, note that these estimates are recovered with cross-sectional variation, so they should not be given a causal interpretation.[21] Overall, some stations have market shares that are very small (below 0.1%), whereas a few enjoy relatively large percentages of the total demand (3.3%). The coefficient related to the ideal refill point is positive and significant; consumers' preferences across stations depend on their distance and follow the pattern suggested by the individual data shown in Figure 2. To measure the substitution effect generated by this parameter, we estimated the cross elasticity with respect to distance between stations. We find that when the consumer's distance from a station increases, the consumer substitutes from the closest stations preceding the station on the driving path to this station. On average, a 1% increase in the distance between the consumer and the station leads to a 0.25% share loss in the station immediately preceding and a 0.04% loss in the station before that; in contrast, the stations following are hardly impacted by the change in distance. As we discussed previously, the empirical distribution of ideal refill points suggests that most consumers on the highway prefer to travel some distance before purchasing fuel. Therefore, a station located further away tends to be preferred over the nearby station, ceteris paribus.

According to the supply coefficients, reported in the right columns of Table 4, the marginal cost estimated by the model is large compared with the final price charged. In terms of percentage of the final price, the first, second, and third quartiles of the distribution of marginal costs across stations and periods are 91.2%, 92.3%, and 93.7%, respectively. This means that stations are charging razor-thin margins.

5.1. Validation and Robustness

The median elasticity above is relatively similar to what has been found in previous research using demand data. Using cardholders' data from a loyalty program, Rossi (2017) finds that the median elasticity across markets outside the highway is −10.3. The estimates found for the marginal costs seem to be confirmed by industry reports, which document marginal costs close to 95% of the price paid by the consumer at the pump (Tabarelli 2010).

Next, to validate the market share estimates, we combine our traffic data with the cardholders' transaction data.[22] After including the 20% of consumers who are estimated to purchase without a card, we find that the market shares of the affiliated stations located in the highway vary between 0.2% and 2.0%. These estimates are similar to most of the market shares estimated by our model.

Furthermore, we also check the predictive validity of the model. We compare the shift in prices from installing the signs observed in the data with those predicted by the model (Figure 3). To control for the volatility of fuel price, we consider differences with respect to the average national price. After the introduction of signs, the observed price at highway stations decreased by €1.80 per liter (from €3.02 to €1.22), whereas the model predicted a reduction of €1.86

Table 4. Estimates of the Full Model

Demand parameters		Supply parameters	
Variable name	Coefficient	Variable name	Coefficient
Brand 1	−3.795 (0.256)	*Brand 1*	0.037 (0.010)
Brand 2	−5.357 (0.225)	*Brand 2*	0.007 (0.009)
Brand 3	−4.645 (0.188)	*Brand 3*	0.027 (0.010)
Brand 4	−5.261 (0.230)	*Brand 4*	0.079 (0.010)
Brand 5	−4.131 (0.275)	*Brand 5*	0.065 (0.012)
Brand 6	−7.174 (0.203)	*Brand 6*	0.014 (0.009)
Brand 7	−6.347 (0.133)	*Brand 7*	—
Area size (square meters)	5.775 (0.499)	*Sign 1*	−0.089 (0.008)
Restaurant	0.340 (0.093)	*Sign 2*	−0.059 (0.009)
ATM	−1.326 (0.120)	*Sign 3*	−0.120 (0.008)
Showers	−0.147 (0.095)	*Sign 4*	−0.112 (0.008)
Park/camp area	−1.355 (0.173)	*Sign 5*	−0.046 (0.008)
Parking spots	−0.694 (0.065)	*Sign 6*	—
No. competitors within 10 *km*	0.058 (0.006)	*Crude oil* × *station* dummy	✓
Refill point	0.015 (0.004)	*Period* dummy	✓
Price	−8.670 (0.681)		
Search cost	0.364 (0.010)		

Notes. The coefficients of the demand model are estimated using nonlinear least squares. The coefficients of the supply model are estimated within the nonlinear search using ordinary least squares. Standard errors are in parentheses.

(from €3.04 to €1.18), suggesting a relatively good fit with the data.

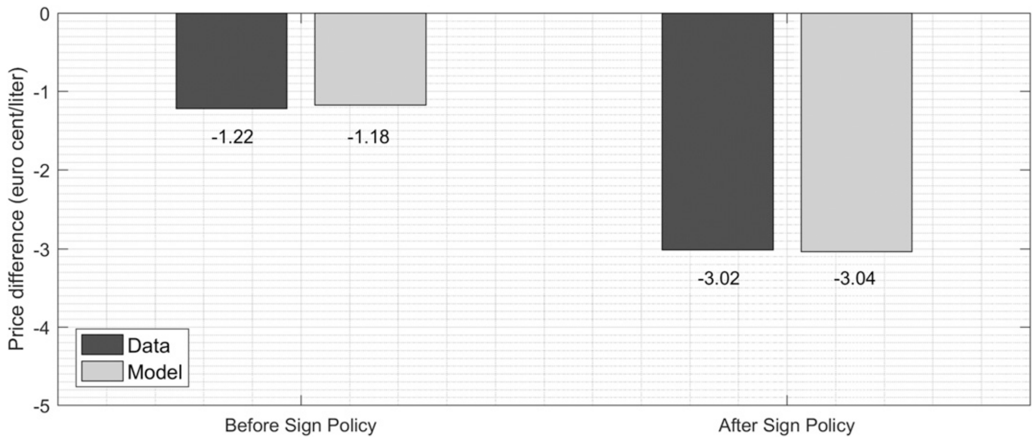
We tested the fit of several alternative models. We tried different specifications for marginal costs, such as changing station and period dummies. The results are relatively robust, as long as the specification is flexible enough to capture variations across weeks and brands. To test the robustness of our results to the price expectation model assumed, as noted previously, we considered alternative specifications. The comparison across estimates, provided in the e-companion, suggest that the results of our model are relatively robust across specifications. We also tested a nonlinear specification of ideal refill point by adding a quadratic term of the variable in Equation (2); the additional estimate, however, is not significant. Finally, we changed the scale factor λ in Equation (13) to check the robustness of our result to different degrees of smoothing.

6. Counterfactuals

6.1. Price Uncertainty and Stations' Margins

How much of the gas stations' margins are the result of consumers' uncertainty about gasoline prices? We run three simulations. In the first simulation, we compute the price equilibrium in which all consumers have full information, that is, consumers are given information on stations' prices at the beginning of their path. We consider this the benchmark case. This scenario is unobserved in

Figure 3. (Color online) Observed vs. Predicted Prices Before and After the Introduction of Price Signs



Note. The prices are reported as differences between the prices of the stations on the highways and the national price.

the data but is of interest to quantify the overall effect of information. In the second simulation, we compute the new equilibrium when price signs are installed. Consumers obtain price information only if they transit in front of a sign. This was the actual market situation in the data after the policy was implemented. In the third simulation, we compute the price equilibrium of gas stations when no price signs are installed. Consumers can obtain prices only by visiting stations. This was the situation before the policy was implemented.

Results are reported in Table 5. For each simulation, we compute the average price and margin across stations in the market. In the table we report the medians across markets and time periods. The first column shows that, as expected, stations charge higher prices as consumers become less informed about prices. The next columns show the levels of margins in each scenario. When all consumers are knowledgeable about prices, the stations obtain a 7.5% margin, on average, which increases to 8.1% when information is instead conveyed through the current signs. With no information available, stations can charge an even higher margin of 10.1%. Note that from the case of no price information, the introduction of signs implies a €0.02 reduction in equilibrium margins, from €0.119 to €0.099 for each liter sold. When consumers are fully informed, margins drop in total by €0.028.

Overall, our analysis suggests that stations charge 31.3% higher margins because of consumers' lack of price knowledge. About two-thirds of this informational rent disappears after the introduction of price signs. The current positioning of the signs conveys price information to consumers coming from previous markets, which represent a large number of the consumers in the market. These consumers have traveled for some time already, so they need fuel and are more likely to stop at the gas stations. However, there are many other consumers who are not reached and may be interested in purchasing fuel as well. These are consumers who get onto the highway after a sign and so do not transit in front of it. The remaining part of the informational rent is eliminated by making the price information accessible to these consumers, as this increases competitive pressure on gas stations.

Table 5. Price Markups Charged by Gas Stations Predicted for Different Levels of Price Uncertainty

		Margin		
	Price	Euros	Percentage	Informational rent (% of margin)
Perfect information	1.295	0.091	7.5	
	[1.277, 1.314]	[0.070, 0.109]	[5.8, 9.0]	
Information on signs	1.304	0.099	8.1	9.3
	[1.289, 1.322]	[0.080, 0.012]	[6.6, 9.6]	[5.9, 13.5]
No information	1.333	0.119	10.1	31.3
	[1.321, 1.350]	[0.107, 0.136]	[9.1, 11.5]	[22.2, 51.8]

Notes. The 95% confidence intervals, reported in brackets, are computed using bootstrapping based on 100 draws. Margins are calculated as $(p - c)/c$.

6.2. Evaluation of Alternative Policies

We now evaluate the effect of the price disclosure policy on welfare. The lack of price knowledge decreases consumers' welfare for two reasons. First, price uncertainty softens competition, and therefore consumers end up paying higher prices for the same good. Second, when prices are unknown, consumers might need to make multiple visits (i.e., pay higher search costs) before purchasing, and/or settle for a less preferred gas station to avoid further search costs. In this section, we quantify the welfare change due to the price information available to consumers via signage.

The proliferation of smartphones and Italian sites similar to Gasbuddy in the United States may suggest a more "crowdsourced" approach to information dissemination in this market. This would certainly be true for consumers able to plan their trips and stops on the highway in advance. Nevertheless, policy makers cannot just rely on smartphones and their diffusion to provide information. Furthermore, for those who have not gathered the information in advance or want to update it, the need to check prices on the phone could distract drivers, which in turn could lead to other negative consequences. Instead, in our analysis, we evaluate the effect of alternative information disclosure policies based on signage. In particular, we compare the current policy, where a sign posts the prices of the next four stations along the road, with an alternative policy, where a sign is installed before each station and posts only the price of the nearby station. A similar disclosure policy is currently in effect for off-highway stations in many U.S. states and other countries as well. According to this policy, gas stations are required to advertise their prices using a nearby sign. Are consumers better off with more price information, or is less information (just one price posted) provided more frequently (before every station) more helpful to consumers?

We simulate four scenarios. In the first scenario, there are no price signs. In the second scenario, there are price signs as currently mandated by the policy, that is, one sign every four stations, posting four prices. In the third scenario, each station is preceded by a sign posting only the price of that station. In the final scenario, which we use as benchmark, consumers are fully informed about prices. In each scenario, we first derive the equilibrium prices, and then, based on these prices, we compute both the profit function of gas stations and the compensating variation for consumers. We obtain the compensating variation by computing the value function that consumers expect at the beginning of their trip and dividing it by the marginal utility of income.

The results of the simulation are reported in Table 6. For each simulation, we compute the average price and profit across stations in the same market, and the average compensating variation across consumers on different paths. In the table, we report the median differences across markets and time periods between each policy scenario and the full-information scenario. Compared with the full-information scenario, the lack of information costs each driver €0.57.[23] The cost is reduced to about €0.19 after the introduction of current signs. This means that the current policy is worth about €0.38 to each customer every time she takes the highway. The availability of information also affects the profitability of gas stations by reducing their market power, and thus their profit. However, the increased welfare of consumers compensates for the profit loss, so the total welfare increases with the level of information available to consumers. In the third row of Table 6, we report the effect of introducing one sign per station. Despite providing consumers with less price information per sign, this policy reaches a better consumer welfare level than the current policy, because of the increase in the number of signs. In terms of compensating variation, this new policy is 50% (19/38) more effective than the policy currently applied. It allows consumers entering the highway at any location to receive some price information. The small difference in compensating variation between this scenario and the full-information scenario suggests that the station-specific signs provide consumers with sufficient price information. Additional price information on subsequent stations is not necessary.

Table 6. Welfare Calculations by Price Uncertainty

	Consumers' welfare (Δ €)	Profits (Δ €)
No signs	−0.567	0.063
	[−0.677, −0.485]	[0.309, 0.023]
Current signs	−0.193	0.050
	[−0.281, −0.073]	[0.223, 0.016]
Station-specific signs	−0.000	0.000
	[−0.000, −0.000]	[0.000, 0.000]

Note. The 95% confidence intervals, reported in brackets, are computed using bootstrapping based on 100 draws.

7. Conclusions

Despite selling a relatively homogeneous good, gas stations charge prices above their costs. One reason they can do this is because of consumers' scarce knowledge of market prices. This softens competition among stations. Using a longitudinal change in price information available to consumers on an Italian highway, this paper measures the effect of price information on gas stations' market power. Stations charge 31% higher margins just because of the (lack of) price information of consumers. Only part (about two-thirds) of this rent is eliminated by the current price advertising. A more traditional disclosure policy, where each station's price is advertised by one sign, would help reduce margins by an additional one-third, as consumers are willing to trade more information per signs with more signs.

Our analysis is based mainly on observed price levels. On the demand side, we observe the traffic flow of consumers and their need for fuel, but not actual purchases. To estimate the demand primitives, we therefore made a number of assumptions on the supply side, including the market conduct of gas stations. However, the nature of interactions between stations could differ over time and across markets. Despite these limitations, we believe that quantifying the value of information to consumers is an important role for marketers (Bachman 2013) and for policy makers. By combining the outcomes of a natural experiment with a structural model, researchers may be better able to quantify the effects of providing such information.

Acknowledgments
The authors thank the senior editor, the associate editor, and the referees, Elisabeth Honka, Nitin Mehta, Vithala Rao, Stephan Seiler, Raphael Thomadsen, and Raluca Ursu, seminar participants at Kellogg, Goethe, IE Business School, Purdue, Marketing Dynamics conference, SICS conference, for helpful comments.

Endnotes
[1] In the United States, the cost of the crude oil alone consists of 57% of the total price paid at the pump. We computed this price change using data available at the U.S. Energy Information Administration (http://www.eia.gov) (see http://www.eia.gov/petroleum/gasdiesel/gaspump_hist.cfm).

[2] U.S. Department of Energy (http://www.eia.gov).

[3] See, for example, Jim Motavalli, "Why do gas stations post their outrageous prices? Because, well, they always have." CBS MoneyWatch (May 6, 2011), https://www.cbsnews.com/news/why-do-gas-stations-post-their-outrageous-prices-because-well-they-always-have/.

[4] The company is not allowed by law to contractually set prices; it can only provide suggestions.

[5] To find the closest path and its driving distance, we use the Google Maps application programming interface.

[6] Other distributions useful for describing the probability of an event occurring over time, such as the lognormal, do not have as good a fit with our data.

[7] The identity of the oil company providing the sample is protected by a confidentiality agreement, so we cannot provide details about the number and locations of the stations affiliated with the company. The reader should consider that this is one of the major players in the Italian gasoline retail market, and its stations fully cover the entire national territory, both inside and outside the highway.

[8] Results are available from the authors.

[9] When using stations outside the highway as the control group, this variable is the difference between the price of the station and the price of the related brand of gasoline at the national level.

[10] Note that because *PrcSign* varies across stations but not over time, its main effect is captured by the stations' fixed effects.

[11] For brevity, we omit the discussion of the other model parameters. Interested readers may obtain that from the authors.

[12] Differently from the traditional sequential search problems (Weitzman 1979), in this model the order in which stations are encountered is not selected by consumers, but exogenously given.

[13] The model setup is a two-stage search model similar to that in Seiler (2013).

[14] The prices of the other highway stations in the next markets are still unknown.

[15] With the addition of consumers' ideal refill points, we model the substitution pattern among inside-good stations directly as a function of their distance from one another.

[16] We also checked the data for any evidence of mixed versus pure strategy. In particular, we checked the percentage of times that each station changing its price also changes its ranking compared with the prices of the other stations in the market. This exercise is similar in spirit to past literature (Chandra and Tappata 2011). Frequent changes would suggest that stations are seeking to vary their price position regularly, thus mixing their strategies. We instead find that the median station changes rankings 35% of the time, which is relatively low. This evidence, however, is in no way conclusive.

[17] Because the consumer reaches all stations in the market in a relatively short time frame, an assumption of independent shocks at each station would be unrealistic and very likely violated.

[18] The scale factor changes the degree of smoothing. We set $\lambda = 0.1$ and assessed the sensitivity of our estimates to the scale factor.

[19] Note that station dummies cannot enter independently, as their coefficients cannot be identified.

[20] To obtain this estimate, we use the median number of liters purchased per stop, which is 46.9. This number was obtained using the cardholders' transaction data. The cost is derived as $0.364 \cdot (46.9/8.670)$.

[21] For example, stations with ATMs might simply be located in more competitive markets.

[22] Note that the transaction data cover a different period than our analysis (October 2008 to May 2009).

[23] This amount considers the median number of liters purchased per stop (46.9), which is obtained using the cardholders' transaction data.

References
Bachman J (2013) The genetically modified burrito: Chipotle tells all. *Bloomberg Business* (June 18), https://www.bloomberg.com/news/articles/2013-06-18/the-genetically-modified-burrito-chipotle-tells-all.

Baye MR, Morgan J, Scholten P (2006) Information, search, and price dispersion. Hendershott T, ed. *Handbook of Economics and Information Systems* (Elsevier, Amsterdam), 323–375.

Bresnahan TF (1987) Competition and collusion in the American automobile market: The 1955 price war. *J. Indust. Econom.* 35(4): 457–482.

Chan TY, Padmanabhan V, Seetharaman P (2007) An econometric model of location and pricing in the gasoline market. *J. Marketing Res.* 44(4):622–635.

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

Dickson PR, Sawyer AG (1990) The price knowledge and search of supermarket shoppers. *J. Marketing* 54(3) 42–53.

Gauri DK, Sudhir K, Talukdar D (2008) The temporal and spatial dimensions of price search: Insights from matching household survey and purchase data. *J. Marketing Res.* 45(2):226–240.

Gaynor M, Propper C, Seiler S (2016) Free to choose? Reform, choice and consideration sets in the English national health service. *Amer. Econom. Rev.* 106(11):3521–3557.

Hastings J (2004) Vertical relationship and competition in retail gasoline markets: Empirical evidence from contract changes in southern California. *Amer. Econom. Rev.* 94(1):317–328.

Hong H, Shum M (2006) Using price distributions to estimate search costs. *RAND J. Econom.* 37(2):257–275.

Honka E (2014) Quantifying search and switching costs in the U.S. auto insurance industry. *RAND J. Econom.* 45(4):847–884.

Houde J-F (2012) Spatial differentiation and vertical mergers in retail markets for gasoline. *Amer. Econom. Rev.* 102(5):2147–2182.

Iyer G, Seetharaman P (2008) Too close to be similar: Product and price competition in retail gasoline markets. *Quant. Marketing Econom.* 6(2):205–234.

McFadden DL (1989) A method of simulation moments for estimation of discrete response models without numerical integration. *Econometrica* 57(5):995–1026.

Mehta N, Rajiv S, Srinivasan K (2003) Price uncertainty and consumer search: A structural model of consumer search formation. *Marketing Sci.* 22(1):58–84.

Moraga-González JL, Sándor Z, Wildenbeest MR (2017) Prices and heterogeneous search costs. *RAND J. Econom.* 48(1):125–146.

Pennerstorfer D, Schmidt-Dengler P, Schutz N, Weiss C, Yontcheva B (2016) Information and price dispersion: Theory and evidence. Working paper, Austrian Institute of Economic Research and Vienna University of Economics and Business.

Rossi F (2017) Lower price or higher reward? Measuring the effect of consumers' preferences on reward programs. *Management Sci.* 64(9):4451–4470.

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

Seiler S (2013) The impact of search costs on consumer behavior: A dynamic approach. *Quant. Marketing Econom.* 11(2):155–203.

Sen B, Shin J, Sudhir K (2011) Demand externalities from co-location: Evidence from a natural experiment. Mimeo, Yale School of Management, New Haven, CT.

Stahl DO (1989) Oligopolistic pricing with sequential consumer search. *Amer. Econom. Rev.* 79(4):700–712.

Tabarelli D (2010) La rete carburanti nel 2009: Il confronto con l'Europa [The fuel network in 2009: Comparison with Europe]. Technical report, Nomisma Energia, Bologna, Italy.

Thomadsen R (2005) The effect of ownership structure on prices in geographically differentiated industries. *RAND J. Econom.* 36(4):908–929.

Varian HR (1980) A model of sales. *Amer. Econom. Rev.* 70(4):651–659.

Weitzman M (1979) Optimal search for best alternative. *Econometrica* 47(3):641–654.