CESifo
AREA CONFERENCES 2020

Economics of Digitization

Munich, 19–20 November 2020

Market Transparency and Consumer Search - Evidence from the German Retail Gasoline Market

Simon Martin

CESifo

Market Transparency and Consumer Search - Evidence from the German Retail Gasoline Market

Simon Martin*

September 2020

Abstract

We study a novel trade-off in market transparency regulation by estimating a structural model of the German retail gasoline market. Transparent environments enable easy price comparisons and match findings. Restricting transparency such that only the cheapest offers are shown induces firms to compete for attention, but matching is inefficient. We find that there is an inverse u-shaped relationship between consumer welfare and market transparency. Consumer welfare is maximal when only the first 20% of prices are shown, which decreases consumer expenditures by 1.2%. Our framework allows estimating games of incomplete information with very lax data requirement.

Keywords: market transparency, consumer search, awareness, consideration sets, retail gasoline prices

JEL Codes: D22, D43, D83, L13, L50

---

*Düsseldorf Institute for Competition Economics (DICE). Email: simon.martin@dice.hhu.de. I would like to thank Daniel García, Alessandro Gavazza, David Genesove, Marc Goñi, Bernhard Kasberger, Maarten Janssen, John Morgan, Karl Schlag, Philipp Schmidt-Dengler, Frank Verboven, Matthijs Wildenbeest, and participants in seminars at the VGSE, University of Mannheim, NHH Bergen, Copenhagen Business School, University of Copenhagen, Universidad Carlos III de Madrid, DICE, the QED Jamboree 2018 (Alicante), the ZEW Summerworkshop 2018 (Mannheim), EARIE 2018 (Athens), and EEA-ESEM 2019 (Manchester) for useful comments and suggestions. Funding from the Austrian Science Fund (P 30922) is gratefully acknowledged.

1 Introduction

Regulating market transparency is an important policy issue in various jurisdictions. The main objective of this regulation is to provide consumers with better information and thereby tighten competition. Several countries have introduced government-operated comparison websites in the last decade, e.g., for financial products, energy markets, and retail gasoline markets. Simplified comparisons should reduce information frictions as a source of market power and lead to lower prices (Varian, 1980, Stahl, 1989, Wolinsky, 1986). However, firms may also find it easier to monitor their competitors, making the market more prone to tacit collusion and thus leading to higher prices (Kühn and Vives, 1995, Ivaldi et al., 2003).

These considerations suggest that full transparency is consumer welfare optimal as long as it primarily affects consumers. In this paper, we provide both theoretical and empirical evidence that this intuition is wrong. It is true that the "comparison effect" makes consumers worse off since they can compare only a smaller set of options. Yet one also needs to take into account the equilibrium response of firms. Restricting transparency by only informing consumers about cheap offers induces firms to compete fiercely for precious spots in the consumers' consideration set. We call this the "attention effect." When the attention effect is sufficiently strong, prices may be lower in a restricted transparency regime.

In addition, there is also a "matching effect" when products are horizontally and vertically differentiated, e.g., because of travel costs. Consumers may be worse off in an environment that induces low prices at the expense of low match quality.

This paper complements the literature on market transparency (Luco, 2019, Dewenter et al., 2016, Ater and Rigbi, 2017, Rossi and Chintagunta, 2016) by describing and quantifying the attention and matching effect. We also make a methodological contribution by developing a framework that allows estimating games of incomplete information using only publicly available data.

Our aim is to understand the overall effect limiting transparency has on prices and consumer welfare. In particular, we are interested in whether the benefits due to the attention effect outweigh the downside owing to the matching effect. We set up a novel model that includes both effects and take it to the data. Retail gasoline markets are an ideal environment for studying these forces since product differentiation can be neatly modelled and estimated. The rich history of prices is readily available. Specifically, we provide an answer to the following question:

What is the effect on prices and match quality should Germany change the market transparency regulation? The economic forces revealed in this paper also apply to numerous other markets and industries with information frictions.

In order to address these questions, we develop and estimate a structural model of the German retail gasoline market. Both the attention and the matching effect arise naturally in our environment. Given estimated model parameters, we conduct counterfactual policy simulations by computing equilibrium prices under different transparency regimes.

Our main finding is that there is an inverse u-shaped relationship between market transparency and consumer welfare. Consumer welfare is maximal at intermediate levels of market transparency. The exact level of course depends on the environment, but for the German retail gasoline market, showing only 20% of stations is optimal. In that case, quantity-weighted prices decrease by 1.2% and consumer welfare increases by 3.9% of firms' current margins. Although prices fall further when less then 20% of stations are shown, consumer surplus decreases due to inefficient matching. Online consumers are only informed about low-quality stations, and high-quality stations price exclusively to their loyal customer segment, making consumers worse off. Hence, the relationship between consumer welfare and market transparency is non-monotonic.

Our model incorporates key features of markets with information frictions. "Online consumers" become informed about stations and prices online. The regulator chooses the "transparency regime," i.e., a cutoff point ρ such that only the ρ% cheapest prices are displayed online. Thus, consideration sets depend endogeneously on prices, giving rise to both the attention and matching effect. Since December 2013, German gasoline stations have been obliged by law to transmit prices to a central database operated by the Market Transparency Unit (MTU). Various websites access the MTU's database and provide consumers with real-time price information. Our model allows us to quantify both the attention and the matching effect and, given estimated parameter values, we conduct counterfactual policy simulations.

These counterfactual scenarios are motivated by observed variation in the transparency regulations of retail gasoline markets across countries. For instance, Germany and France provide websites that are fully transparent. Consumers can learn all the prices they might possibly be interested in. Conversely, Austria's website provides only restricted transparency: For any given address, only the first half of the nearby prices are shown. We explain the reason for this restriction in detail below.

We find that the attention effect is very strong. In the first counterfactual scenario we simulate Germany adopting the Austrian regime of showing only the cheapest 50% of stations. Quantity-weighted prices decrease by 0.6% and average prices decrease by 0.5%. Decomposing the total effect, we find that the comparison effect increases average prices by 1.1%. Since the attention effect is stronger, decreasing average prices by 1.6%, the net effect on prices is negative.

The undesired matching effect of restricting transparency stems from low-quality firms with lower marginal costs. These are more inclined, on average, to charge low prices in order to appear on the restricted price list and hence in the consumers' consideration set. Thus, the price list more often contains low quality stations, making consumers fuel their car at cheaper stations with fewer amenities and possibly further away from them. This partially mitigates the positive effect on consumer surplus through lower prices. The increase in consumer surplus is exclusively through uninformed consumers who benefit from lower prices. Online consumers, on the contrary, are actually worse off under restricted transparency due to inefficient matching.

Our paper contributes to the literature by overcoming both theoretical and empirical challenges in novel ways. In our model, intertemporal equilibrium price dispersion arises due to private-information marginal cost shocks, resulting in a game of incomplete information. This approach eliminates Bertrand-type undercutting motives and gives rise to pure strategy equilibria. Our environment allows for a close link between theory and reality with respect to the representation of consideration sets. Online consumers are informed exactly about all the stations displayed online. We thereby avoid the necessity of introducing additional noise in the information acquisition and hence an additional 'information elasticity' parameter to be estimated (Sovinsky Goeree, 2008, Honka et al., 2017, Dinerstein et al., 2018).

Empirically, a major challenge in our environment is the unavailability of demand-side data or market shares. Therefore, standard approaches like Berry et al. (1995) that simultaneously estimate supply- and demand-side parameters and implied elasticities are not feasible. Our approach is based on Thomadsen (2005). We exploit the fact that different brands are observed repeatedly in different market configurations. For instance, a Shell and a BP station may compete with each other in two different markets, where the distances between them differ. We assume that demand is deterministic given prices and hence the only source of noise stems from marginal cost shocks. Then we can use the firms' first-order conditions to estimate the structural parameters due to the non-linearities in optimal pricing with respect to distance. Under a similar argument, variation in the number of stations per market and variation in the input prices in a given market over time independently identify the parameters of interest. We combine all these sources of variation and additionally employ macro moments to enhance efficiency.

This estimation approach requires detailed information about consumers' locations in order to proxy for demand characteristics. We exploit building information publicly available on Open Street Maps to construct an estimate of population density in 1x1km large cells. Distances to gas stations vary across cells, allowing for very rich substitution patterns. Alternative approaches proposed in the literature are based on consumers' commuting patterns (Houde, 2012, Pennerstorfer et al., 2020), which require highly detailed commuting data that is rarely available on a large scale.

Similar to approaches used in the auctions literature (Guerre et al., 2000, Athey and Haile, 2002, 2007), we apply a two-stage estimation routine. In the first stage, we fit the equilibrium price distribution for each station, conditional on the oil price. These first-stage estimates are then used as equilibrium beliefs about competitors' prices in the second stage. Finally, we back out the implied marginal cost shock from the firms' first-order conditions, given a parameter vector, and proceed with a GMM estimation. Since the equilibrium of our incomplete information game is somewhat involved and requires iteratively solving a set of integral equations, full information approaches are essentially unfeasible. Our two-step approach eliminates the computational burden of solving integral equations in each estimation step.

The Austrian regulator enacted a restricted transparency regime due to concerns of tacit collusion in a dynamic setting (Ivaldi et al., 2003, Albæk et al., 1997, Schultz, 2005, Petrikaitė, 2016). Nevertheless, we are convinced that considering a static setting is appropriate in our environment. The retail gasoline market is already under the close scrutiny of competition authorities in several OECD countries, including Austria (Bundeswettbewerbsbehörde, 2011) and Germany (Bundeskartellamt, 2011, Bundesministerium für Wirtschaft und Energie, 2018). Hence, even if more transparency could lead to more collusion and thus higher prices in theory, it seems very unlikely that this would be the case in practice. Additionally, even if average expected prices increased, this still need not be detrimental for consumer welfare; what we actually should be concerned about is the average price that consumers actually pay, and how far they have to drive in order to make a purchase. Thus, instead of explicitly modelling a dynamic environment, we estimate the static model and provide convincing evidence that this model adequately captures key market characteristics. Finally, we provide compelling evidence that *reducing* transparency is in the consumers' interest, even if it affects only the consumer side. If there is an additional response on the firms' side, the case for reducing transparency becomes even stronger.

We proceed as follows. In the next subsection, we relate our paper to other existing literature. Section 2 provides an overview of the industry background and the data, as well as the data collection process and descriptive evidence. The model is developed and described in Section 3. The estimation and identification strategy as well as parameter estimates are presented in Section 4. We show results from counterfactual policy analysis and several robustness checks in Section 5, and conclude in Section 6.

1.1 Related literature

Our paper relates to various strands of the literature. Government-mandated transparency initiatives have recently triggered several empirical papers (Luco, 2019, Dewenter et al., 2016, Ater and Rigbi, 2017, Rossi and Chintagunta, 2016, Montag and Winter, 2020). These papers typically conduct a difference-in-differences analysis comparing price levels before and after the introduction of a transparency website, which allows considering the two extremes of the spectrum of either no transparency versus full transparency. As our results show, intermediate levels of transparency may actually be optimal from a consumer surplus perspective. Our structural estimates also reveal important insights on the underlying channels and allow additional counterfactual simulations.

There is a vast literature on retail gasoline markets. Eckert (2013) provides a comprehensive overview. Besides Luco (2019) and Rossi and Chintagunta (2016), Nishida and Remer (2018) is close to this paper. Nishida and Remer (2018) estimate the consumers' search cost distribution for geographically isolated markets. They show that, theoretically, policies that decrease both the mean and variance of the search cost distribution in a particular way may lead to higher prices. Measuring how exactly transparency initiatives alter the search cost distribution is almost impossible. In our model information acquisition reflects exactly how price comparison websites typically operate: Consumers are more likely to be informed about lower prices. Firms in Nishida and Remer (2018) are vertically differentiated. We also allow for horizontal differentiation and thus the matching and the attention effect arise.[1,2]

In the literature to estimate search costs, it is common to exploit the indifference condition of firms in a mixed-strategy equilibrium. This enables estimations when market shares are unavailable (Hong and Shum, 2006, Moraga-González and Wildenbeest, 2008). Wildenbeest (2011) extends this setup to a model of vertical differentiation. This approach is not applicable to our setting where stations are also horizontally differentiated due to their spatial location. Since in our setting firms play an equilibrium in pure strategies, estimation based on indifference conditions in mixed-strategy equilibria is not feasible.

The theoretical literature on information frictions is very broad and has become increasingly relevant recently due to the rise of the Internet. Also in our model some consumers are not informed, whereas others have the possibility of getting informed through a central website that controls the flow of information. An important difference is that in many of these models firms play mixed strategy equilibria (Varian, 1980, Burdett and Judd, 1983) whereas in our model there are private-information marginal cost shocks and firms play pure strategies given their cost shock. Endogenous awareness sets are already considered, for example, in Butters (1977), Sovinsky Goeree (2008), Honka et al. (2017), and Moraga-González et al.

[1]Other re-occurring themes in the literature on gasoline markets, not specifically discussed here, include asymmetric pass-through of input-cost shocks (Borenstein and Shepard, 1996, Borenstein et al., 1997, Chandra and Tappata, 2011), Edgeworth cycles (Noel, 2007a,b, Wang, 2009, Eibelshäuser and Wilhelm, 2017), regulation of frequency of price adjustment (Obradovits, 2014, Dewenter and Heimeshoff, 2012) and consumer learning about common input prices (Dana, 1994, Janssen et al., 2011). We are mostly interested in price levels in this paper. In order to purge out the possible confounding effects of intra-day pricing dynamics, we only consider prices at 5pm every day. The issue of learning does not arise in our setting because consumers only consider stations whose price quote they have already obtained online.

[2]There are also several empirical papers specifically about the German market (Haucap et al., 2016, 2017, Dewenter and Heimeshoff, 2012, Dewenter et al., 2016, Montag and Winter, 2020, Horvath, 2019) and the Austrian retail gasoline market (Pennerstorfer and Weiss, 2013, Pennerstorfer et al., 2020). Since the scope of these papers is very different to ours we do not discuss them in detail here.

(2018). In these models awareness typically stems from the explicit advertising expenditures of firms. In our model, the sole means through which firms get the chance to advertise their prices are with the prices themselves: The probability of making a consumer aware of their product depends wholly on prices, since the government website ranks cheaper products higher. Finally, there are several recent approaches to estimate heterogeneity in consideration sets across consumers (Heiss et al., 2016, Abaluck and Adams, 2020, Barseghyan et al., 2020). In our model all consumers who visit the website have the same consideration set.

Online price comparison websites have also strongly influenced the literature on prominence and ordered search (Armstrong et al., 2009, Armstrong and Zhou, 2011, Armstrong, 2017). Our theoretical result, that prices decrease as firms compete for prominence (attention) through prices, is reminiscent of the main results in Armstrong and Zhou (2011). We additionally introduce the matching effect and explain why it causes consumer welfare to be non-monotonic in transparency. Although conceptually different, our finding that higher transparency may lead to higher or lower prices is reminiscent of the ambiguous comparative statics in prices with respect to search costs in Moraga-González et al. (2017) and Moraga-González et al. (2020). Since we additionally allow for vertical differentiation and endogeneous prominence, we provide additional insights in terms of consumer welfare.

Dinerstein et al. (2018) consider the role of platform design in online markets and show that putting more weight on the price when displaying search results intensifies competition. When products are sufficiently differentiated, this comes at the expense of inefficient matching. This is similar to our finding that tougher competition for a spot in the consumers' consideration set may lead to lower prices relative to full transparency, but worse matching. A key difference is that Dinerstein et al. (2018) show that giving *different* information to consumers may be beneficial for them; whereas we show that giving consumers even *less* information may be beneficial for them. Dinerstein et al. (2018) consider online markets, whereas we show the relevance of market transparency initiatives in offline markets. Additionally, we establish a direct link between actually observed rankings and the consumers' consideration set.

2 Industry background and data

In this section, we explain the industry background as well as relevant market features. We show that there are persistent price differences as well as intertemporal price dispersion, motivating our model with private-information marginal cost shocks and both horizontal and vertical differentiation.

2.1 Market description and transparency regulation

In the German retail gasoline market, there are five major brands that operate across the entire country and are strongly vertically integrated. Following Bundeskartellamt (2011), these are referred to as oligopoly players (Aral (BP), Shell, Total, Esso, and Jet). The market leader Aral has a market share of around 16% across Germany (according to the number of stations in 2017).[3] The combined market share of the five oligopoly brands is 51%. In line with the definition used by Haucap et al. (2017), we additionally distinguish between 'integrated' brands (Orlen (Star), Agip, HEM and OMV; combined market share 12%). All other brands are referred to as 'others.'

Retail gasoline stations in Germany are obliged to transmit prices to a central database operated by the Market Transparency Unit (MTU, German Markttransparenzstelle für Kraftsstoffe MTS-K), a subunit of the German competition authority (Bundeskartellamt). This regulation was enacted in December 2013. Contrary to the Austrian model where the regulatory body also operates the website to inform consumers, in Germany the MTU only operates the database. Information to consumers is provided through privately operated websites and mobile apps, which are given access to the MTU's central database. There are 62 officially registered websites and mobile apps.[4] Gasoline stations in Germany are entirely unconstrained in their price setting; in particular, they may increase their prices arbitrarily often on any given day. A screenshot of one German price information provider (ADAC) is shown in Figure 1. All stations and prices close to the stated address are shown. On the Austrian price comparison website, only the cheapest stations are shown (see Figure 2 for a screenshot of the market around Vienna, Austria).

Our analysis is based on the entire history of prices set in 2017, provided by the

[3] https://www.mwv.de/statistiken/tankstellenbestand/

[4] https://www.bundeskartellamt.de/EN/Economicsectors/MineralOil/MTU-Fuels/mtufuels_node.html, accessed June 2018.

German website Tankerkönig,[5] which in turn uses prices published by the German MTU.

In addition to prices and rankings, we also observe the brand and various characteristics for each station. We complement the dataset with additional regional information on the population density from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR).[6] We proxy for input prices using the Brent Crude Oil Price published by the US Energy Information Administration[7] and calculate the price in Euro cents per liter based on the exchange rate published by the ECB (Janssen et al., 2011).[8] Website usage statistics are based on figures from the German Audit Bureau of Circulation (IVW).[9] Total gasoline and diesel consumption data were obtained from the Federal Office for Economic Affairs and Export Control.[10] Finally, we use building and road network information publicly available from Open Street Maps.

We delineate markets based on municipality boundaries because it allows for clean allocations and guarantees non-overlapping of markets, which is a key drawback of markets defined by circles around stations (Pennerstorfer et al., 2020). Considering geographically isolated markets (Nishida and Remer, 2018) also avoids overlapping markets, but is still somewhat arbitrary and may eliminate otherwise relevant markets. We focus on municipalities with at least three stations and at most 315 1x1km large grid cells since we believe that in these markets our key consideration of how market transparency shapes consumer choice plays the biggest role. In very small markets, information acquisition online is less relevant to start with. In larger cities there may be several additional forces at play that we may not be able to address in a satisfactory manner. We thus omit these very large municipalities altogether. Robustness checks are shown in Section 5.1. Since gasoline stations on the highway are considered as a separate market by the German Bundeskartellamt, we also omit them. In our sample there are 8,979 gasoline stations in 1,479 markets. Table 1 gives an overview of the station characteristics.

[5] http://tankerkoenig.de/

[6] http://www.inkar.de/

[7] https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm

[8] http://sdw.ecb.europa.eu/quickview.do?SERIES_KEY=120.EXR.D.USD.EUR.SP00.A

[9] http://ausweisung.ivw-online.de/

[10] http://www.bafa.de/DE/Energie/Rohstoffe/Mineraloelstatistik/mineraloel$_$node.html

2.2 Intertemporal and cross-sectional price dispersion

Two key features of the German retail gasoline market are worth mentioning. First, there are persistent price differences across stations and regions (see Table 1). Second, both prices and the price ranking of stations vary over time. If consumers knew which station was currently the cheapest, they might shop accordingly. In our sample, the average number of daily price adjustments per station is 4.9 on weekdays.

Intra-day price variation may result from systematic demand variation throughout the day, e.g., morning and evening commuting patterns. In order to eliminate these concerns, we restrict consideration to 5pm on weekdays, where most consumers fuel their cars (Montag and Winter, 2020). Figure 3 shows the distribution of the standard deviations of ranking across stations, which is very volatile. As an example, figures 4 and 5 show the price and the price ranking, respectively, of the Esso station in Plattling, Bavaria, over the first week of August 2017. The red vertical bars denote noon on each of these days. Prices and rankings are very volatile during the day and fluctuate much more than the oil price. One reason explaining these fluctuations may be deterministic traffic flows during the day. Figure 6 shows the ranking of the Esso station in Plattling, daily at 8am, which also fluctuates quite heavily. Since traffic flows should be the same at the same time during weekdays, they cannot explain the observed volatility.

The observed volatility in rankings is consistent both with firms playing mixed strategies as well as with station-specific shocks that induce price changes. Chandra and Tappata (2011) develop a test for mixed strategies and find clear support for it in their data. They investigate rankings at the station-pair level. Stations pairs are either "at the same corner" (which they define as being less than 270 feet away, i.e., 0.082km, apart from each other) or not. Since information frictions should not be present for stations at the same corner, differences in prices for those stations should be driven exclusively by systematic differences. If stations at the corner change the relative price rankings less frequently, this can be interpreted as a price setting that is consistent with mixed strategies.

We implement the test of Chandra and Tappata (2011) and apply it to our data. Figure 7 shows the CDF of rank reversals by distance as in Chandra and Tappata (2011, Figure 7; see appendix for details). Table 2 shows coefficients from regressing the dependent variable on an indicator for whether a station pair is on the same corner or not (Chandra and Tappata, 2011, Table 6). In contrast to Chandra and Tappata, we do not find evidence that station pairs on the same corner are systematically different from other station pairs. Fluctuations in rankings resulting from shocks to marginal costs appear to be a better explanation for our data.

2.3 Descriptive evidence

We are interested in the effect of transparency regulation on prices. Currently, Germany provides full transparency, whereas transparency in Austria is restricted. Time series of the weekly average prices for Germany and Austria including and excluding taxes are shown in Figure 8, respectively. These figures are based on the Weekly Oil Bulletin published by the European Commission.[11] Gross prices for diesel are much lower in Austria (mean 119.0 CPL) than in Germany (126.0 CPL). These differences are partially due to different tax rates. Excise duties on diesel are currently 39.7 CPL (plus 20% Value Added Tax) in Austria and 47.04 CPL (plus 19% VAT) in Germany. In both countries, VAT is levied on the price including the excise tax.[12] Even accounting for different tax rates, net prices in Austria are still lower (58.1 CPL vs. 58.8 CPL). We interpret these differences as evidence in favor of restricted transparency. We cannot, however, claim causality based on these numbers. Systematic price differences between countries might arise due to several other reasons, e.g., regulation of price adjustment patterns (Obradovits, 2014) or other institutional patterns. Therefore, we cannot compare prices between countries, but need to consider the effects in isolation within a particular country.

Thus, we next consider the main determinants of gasoline prices in Germany. We present results from reduced-form regressions with prices per station as dependent variable in Table 3. Naturally, the oil price (brent) is a key driver of price setting. Oligopoly and integrated brands charge substantially higher prices than other brands. Stations on the highway are even more expensive, lending support to the claim that they form a separate market Bundeskartellamt (2011). The competitive and spatial environment also play key roles. Prices decrease in the number of stations. A prominent location close to consumers (distance-weighted population) makes a gas station relatively more attractive and allows it to charge higher prices.

[11] https://ec.europa.eu/energy/en/data-analysis/weekly-oil-bulletin

[12] https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Service/Einfach_erklaert/2018-01-11-grundlagen-benzinpreis.html

These results indicate that there is some relationship between prices, consumer locations, and firm locations. In order to isolate the supply- and demand-side characteristics, we develop and estimate a structural model in the following sections.

2.4 Consumer locations

Consumer locations and the distance to stations are among the determinants governing consumer demand from a respective gasoline station. We do not observe the exact location, commuting patterns, etc., for each consumer at each point in time. We proxy for consumer locations by assigning them to 1x1km large cells, based on buildings information. Thomadsen (2005) uses census track data, which is readily available for small census blocks in the US. An advantage of our approach over methods considering only reported place of residence is that it takes into account also likely work locations. Spatially matching almost 32m buildings to almost 1m grid cells is a computationally heavy task that we accomplish using ArcGIS software and the ArcPy Python plugin.

Our starting point is population data per municipality obtained from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). For an illustration of our approach, consider the municipality Plattling, Bavaria, shown in Figure 9. We obtain buildings data freely available from Open Street Maps. The buildings in Plattling are shown in Figure 10. We assign the municipality's population to buildings proportional to the building ground area. Since we cannot identify either the building height or the building type, we assume that the building height within a municipality is uniform. We then project an artificial fishnet over the entire area, creating 1x1km large cells (see Figure 11). Finally, we calculate the sum of the proportional population per building in each of these cells. The representative consumer per cell, used for distance calculations, resides in the centroid of the cell.

3 Model

In light of the findings of the previous section, we present a model reflecting key features of retail gasoline markets. We assume that the mean utility per brand (chain) is the same across markets, allowing for identification even when market shares are unavailable (Thomadsen, 2005). We depart from Thomadsen (2005) by allowing the consideration set to be formed endogenously. In contrast to similar approaches by Butters (1977), Honka et al. (2017) or Sovinsky Goeree (2008), consideration sets do not respond to the explicit advertising expenditures of the firms but are formed only as functions of the prices chosen by firms, reflecting the actual mechanisms used by the price comparison websites.

Note that we present a static model, hence there is no time dimension and therefore also no time subscript for variables. It is implicitly understood that prices change over time, whereas general market characteristics are time-invariant. We also treat each station as an independent decision-maker since stations from the same brand very rarely interact in a given market.

More specifically, a geographic market M consists of $n \geq 1$ stations. Given awareness of station i with brand b, the indirect utility of consumer j from firm i is given by

$$V_{i,j} = X_i'\beta - d_{i,j}\delta - p_i\gamma + \eta_{i,j} \tag{1}$$

where X_i are a firm's brand, $d_{i,j}$ the driving distance, p_i firm i's price and $\eta_{i,j}$ a type 1 extreme value error term.

Consumers discretely choose one station in their consideration set S. Given a consideration set S consisting of a set of stations $s \in S$ and an outside option with utility normalized to 0, the choice probabilities of consumer type j take the usual logit form:

$$D_{s,j}(P, X, S|\beta, \delta, \gamma) = \frac{e^{X_s'\beta - d_{s,j}\delta - p_s\gamma}}{1 + \sum\limits_{k \in S} e^{X_k'\beta - d_{k,j}\delta - p_k\gamma}}. \tag{2}$$

There are two types of consumers: A fraction $1 - \mu$ are uninformed consumers and the remaining fraction μ are potentially informed online consumers. Due to data limitations, we do not explicitly model the choice of becoming informed and thus take these fractions as exogenously given.

For uninformed consumers, the consideration set consists of only one firm in market m. The probability of being the unique firm in the consideration set is assumed to be proportional to the distance to the firm.[13] Thus, demand for firm i from uninformed consumers is given by

$$D^U_{i,j}(P, X|\beta, \delta, \gamma) = \frac{D_{i,j}(P, X, \{i\}|\beta, \delta, \gamma)}{n-1} \left(1 - \frac{d_{i,j}}{\sum\limits_{j \in M} d_{k,j}}\right) \tag{3}$$

[13] In Section 5.1 we show that our main findings are robust to alternative specifications.

Online consumers are informed about the set S of k cheapest stations, where $k = \lceil \rho n \rceil$. $\rho \in [0, 1]$ is referred to as the "transparency regime" and is chosen by the regulator. In the baseline specification, $\rho = 1$ (the current German setting with full transparency). Given a vector of prices P and given that p_i is among the k cheapest prices, demand from online (web) consumers is:

$$D^W_{i,j}(P, S, X, m|\beta, \delta, \gamma) = D_{i,j}(P, X, S|\beta, \delta, \gamma) \tag{4}$$

and $D^W_{i,j} = 0$ if $rank(p_i, P) > k$.

Similar to the approach chosen by Thomadsen (2005), we discretize consumers' locations into a set of block cells B and compute distances, and consequently demand per block cell b, weighted by a population measure $\kappa(b)$. Taken together, given prices P, demand for firm i is given by

$$D_i(P, S, X, m|\beta, \delta, \gamma) = \sum_{b \in B} \kappa(b) \left(\mu D^W_{i,b}(\cdot) + (1 - \mu) D^U_{i,b}(\cdot) \right) \tag{5}$$

Upon realization of a firm-specific private-information cost-shock ε_i, firms simultaneously choose prices, taking the distribution of prices of their competitors as given. Marginal cost shocks can also be interpreted more broadly as profitability shocks, e.g., stemming from the sales of secondary products such as sandwiches, chocolate bars, and car wash facilities. Therefore, profits of firm i are given by

$$\pi_i = (r(p_i) - c_i)\mathbb{E}_P(D_i(\cdot)) \tag{6}$$

where $r(p_i)$ denotes the net revenue at price p_i, i.e., $r(p_i) = p_i/(1 + VAT) - excise$. Marginal costs c_i consist of a common part (across all firms, i.e., the oil price), a chain k-specific part and an additional firm-specific shock:

$$c_i = c_a + c_k + \varepsilon_i \tag{7}$$

where c_a may vary over time and ε_i is drawn from a known distribution $H(\varepsilon_i)$, distributed independently across firms.

Note that the expectation in (6) is taken w.r.t. to the prices of P, i.e., the vector of prices of all the competitors. Due to the linearity in (5), we can decompose expected demand into separate terms for uninformed and online consumers, respectively. Since uninformed consumers evaluate station i only against the outside option, the competitors' prices are irrelevant, and thus the expression can be taken out of the expectation altogether:

$$\mathbb{E}_P(D_i(\cdot)) = D^U(i, j)(\cdot) + \sum_{b \in B} \kappa(b) \mu \mathbb{E}_P(D^W_{i,b}(\cdot))$$

For n stations in the market, there are $n!$ possible rank configurations $\mathcal{R} \in \{1, 2, \ldots, n\}^n$ with the interpretation that the i-th element of $q \in \mathcal{R}$ denotes the ranking of station i (assuming no ties, which emerge with probability 0 for continuous price distributions). The expected demand from online consumers is positive only whenever station i is among the cheapest k stations, and 0 otherwise. Station i only cares about rank orderings $\mathcal{R}_i^k$ where i is among the top k, i.e., $\mathcal{R}_i^k = \{q \in \mathcal{R} | rank(\mathcal{R}, i) \leq k\}$. We write the expected demand from online consumers as follows:

$$\mathbb{E}_P(D_{i,b}^W(\cdot)) = \sum_{q \in \mathcal{R}_i^k} Pr(q|p_i)\mathbb{E}_{P|q}(D_{i,b}^W(\cdot))$$

where, denoting the equilibrium price distribution as $G_j(p_j)$, we have that

$$\mathbb{E}_{P|q}(D_{i,b}^W(\cdot)) = \int_{-\infty}^{p_i} \ldots \int_{p_i}^{\infty} D_{i,b}^W(\cdot) dG_k(p_k) \ldots dG_1(p_1)$$

depending on the identity of stations in ranking q.

Taking first-order conditions in (6), we see that in equilibrium it has to hold that for each firm i

$$\frac{\mathbb{E}_P(D_i(\cdot))}{1 + VAT} + (r(p_i) - c_a - c_k - \varepsilon_i)\frac{\partial \mathbb{E}_P(D_i(\cdot))}{\partial p_i} = 0 \tag{8}$$

which we can rewrite in vector notation as

$$\tilde{D}(P) + \Omega(r(P) - C - \varepsilon) = 0 \tag{9}$$

where $\tilde{D}(P) = D(P)/(1 + VAT)$ and the diagonal matrix Ω consists of $\frac{\partial \mathbb{E}_P(D_i(\cdot))}{\partial p_i}$ on the diagonal and zeros everywhere else.

Conditional on the structural parameters $\theta = (\beta', \delta, \gamma, c')$ and observables X, this can be rewritten as

$$\tilde{D}(P, X|\theta) + \Omega(P, X|\theta)(r(P) - C - \varepsilon) = 0 \tag{10}$$

or alternatively be solved for ε which enters as a residual into our GMM estimation:

$$\varepsilon = r(P) - C + \Omega(P, X|\theta)^{-1}\tilde{D}(P, X|\theta). \tag{11}$$

3.1 Consumer welfare

We measure the change in consumer welfare by compensating variation (Nevo, 2000a, Brenkers and Verboven, 2006). As pointed out by Nevo (2000a), this is a good measure of consumer welfare under the assumption that in the counterfactual scenario, both unobserved components and the utility form remain unchanged. Both of these assumptions are justified in our environment.

Small and Rosen (1981) derive the compensating variation for deterministic prices when each consumer i's demand is given by logit demand. We extend their calculation to account for stochastic prices and consideration sets S and obtain

$$CV_i = -\frac{1}{\gamma}\left(\int_P \log\left(\sum_{j\in S^{post}(P)} V_{i,j}^{post}\right) dG^{post}(P) - \int_P \log\left(\sum_{j\in S^{pre}(P)} V_{i,j}^{pre}\right) dG^{pre}(P)\right). \tag{12}$$

Aggregating across the entire population and accounting for heterogeneous consideration sets of uninformed and online consumers, we obtain the mean compensating variation:

$$MCV = \frac{\sum_{b\in B}\kappa(b)(\mu CV_b^W + (1-\mu)CV_b^U)}{\sum_{b\in B}\kappa(b)}. \tag{13}$$

4 Estimation

4.1 Identification

We are interested in the supply- and demand-side parameters (which in turn imply estimated elasticities) that generated the data observed. A major challenge for identification is that we do not observe market shares, rendering standard estimation routines infeasible. Our approach is based on the work of Thomadsen (2005) instead. Separate identification of the supply- and demand-side parameters is possible since we observe different market configurations and make the assumption that brand mean utilities and costs are the same across markets. This is justified in our environment since stations of the same brand have the same design and typically also the same set of amenities. Additionally, we assume that demand is entirely deterministic given prices, which is a special case of the approach in MacKay and Miller (2019). In the absence of demand shocks, we can estimate from the first-order conditions in the supply relation.

To illustrate, consider that we observe prices from markets differing in their number of firms n (see appendix for more details). For logit demand, equilibrium prices are non-linear functions of n. Since there are only marginal cost shocks and no demand shocks by assumption, we can invert the first-order conditions for optimal pricing, use these shocks as GMM residuals, and jointly identify supply and demand using higher order moments of n.

Alternatively, supply and demand parameters would also be simultaneously identified if we only observed a monopolist firm's prices and varying input prices over time. Provided the equilibrium prices are non-linear in input prices, we could use higher order moments of input prices. This aspect is absent in Thomadsen (2005) who does not observe firms over time but gives additional identifying power in our setting. We combine both sources of variation and moreover employ macro moments to enhance efficiency.

Two parameters are not identified within our model and thus need to be specified upfront. As is common in the literature, we need to define the total market size (Nevo, 2000b). We assume that the maximum potential monthly gasoline demand is given by the highest observed value, i.e., from March 2017 (see Table 14). This yields a total market share of all inside goods of 92.9%. Methodologically similar approaches are used by Manuszak (2010) and Houde (2012).

Moreover, our data does not allow us to identify the search cost distribution (MacMinn, 1980, Myśliwski et al., 2017). Hence, we cannot endogenize the consumers' search decision and take the fractions of uninformed and online consumers as given. We base our calculation on the actual usage data of some of the most popular price information services in Germany (see appendix for details). Our proxy for informed consumers is given by $\mu = 20.5\%$. The proxy compares very well to survey data from 2016 conducted by the German Federal Ministry of Economics and Technology, according to which "around one quarter of the survey participants use [price] information services always or occasionally" (Bundesministerium für Wirtschaft und Energie, 2018, p. 11).

4.2 Estimation technique

We apply a two-step estimation technique to estimate the structural parameters θ. Given θ and consistent estimates for equilibrium beliefs about competitor prices, we can evaluate the right-hand side of (11) and hence obtain an expression for ε. Our instruments Z need to satisfy $\mathbb{E}(\varepsilon_j|Z_j) = 0$. We use brand dummies as supply shifters, which, by definition, are independent of the unobserved component of marginal costs ε (Thomadsen, 2005). Our demand shifters consist of other market characteristics, such as number of stations, consumer demographics, and interactions with the brand dummies. Finally, we also use the oil price in interactions with the number of stations.

Given instruments Z, we consistently estimate θ via a Method of Simulated Moments (MSM) approach (McFadden, 1989, Gourieroux and Monfort, 1996, Train, 2009). Simulation is necessary since the multi-dimensional integral over competitor prices does not admit a closed-form solution.

More specifically, our first estimation step for parametric estimation of the joint competitor price distribution $\hat{G}(p)$ works as follows. Observed prices are drawn from some distribution $G_i(p|c_a)$ for each firm i, where c_a is the current oil price. In principle, $\hat{G}(p)$ is identified non-parametrically. Due to data availability, we choose a parametric approach. We experimented with different distributions and found the best fit using a normal distribution. So we assume that for each firm, prices are drawn from a normal distribution, i.e., $p_i|c_a \sim N(\mu_i(c_a), \sigma_i(c_a))$ where $\mu_i = b_{i,0} + b_{i,1}c_a$ and $\sigma_i = b_{i,2} + b_{i,3}c_a$. We found that the variance does not change with input prices, so we set $b_{i,3} = 0$.

Given our incomplete-information setting, price observations are independent from each other (conditional on observables). Thus, for each firm i, we can estimate the parameter vector $b_i = (b_{i,0}, b_{i,1}, b_{i,2})$ independently by maximizing the log likelihood function

$$LL(b_i) = \sum_t \log \phi(p_{i,t}, c_t; b_i)$$

where ϕ denotes the normal density.

In the second step, we plug $\hat{G}(p)$ into the first-order conditions (8) and solve for $\varepsilon(\theta)$. We have that $\mathbb{E}(\varepsilon(\theta)) = 0$. Since $\hat{G}(p)$ is a consistent estimate for $G(p|\theta)$, ε can then be used to form sample moments for MSM estimation in the usual manner.

This implies that in equilibrium it needs to hold that

$$\varepsilon = r(P) - C + \Omega(P, X|\theta)^{-1}\tilde{D}(P, X|\theta)$$

where Ω and $\tilde{D}$ are functions of $\mathbb{E}_P$. Given consistent estimates of $\hat{G}(p)$, we can write this as

$$\hat{\varepsilon}(\theta) = r(P) - C + h(P, X, \hat{G}(p)|\theta).$$

We then formulate sample moments

$$M(\theta) = \frac{1}{N}\sum_{i=1}^{N} z'\hat{\varepsilon}_i(\theta)$$

and find θ that minimizes the following expression:

$$\arg\min_{\theta} M(\theta)' W M(\theta)$$

for W a positive semi-definite matrix.

We add macro-moments regarding total quantity to aid estimation efficiency (Imbens and Lancaster, 1994). Although our two-step approach does not require solving integral equations in every step, the computational burden is still substantial since there is logit demand from each customer block cell.

Similar to Ryan (2012), we derive standard errors by repeatedly taking subsamples of random time periods, all of them including our entire set of stations (Politis and Romano, 1994).

As described in Section 2, we distinguish between five oligopoly brands, integrated brands, and other brands. We thus estimate seven mean brand utilities and seven mean marginal costs parameters.

4.3 Results

The main estimation results are shown in Table 4. Indifference with an Aral station charging a price of 120 Euro cents and at the same distance is shown in Table 5. In line with intuition, the mean brand utilities for oligopoly brands are higher than those of integrated or other brands. The implied travel costs are 9.89 Euro cents per kilometer to fill a 50 liter tank. At an average consumption of 6 liters per 100 kilometer and an average price of 113 Euro cents per liter, gasoline costs are 6.78 Euro cents per kilometer. Our estimates imply that consumption accounts for roughly two thirds of estimated travel costs.

We present own- and cross-price elasticities for the market in Plattling, consisting of four stations (2 other brands (Billmeier and Globus), one Total and one Esso; see Table 6 and Figure 9 for the spatial distribution of the market). The air distance from the Total station is 2.3km, 2.0km, and 1.6km, to the Billmaier, Globus, and Esso station, respectively. Consequently, the demand elasticity of Total (3rd row) is stronger w.r.t. Globus than to Billmeier (columns 2 and 1), both of which are unbranded. We interpret these estimates as supportive evidence of the credibility of our approach.

4.4 Model validation

In order to show that our model adequately captures the underlying market, we first compare the actual observed prices with the equilibrium prices implied by our parameter estimates and the model. We compute equilibrium prices by storing the residuals $\varepsilon_{i,t}$ in the estimation routine and compute the equilibrium pricing function $p_i(\varepsilon_{i,t})$. We repeat regressing the price on a set of station and market covariates (see Table 7). The coefficients from the equilibrium model (column 2) are very similar, both in sign and in magnitude, to the original coefficients (column 1).

We also validate our model out-of-sample by predicting prices in periods not used during estimation (see Table 8). Our approach predicts both in-sample and out-of-sample prices very well, providing confidence in the validity of our model. We are thus confident that our model does indeed capture the most relevant market characteristics.

5 Counterfactual policy simulation

Given the parameter estimates obtained in the previous section, we conduct the following counterfactual policy simulations: What is the effect on prices and consumer welfare as we change the transparency regime ρ?

Estimates from the previous section were obtained by setting $\rho = 1$, i.e., full transparency. Using the estimated parameter values, we compute equilibrium prices under the current transparency regime as well as under alternative transparency regimes. For welfare calculations, we assume that the average car drives 14,000km per year, consuming 6.4 liters per 100 km.[14]

The underlying game is an incomplete information game, where we are interested in the equilibrium pricing function $p_i(\varepsilon_i; \theta)$ for each firm i and each marginal cost shock ε_i. This game does not admit a closed-form solution since the first-order conditions involve integral equations with respect to the equilibrium prices by competitors. We approximate the solutions by iteratively solving the system of integral equations (Richardson, 2004) at a finite number of points. In each step, we draw prices from the competitor's price distribution to calculate a given firm's expected profit. We then fit a fifth-degree polynomial over the firm's profit function in order

[14]https://www.kba.de/DE/Statistik/Kraftverkehr/VerkehrKilometer/verkehr_in_kilometern_node.html

to smooth out discrete sampling error. The pricing function $p_i(\varepsilon_i)$ is obtained by maximizing this smooth profit function and is then used for the next iteration step.

The estimated parameters result in a distribution of residuals in (11). Based on the residuals we fit a distribution of marginal cost shocks. The best fit is given by a normal distribution with mean 0, truncated symmetrically at -20 and 20, respectively, and a variance of 44.67. Like Ryan (2012), we construct confidence intervals based on a subsample bootstrap as suggested by Politis and Romano (1994).

The results from the counterfactual simulation of restricting transparency to showing only the cheapest 50% of prices are shown in Table 9. The mean price decreases by 0.5% and the quantity-weighted price decreases by 0.6%. Less efficient matching and further driving distances reduce the positive effect on consumer welfare, leading to a net effect of consumer welfare increasing by 2.6% of firms' current margins. The increase in consumer surplus is exclusively through uninformed consumers (79.5% of consumers) who benefit from lower prices, while leaving their matching quality unaffected. Conversely, informed consumers (20.5%) are worse off under restricted transparency due to inefficient matching.

Next, we decompose the total effect into a comparison and an attention effect (see Table 10). The comparison effect is computed by showing consumers only 50% of the prices, but crucially, a random sample instead of the cheapest ones. This counterfactual isolates the negative aspects of the policy change, namely making it harder for consumers to compare offers. The estimated comparison effect is a price increase of 1.1%. The attention effect, defined as the effect of firms competing for consumers' attention when only the cheapest prices are shown, is computed as the difference between the total effect and the comparison effect. Mean prices decrease by 1.6%. These considerations reveal an important difference between *how much* information versus exactly *which* information is given to the consumers.

The underlying mechanism is illustrated by the equilibrium pricing functions of two stations in Plattling, shown in Figure 12. As transparency decreases (the dotted lines in the figure), firms compete more aggressively for attention when they receive a favorable cost draw. However, when marginal costs are high, firms do not compete for consumers with elastic demand anymore and focus on the loyal consumers segment instead, charging relatively high prices. This leads to an increase in price dispersion and a decrease in average prices.

We conduct additional counterfactual simulations where we allow for different fractions of prices shown (see Figure 13). As we reduce the transparency regime ρ relative to the baseline scenario with $\rho = 1$, both average and quantity-weighted prices decrease up to $\rho = 0.1$ and price dispersion increases. At $\rho = 0$, consumers using the website cannot obtain any information and hence firms price exclusively for uninformed consumers, resulting in very high prices. However, besides the effect on prices, different transparent regimes also affect match quality. Thus, consumer welfare is not maximized at $\rho = 0.1$, even though prices are very low. Indeed, the optimum from a consumers' point of view is attained when transparency is intermediate, i.e., $\rho = 0.2$. In that case, prices are already substantially lower than under full transparency and match quality is still sufficiently high.

Finally, we simulate different fractions of informed consumers (see Figure 14). In the baseline scenario, $\mu = 0.205$. Prices decrease monotonically in μ, whereas consumer welfare monotonically increases. In order to reach the same increase in consumer welfare as achieved by a reduction in transparency to $\rho = 0.1$, a fraction of informed consumer of $\mu' \approx 0.25$ would be needed. This would require educating consumers toward using price comparison websites more frequently. Conversely, reducing transparency to facilitate competition for attention merely requires small changes to the websites.

5.1 Robustness checks

In this section, we describe the following four sets of robustness checks, namely accounting for (i) sample selection, (ii) fraction of informed consumers, (iii) joint market share of the inside goods, and (iv) assignment of uninformed consumers.

We first present a set of robustness checks concerning our sample selection. Our main specification is estimated for municipalities with at least three stations and at most 315 block cells. Table 11 shows our descriptive regressions for different sample selection criteria. All coefficients of interest are very robust across specifications.

Our main estimation and counterfactual policy simulation are conducted under the assumption that the fraction of informed consumers $\mu = 0.205$, and that the combined market share of all inside goods is 0.929. We estimate the model (see Table 12) and compute the counterfactuals (see Table 13) also for alternative assumptions about the fraction of informed consumers and the market share of inside goods. Of course, the parameter estimates change when we make alternative assumptions, but our main findings are very robust to these alternative specifications. Across specifications, decreasing transparency from a baseline scenario with full transparency to restricted transparency with $\rho = 0.5$ leads to lower prices and higher consumer welfare.

We also investigate an alternative specification in which uninformed consumers do not purchase from the station that is the closest, but from the station that had the cheapest average price in the past. As expected, parameters estimates are sensitive to this making this adjustment. But again, the main result remains very robust: Decreasing transparency makes consumers better off.

6 Conclusion

In this paper, we develop and estimate a structural model of the German retail gasoline market. Prices may be lower in less transparent settings. The main reason is that when firms know that consumers will be informed about the cheapest offers only, this opens the additional channel of competing for attention. However, matching is worse in restricted transparency regimes, making the net effect on consumer welfare ambiguous even if prices decrease.

According to our counterfactual policy simulations, we find that prices decrease and consumer welfare would increase substantially if Germany were to adopt a more restricted transparency regime. We also show that transparency should not be too restricted, because eventually the matching effect dominates and hence consumers are worse off. Our decomposition of the comparison and the attention effect reveals an important difference between *how much* information versus exactly *which* information is given to consumers.

Besides regulating transparency, our analysis also makes a strong case for incentivizing consumers to obtain information online. While spending time to inquire about prices may be costly for the individual, the induced price pressure on firms is even stronger than the effect due to transparency restrictions.

The implications of our findings are much more general than a sole applicability to the German retail gasoline market. In previous studies, as well as in most policy debates, contributing to allowing consumers to make better informed choices was seen as the main objective. As this paper shows, however, this is only one part of the story. What policymakers should also keep in mind is that the appropriate transparency regime must provide strong incentives for firms to compete for attention and still allow consumers to find good matches.

We develop a framework that allows estimating games of incomplete information even when richer demand-side data are not available. The modelling and estimation technique may serve as a building block for future research. Frequently, prices are very easy to obtain whereas demand data are not. Using publicly available geo-spatial data can be useful in numerous other settings.

The intuition developed in this paper provides a rough guidance for optimal transparency regulation. Conducting other country- and industry-specific studies appears to be a fruitful area for future research in order to improve our understanding of how exactly the optimal transparency regime changes as a function of market characteristics. Richer demand side data would allow endogenizing a consumer's decision to search and estimating the search cost distribution.

Appendix

A Tables

Table 1: Summary statistics for gasoline stations

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
# stations	15.649	34.164	3	221	8979
Mean price	112.907	3.286	98.900	150.9	8979
Pop. dens.	709.041	650.642	24.624	4111.472	8979
Pop.	98100.511	283859.636	592	1830584	8979
Muni. area.	102.112	117.253	2.89	755.09	8979
Aral	0.16	0.367	0	1	8979
Shell	0.119	0.324	0	1	8979
Total	0.054	0.225	0	1	8979
Esso	0.069	0.254	0	1	8979
Jet	0.053	0.225	0	1	8979
Integrated	0.105	0.307	0	1	8979
Other brand	0.439	0.496	0	1	8979

Notes: This table presents summary statistics for key variables per station in our sample. There is substantial variation in the mean price across stations.

Table 2: Regressions rank reversals and price dispersion

Sample	Dep. Var.	OLS
Station pairs	r_{ij}	-0.0006
within 1 mile (1.61km)		(0.0043)
$N = 10,814$	σ_{ij}	0.0766
		(0.5345)
Station pairs	r_{ij}	-0.006
within 2 miles (3.22km)		(0.0045)
$N = 27,457$	σ_{ij}	0.2150
		(0.4706)

Notes: This table shows coefficients from regressing the dependent variable on an indicator for whether a station pair is on the same corner or not (Chandra and Tappata, 2011, Table 6). r_{ij} denotes the rank reversal between station i and j, and σ_{ij} the standard deviation of price spreads. In contrast to Chandra and Tappata (2011), where all these coefficients are negative and statistically significant, we do not find evidence that station pairs on the same corner are systematically different from other station pairs.
Standard errors in parentheses.

Table 3: Descriptive regressions for determinants of prices

	(1)	(2)	(3)
	Price	Price	Price
Brent (CPL)	1.01***	1.01***	1.01***
	(0.00)	(0.00)	(0.00)
# stations	-0.03***	-0.03***	
	(0.00)	(0.00)	
# olig. stat.	0.00***	0.00***	
	(0.00)	(0.00)	
# integr. stat.	0.00***	0.00***	
	(0.00)	(0.00)	
Oligopoly brand	1.56***	1.42***	1.53***
	(0.01)	(0.01)	(0.00)
Integrated brand	0.42***	0.32***	0.39***
	(0.01)	(0.01)	(0.01)
Pop. wgt. by dis.	0.00***	0.00***	0.33***
	(0.00)	(0.00)	(0.01)
Muncipality. area	4.59***	4.38***	
	(0.05)	(0.05)	
Population	-4.07***	-3.63***	
	(0.14)	(0.13)	
On highway	5.68***		
	(0.02)		
Constant	81.29***	81.48***	80.97***
	(0.03)	(0.03)	(0.03)
Observations	2220365	2160476	2160476
Market FE	No	No	Yes
Observations	2220365	2160476	2160476

Standard errors in parentheses

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

Notes: This table presents panel regression results for prices at 8am on weekdays. Column (1) shows that highway stations are more expensive, justifying the claim that they operate in a separate market. Thus, highway stations are excluded in columns (2)-(3). Oligopoly and branded stations are more expensive on average. The competitive environment plays a relevant role: The coefficient on # stations is negative, and the coefficient on population weighted by distance is positive. We do not give a causal interpretation to these figures (see main text for details). Standard errors in parentheses.

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

Table 4: Estimation results of supply and demand parameters

Variable		S.E.
Price sensitivity (γ; EUR per liter)	25.27	0.03
Distance disutility (δ; EUR per km)	0.05	0.05
Implied travel costs for 50 liters (EUR cent)	9.89	
V_{Aral}	31.81	3.18
V_{Shell}	31.79	3.15
V_{Total}	31.63	3.08
V_{Esso}	31.55	3.09
V_{JET}	31.23	3.09
$V_{integrated}$	31.37	3.13
V_{other}	31.31	3.05
c_{Aral}	3.73	1.28
c_{Shell}	4.27	1.44
c_{Total}	2.80	1.58
c_{Esso}	2.86	1.39
c_{JET}	1.51	1.33
$c_{integrated}$	2.53	1.22
c_{other}	1.88	1.14
Median own price elasticity	-24.47	
Median outside diversion	7.33%	

Notes: This table presents the estimated supply- and demand-side parameters for seven brands, including standard errors (see description in the main text). V_x and c_x denote brand x's mean utility and marginal cost (in CPL on top of the oil price), respectively.

Table 5: Indifference prices across brands

Brand	Indifference Price
Shell	119.92
Total	119.29
Esso	119.00
JET	117.72
Integrated	118.27
Other	118.04

Notes: This table shows the price at which a consumer is indifferent between an Aral station charging a price of 120 and a station of another brand, given both stations are equal distances away, as implied by the estimates in Table 4.

Table 6: Own- and cross-price elasticities for the market in Plattling

	Other 1	Other 2	Total	Esso
Other 1 (Billmeier)	-19.08	7.72	8.19	7.65
Other 2 (Globus)	7.86	-19.11	8.12	7.58
Total	8.00	7.79	-19.00	7.73
Esso	8.01	7.79	8.27	-19.54

Notes: This table presents own- and cross-price elasticities for the market in Plattling. The cell i, j shows the percentage change in demand of station i resulting from a 1% price increase of station j.

Table 7: Determinants of observed and computed prices

Dependent variable	Observed price	Computed price
	(1)	(2)
Brent (CPL)	0.95***	0.78***
	(0.00)	(0.00)
# stations	-0.01***	-0.06***
	(0.00)	(0.00)
Oligopoly brand	1.22***	1.31***
	(0.01)	(0.02)
Integrated brand	0.12***	0.45***
	(0.02)	(0.02)
Pop. wgt. by dis.	0.00***	0.00
	(0.00)	(0.00)
Municipality area	4.45***	4.80***
	(0.13)	(0.14)
Population	-3.14***	9.28***
	(0.36)	(0.38)
Constant	83.27***	88.79***
	(0.07)	(0.08)
Observations	195382	195382

Notes: This tables presents results of regressing observed prices (column 1) and computed equilibrium prices (column 2) on market and station covariates. The main coefficients from the equilibrium model are very similar, both in sign and in magnitude, to the original coefficients. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Determinants of observed and computed prices, out-of-sample

Dependent variable	Observed price	Computed price
	(1)	(2)
Brent (CPL)	1.07***	0.77***
	(0.00)	(0.00)
# stations	-0.02***	-0.02***
	(0.00)	(0.00)
Oligopoly brand	1.22***	1.34***
	(0.01)	(0.01)
Integrated brand	0.09***	0.48***
	(0.02)	(0.02)
Pop. wgt. by dis.	0.00***	-0.00***
	(0.00)	(0.00)
Municipality area	4.38***	4.06***
	(0.13)	(0.12)
Population	-3.26***	7.69***
	(0.34)	(0.34)
Constant	80.05***	89.02***
	(0.08)	(0.08)
Observations	228848	228848

Notes: These tables presents results of regressing observed prices (out-of-sample, column 1) and computed equilibrium prices (column 2) on market and station covariates. The subsample consists of a random sample of 10% of periods not used during estimation. Our approach predicts both in-sample and out-of-sample prices very well, providing confidence in the validity of our model. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Counterfactual policy simulation - Main results

Variable	Baseline $\rho = 1$	Counterfactual $\rho = 0.5$		
			Δ	Δ 95 % C.I.
Mean price (CPL)	113.74	113.21	-0.48 %	[-0.63, -0.40]
Quantity-weighted price (CPL)	111.82	111.13	-0.62 %	[-0.74, -0.60]
Welfare uninformed cons.			3.42	[3.23, 4.32]
Welfare informed cons.			-0.47	[-0.56, -0.13]
Mean consumer welfare			2.60	[2.42, 3.40]
Mean price variance	19.25	22.09	14.79 %	[13.89, 16.12]
Mean driving distance (km)	3.41	3.41	0.16 %	[0.13, 0.20]
Profits			-1.93 %	[-2.55, -1.83]
Margins (CPL)	14.93	14.56	-2.42 %	[-2.90, -2.30]
Market share Q olig. (%)	48.02	47.63	-0.63 %	[-0.78, -0.58]
Market share Q integr. (%)	10.70	10.75	0.47 %	[0.32, 0.51]
Market share Q other. (%)	41.28	41.62	0.82 %	[0.59, 0.83]
Total demand			0.47 %	[0.41, 0.57]
Total welfare (EUR)			3.36	[2.12, 3.55]

Notes: This table presents changes in market outcomes as a result of a transparency reduction from $\rho = 1$ to $\rho = 0.5$. Consumer welfare is measured in compensating variation and reported relative to baseline firm margins. Total welfare and profit calculations are computed based on annual expenditures per car. Prices are lower under restricted transparency, and consumer welfare is higher. Although profits decrease, the positive effect on consumer welfare dominates, and hence total welfare is also higher under restricted transparency. The 95% confidence interval is obtained by a subsample bootstrap (see main text for details).

Table 10: Counterfactual policy simulation - Comparison effect vs. attention effect

	$\Delta\rho = 0.5$	Δ Comparison eff.	Δ Attention eff.
Mean price (CPL)	-0.48 %	1.09 %	-1.57 %
Quantity-weighted price (CPL)	-0.62 %	1.66 %	-2.28 %
Welfare uninformed cons.	3.42	-8.78	12.2
Welfare informed cons.	-0.47	-19.81	19.34
Mean consumer welfare	2.60	- 11.04	13.64
Mean price variance	14.79 %	-31.71 %	46.5 %
Mean driving distance	0.16 %	-0.10 %	0.26 %
Profits	-1.93 %	1.22 %	-3.15 %
Margins (CPL)	- 2.42 %	6.49 %	-8.91 %
Market share Q olig. (%)	-0.63 %	0.22 %	-0.85%
Market share Q integr. (%)	0.47 %	0.57 %	-0.10 %
Market share Q other. (%)	0.82 %	-0.41 %	1.23 %
Total demand	0.47 %	-7.04 %	7.51
Total welfare (EUR)	3.36	-14.69	18.05

Notes: This tables presents changes in market outcomes as a result of a transparency reduction from $\rho = 1$ to $\rho = 0.5$, decomposed into comparison and attention effect. The comparison effect is computed through a counterfactual scenario in which online consumers only learn 50% of the prices, but these prices are random instead of ranked. The attention effect is the difference between the total counterfactual effect of restricted transparency $\rho = 0.5$ and the comparison effect. The attention effect dominates the comparison effect, leading to lower prices and higher consumer welfare under restricted transparency.

Table 11: Descriptive regressions - robustness checks

	(1)	(2)	(3)	(4)
	Price	Price	Price	Price
Brent (CPL)	1.01***	1.01***	1.01***	1.02***
	(0.00)	(0.00)	(0.00)	(0.00)
# stations	-0.03***	-0.02***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
# olig. stat.	0.00***	0.00***	-0.00	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
# integr. stat.	0.00***	0.00***	0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Oligopoly brand	1.42***	1.49***	1.56***	1.58***
	(0.01)	(0.01)	(0.01)	(0.01)
Integrated brand	0.32***	0.18***	0.25***	0.42***
	(0.01)	(0.01)	(0.01)	(0.02)
Pop. wgt. by dis.	0.00***	0.00***	0.01***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Municipality area	4.38***	5.25***	4.41***	8.46***
	(0.05)	(0.06)	(0.06)	(0.13)
Population	-3.63***	-4.47***	-8.30***	-5.55***
	(0.13)	(0.14)	(0.16)	(0.30)
Constant	81.48***	81.17***	81.52***	80.98***
	(0.03)	(0.03)	(0.03)	(0.05)
# stations	≥ 3	≥ 5	$5 \leq n < 50$	≥ 5
# block cells	≤ 315	≤ 315	≤ 315	< 100
Observations	2160476	1551382	1478257	583777

Standard errors in parentheses

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

Notes: This tables presents robustness checks for our descriptive regressions with respect to different sample selection criteria. All coefficients are very robust across specifications.

Table 12: Robustness checks for estimation results of supply- and demand parameters

	(1)	(2)	(3)	(4)	(5)	(6)
Informed consumers μ	0.205	0.15	0.25	0.205	0.205	0.205
Market Share Inside Goods	0.929	0.929	0.9	0.957	0.929	0.929
Uninformed assignment	Dist.	Dist.	Dist.	Dist.	Dist.	Price
Price sensitivity (γ)	25.27	25.38	25.47	25.74	24.97	25.44
Distance disutility per km (δ)	0.05	0.03	0.03	0.05	0.01	0.03
Implied travel costs for 50 liters (EUR cent)	9.89	5.50	6.78	9.95	1.64	6.70
V_{Aral}	31.81	31.71	31.78	31.76	31.75	31.78
V_{Shell}	31.79	31.74	31.77	31.73	31.86	31.77
V_{Total}	31.63	31.51	31.54	31.52	31.56	31.54
V_{Esso}	31.55	31.64	31.56	31.54	31.60	31.56
V_{JET}	31.23	31.43	31.21	31.21	31.44	31.22
$V_{integrated}$	31.37	31.32	31.34	31.30	31.38	31.35
V_{other}	31.31	31.27	31.30	31.27	31.29	31.31
c_{Aral}	3.73	0.28	5.82	5.73	0.81	3.92
c_{Shell}	4.27	0.54	5.95	5.88	1.48	4.05
c_{Total}	2.80	-0.22	5.14	5.20	0.28	3.21
c_{Esso}	2.86	-0.71	4.97	4.75	0.64	2.98
c_{JET}	1.51	-1.84	3.61	3.31	-0.04	1.52
$c_{integrated}$	2.53	-0.84	4.57	4.48	-0.13	2.65
c_{other}	1.88	-1.50	4.28	3.94	-0.27	2.23

Notes: This table presents robustness checks for the estimation results of supply- and demand parameters. The main specification is in column (1) (see Table 4). Specifications (2) and (3) are estimated with a modified fraction of informed consumers μ. In specification (4) and (5), the market share of inside goods is changed. In specification (6) uninformed consumers are assigned to stations based on average prices instead of distance.

Table 13: Robustness checks for counterfactual policy simulation

	(1)	(2)	(3)	(4)	(5)	(6)
Informed consumers μ	0.205	0.15	0.25	0.205	0.205	0.205
Market Share Inside Goods	0.929	0.929	0.929	0.9	0.957	0.929
Uninformed assignment	Dist.	Dist.	Dist.	Dist.	Dist.	Price
Mean price (CPL)	-0.48 %	-0.63 %	-0.40 %	-0.33 %	-0.78 %	-0.50 %
Quantity-weighted price (CPL)	-0.62 %	-0.72 %	-0.63 %	-0.53 %	-0.83 %	-0.66 %
Welfare uninformed cons.	3.42	3.87	3.33	2.73	4.83	3.79
Welfare informed cons.	-0.47	-0.64	-0.19	-0.82	-0.14	-0.27
Mean consumer welfare	2.60	3.18	2.43	1.97	3.86	2.94
Mean price variance	14.79 %	13.60 %	16.27 %	13.74 %	11.71 %	15.49 %
Mean driving distance	0.16 %	0.16 %	0.16 %	0.18 %	0.13 %	0.07 %
Profits	-1.93 %	-2.19 %	-2.08 %	-1.58 %	-2.93 %	-2.20 %
Market share Q olig. (%)	-0.63 %	-0.77 %	-0.69 %	-0.75 %	-0.54 %	-0.67 %
Market share Q integr. (%)	-0.47 %	0.42 %	0.47 %	0.44 %	0.38 %	0.46 %
Market share Q other. (%)	0.82 %	0.80 %	0.69 %	0.77 %	0.55 %	0.66 %
Total demand	0.47 %	0.61 %	0.30 %	0.38 %	0.55 %	0.50 %
Total welfare (EUR)	3.36	3.63	1.93	1.62	4.11	2.73

Notes: This table presents changes in market outcomes as a result of a transparency reduction from $\rho = 1$ to $\rho = 0.5$. The main specification is in column (1) (see Table 9). Specifications (2) and (3) are estimated with a modified fraction of informed consumers μ. In specification (4) and (5), the market share of inside goods is changed. In specification (6) uninformed consumers are assigned to stations based on average prices instead of distance. Naturally, the exact numbers vary depending on the assumptions, but there is a clear and consistent picture across specifications: Reducing transparency leads to lower prices and higher consumer welfare.

Table 14: Total diesel consumption in Germany in 2017

Month	Consumption (tons)
Jan	2,954,270
Feb	2,801,861
Mar	3,465,649
Apr	3,146,429
May	3,346,302
Jun	3,316,595
Jul	3,299,811
Aug	3,408,483
Sep	3,264,392
Oct	3,306,754
Nov	3,351,100
Dec	3,040,903
Total	38,702,549

Notes: This table shows total diesel consumption in Germany in 2017 in tons per month. 38.7 million tons equal 46.1 billion liters. Source: Federal Office for Economic Affairs and Export Control

B Figures

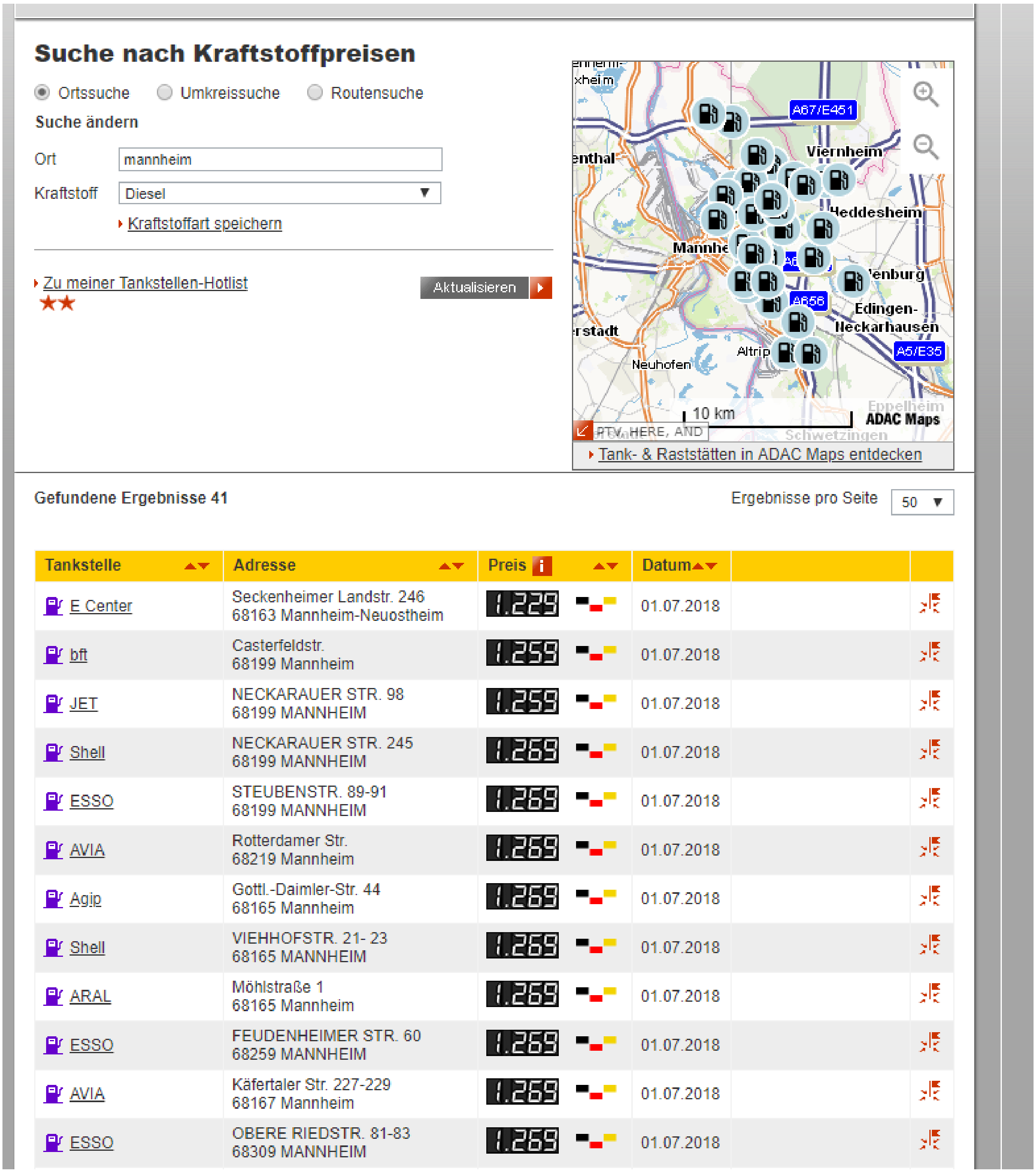


Figure 1: German fuel price comparison website

Notes: This figure shows a screenshot of the German fuel price comparison website ADAC. All stations and prices close to the stated address (Mannheim) are shown, i.e., the website is fully transparent.

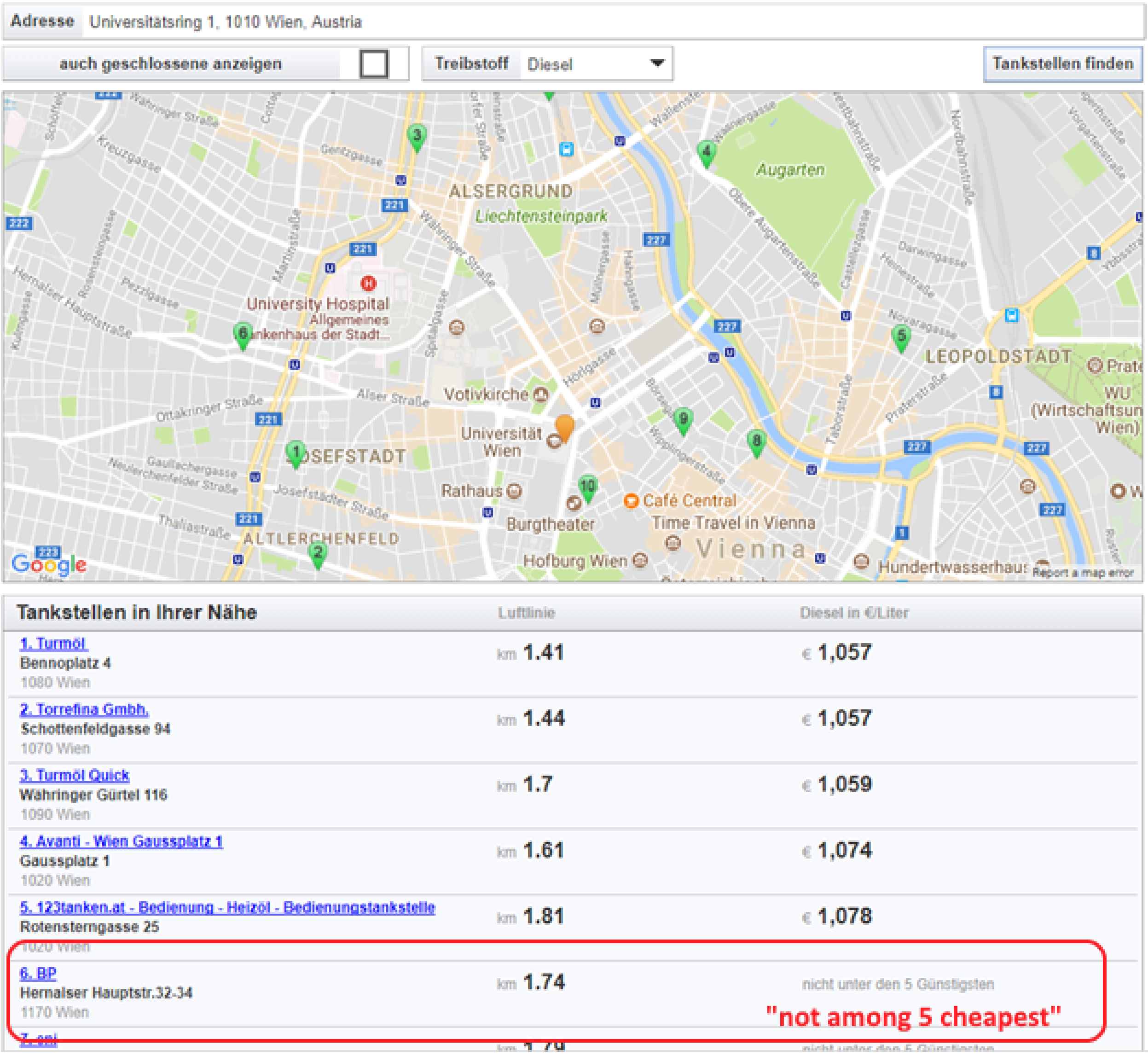


Figure 2: Austrian fuel price comparison website

Notes: This figure shows a screenshot of the Austrian fuel price comparison website Spritpreisrechner (the red box is added by the author). Only the cheapest prices are shown. For the remaining stations, only the ranking is shown, but not the price, i.e., transparency is restricted.

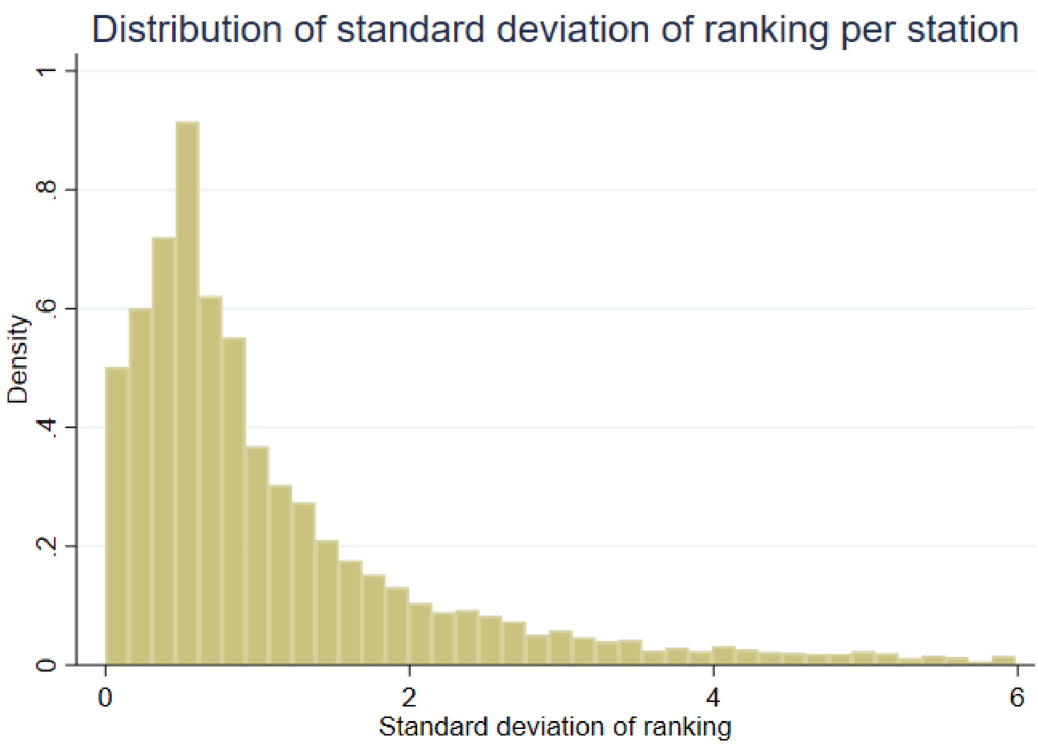


Figure 3: Distribution of standard deviations of ranking across stations

Notes: This figure shows the distribution of standard deviations of ranking across stations. The ranking of most stations fluctuates quite frequently, i.e., consumers are inherently uncertain about current price rankings. The top fifth percentile of stations was omitted from this graph for better readability.

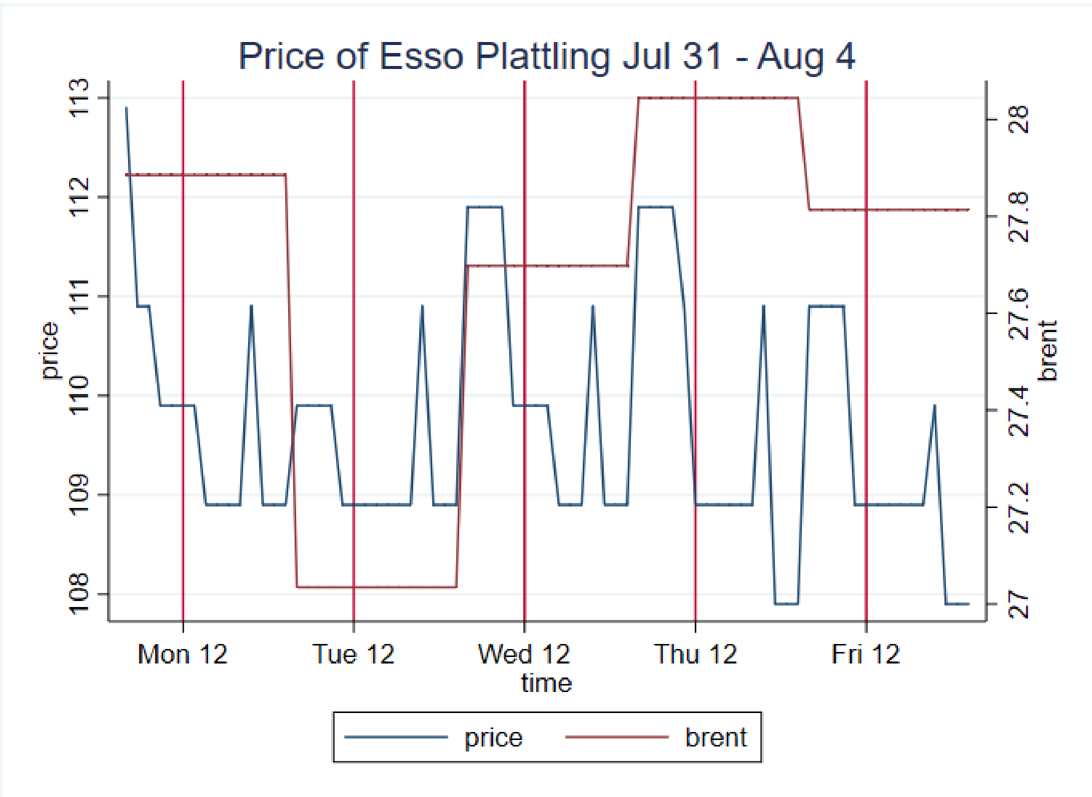


Figure 4: Prices of Esso in Plattling over the first week of August 2017.

Notes: This figure shows the prices of Esso in Plattling over the first week of August 2017, in contrast with the oil price (brent). The diesel price fluctuates frequently even throughout the day, i.e., irrespective of current oil price movements.

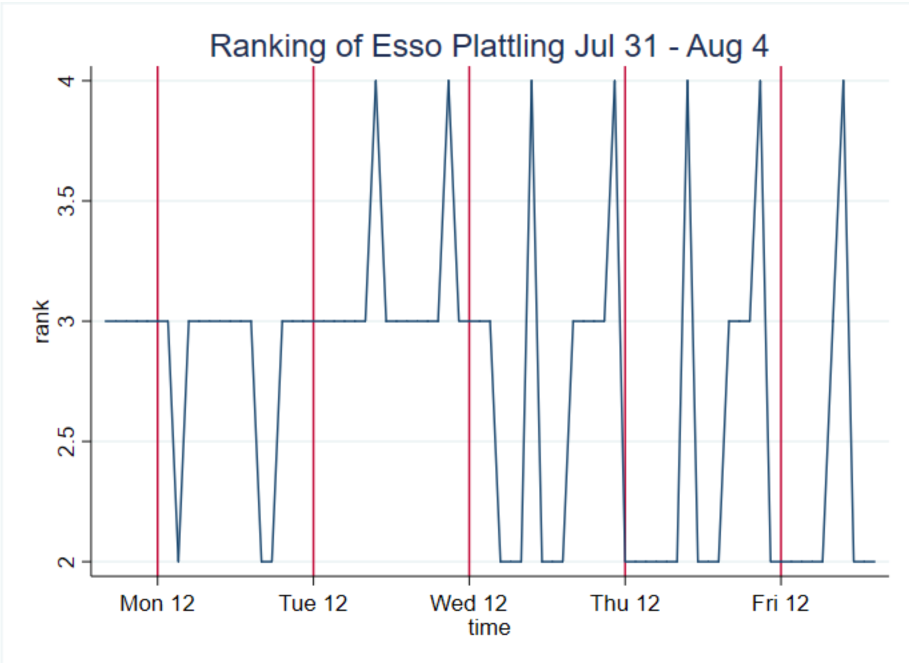


Figure 5: Price ranking of Esso in Plattling over the first week of August 2017

Notes: This figure shows the price ranking of Esso in Plattling over the first week of August 2017. The market consists of four stations. The price ranking of the Esso station fluctuates frequently throughout the day.

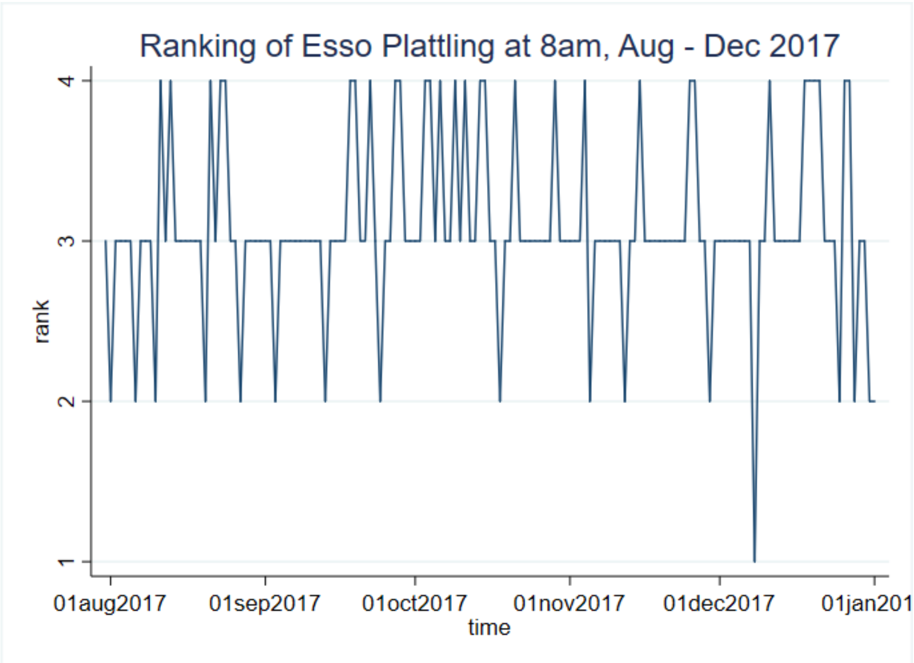


Figure 6: Price ranking of Esso in Plattling at 8am, Aug - Dec 2017

Notes: This figure shows the price ranking of Esso in Plattling daily at 8am, from August to December 2017. The market consists of four stations. The price ranking of the Esso station fluctuates heavily. Hence, consumers are uncertain about price rankings.

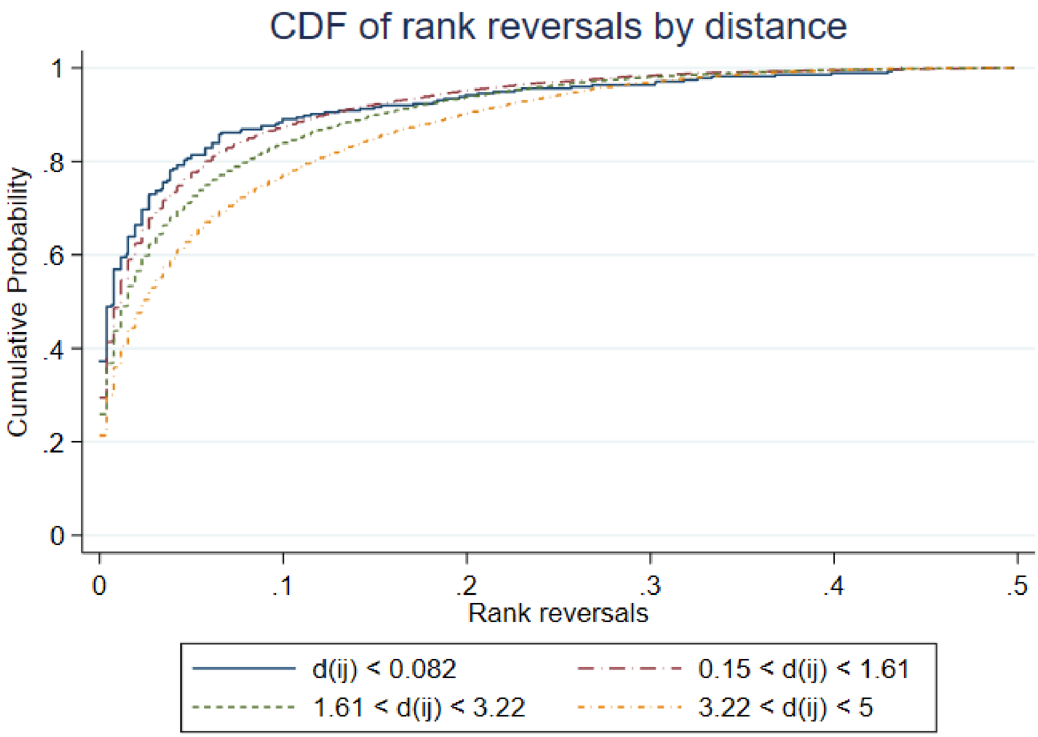
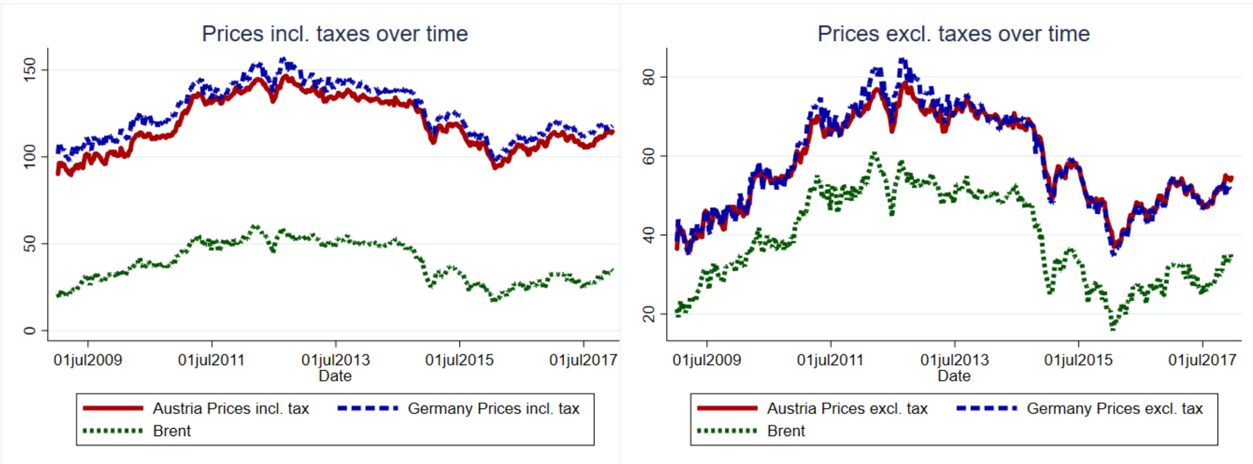


Figure 7: CDF of rank reversals by distance

Notes: This figure shows rank reversals by distance as in Chandra and Tappata (2011, Figure 7). The first group corresponds to the "same corner" group in Chandra and Tappata (2011). In our data, there is no systematic difference between station pairs that are on the same corner or not.



(a) Prices including taxes over time

(b) Prices excluding taxes over time

Figure 8: Prices in Austria and Germany over time.

Notes: This figure shows gasoline prices in Austria and Germany over time, including (panel (a)) and excluding taxes (panel (b)). Excise duties on diesel are 39.7 CPL (plus 20% Value Added Tax) in Austria and 47.04 CPL (plus 19% VAT) in Germany. Both including and excluding taxes, average prices are lower in Austria than in Germany.

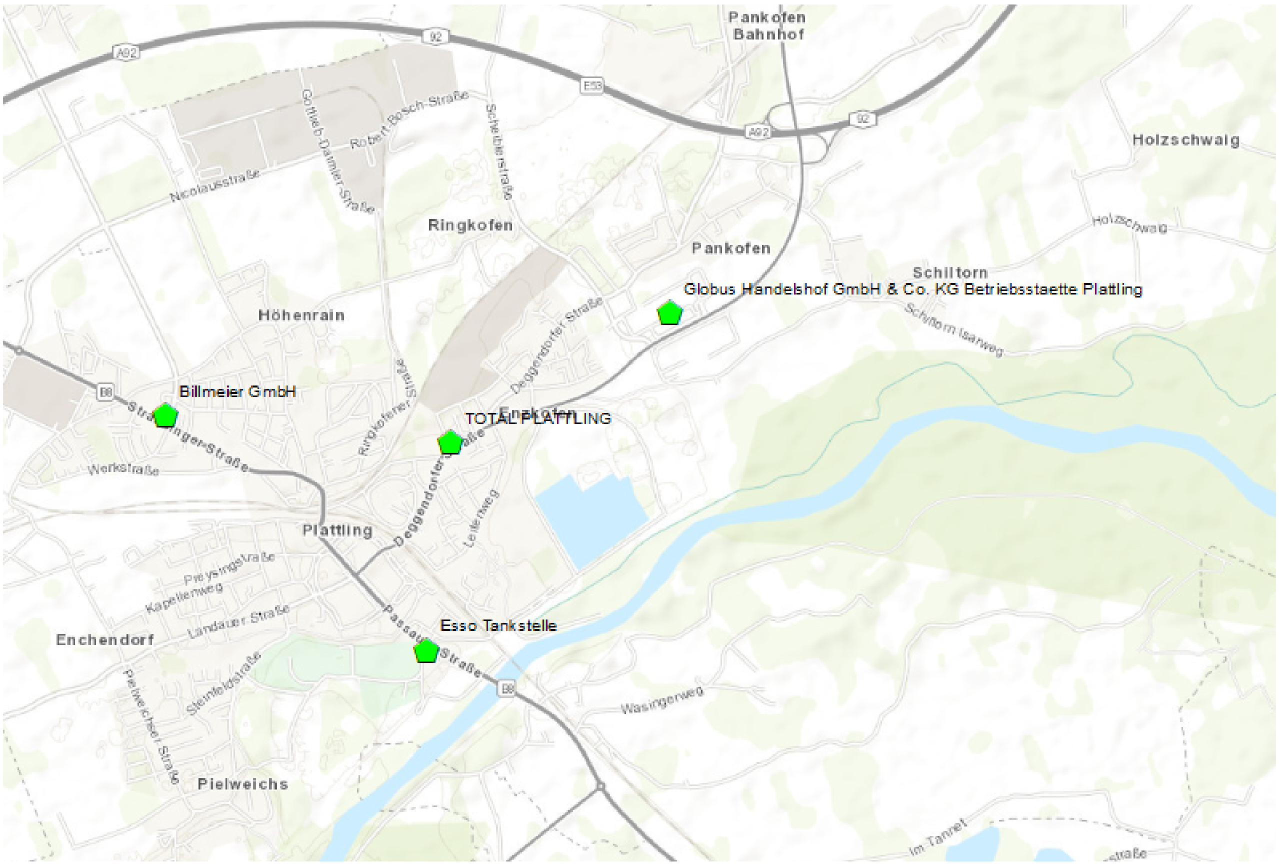


Figure 9: Market around Plattling

Notes: This figure shows the market around Plattling. The air distance from the Total station is 2.3km, 2.0km, and 1.6km, to the Billmaier, Globus, and Esso station, respectively.

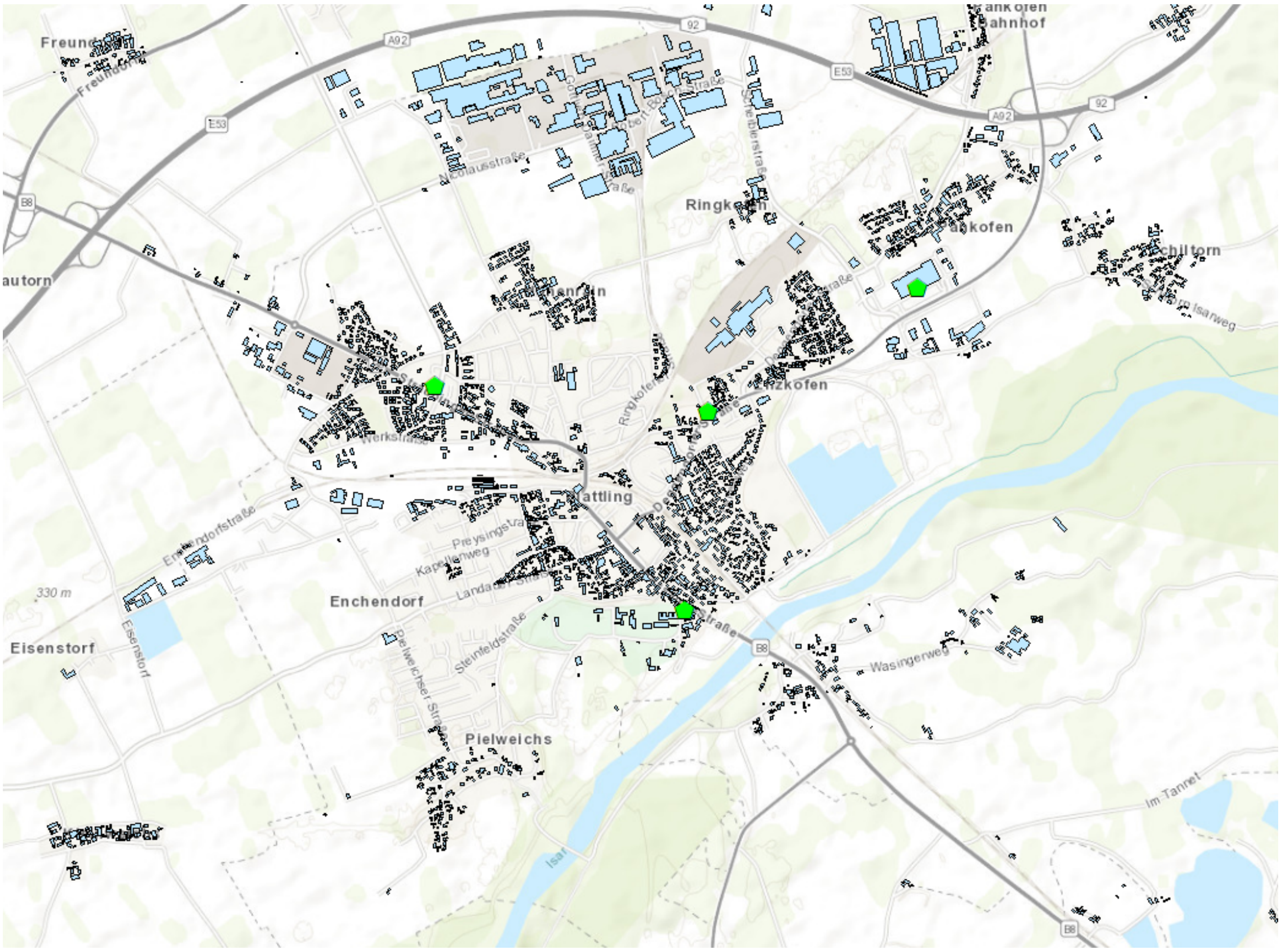


Figure 10: Buildings in Plattling

Notes: This figure illustrates the first step of our approach for proxying for demand. The figure shows the buildings in Plattling. Green pentagons denote gasoline stations.

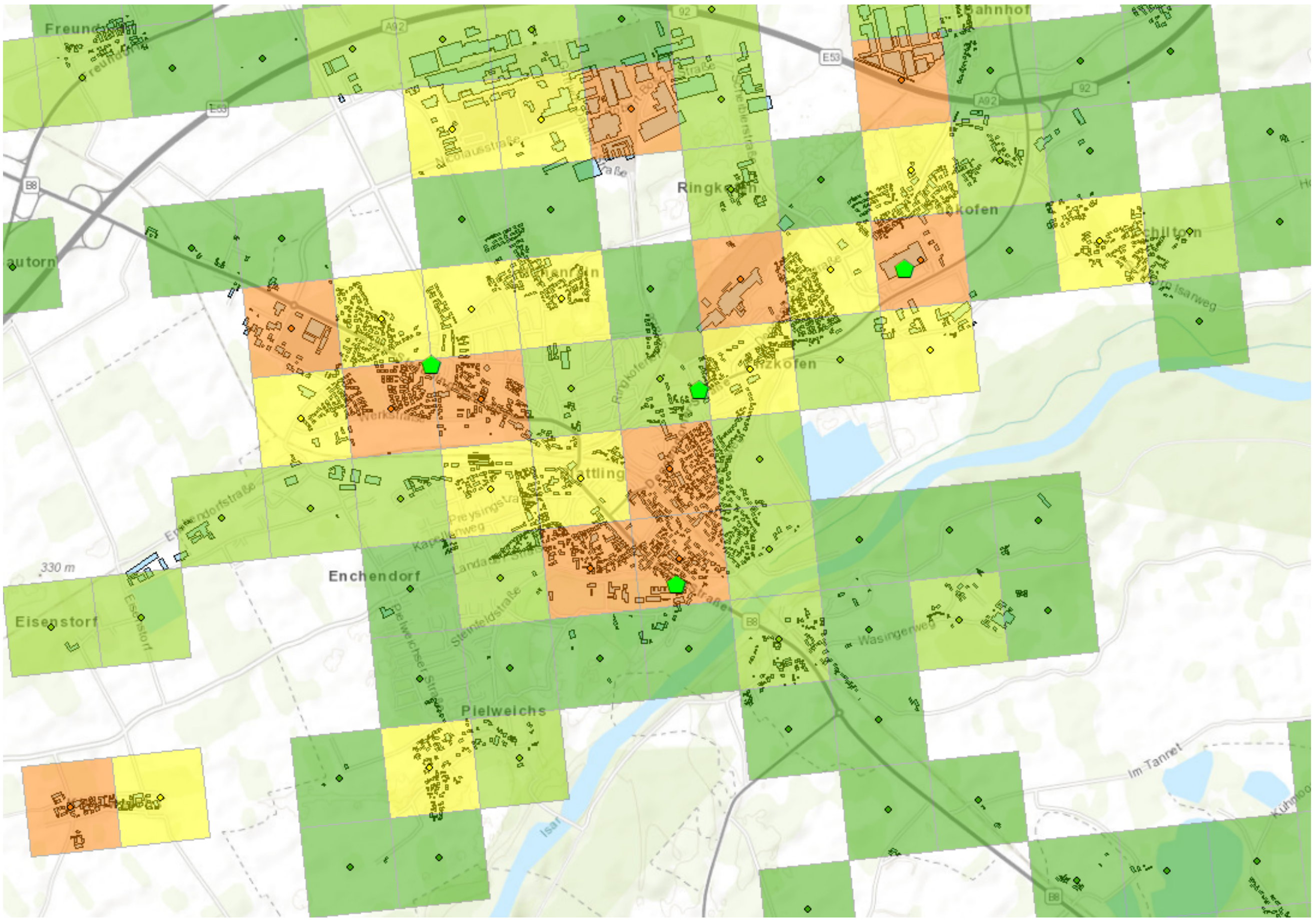


Figure 11: Grid cells in Plattling

Notes: This figure illustrates the second step of our approach for proxying for demand. The figure shows 500x500 meter large grid cells in Plattling, color coded by estimated population density (green = low density, red = high density). Green dots represent centroids of the cells, used subsequently for demand estimation. Green pentagons denote gasoline stations.

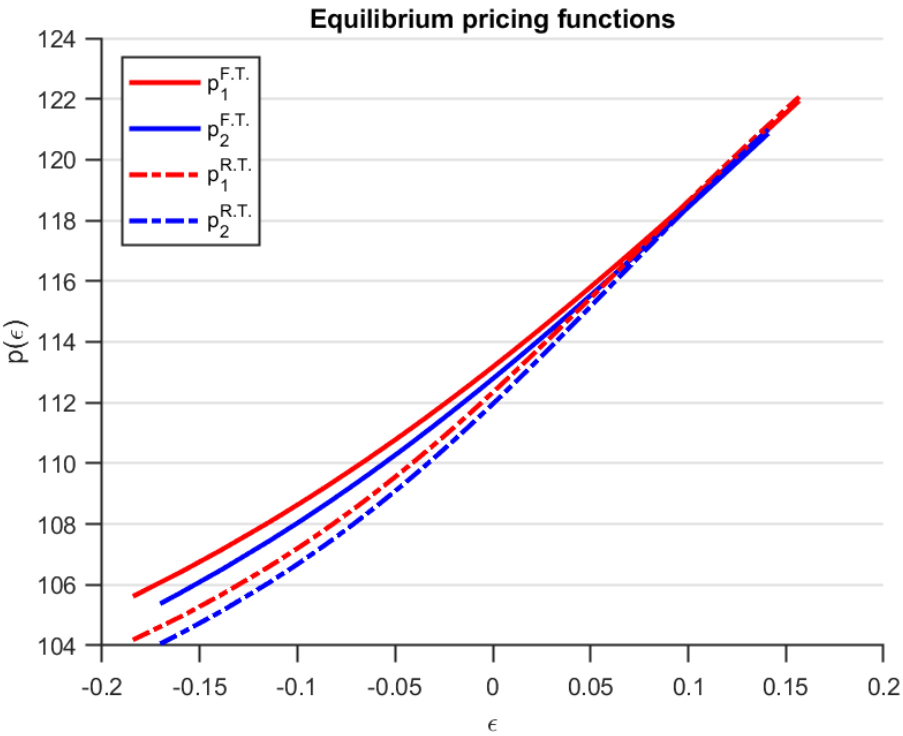


Figure 12: Equilibrium price functions

Notes: This figure shows the equilibrium price functions of two stations (Total and Esso) in Plattling, given marginal cost shocks ε. The solid lines are computed under the baseline scenario with full transparency $\rho = 1$ (F.T.), whereas the dashed lines are under restricted transparency $\rho = 0.5$ (R.T.). Prices are lower, on average, but price dispersion increases. Only firms that obtained a low cost draw attempt to attract online consumers, and otherwise they charge quasi-monopoly prices to their loyal consumer segment.

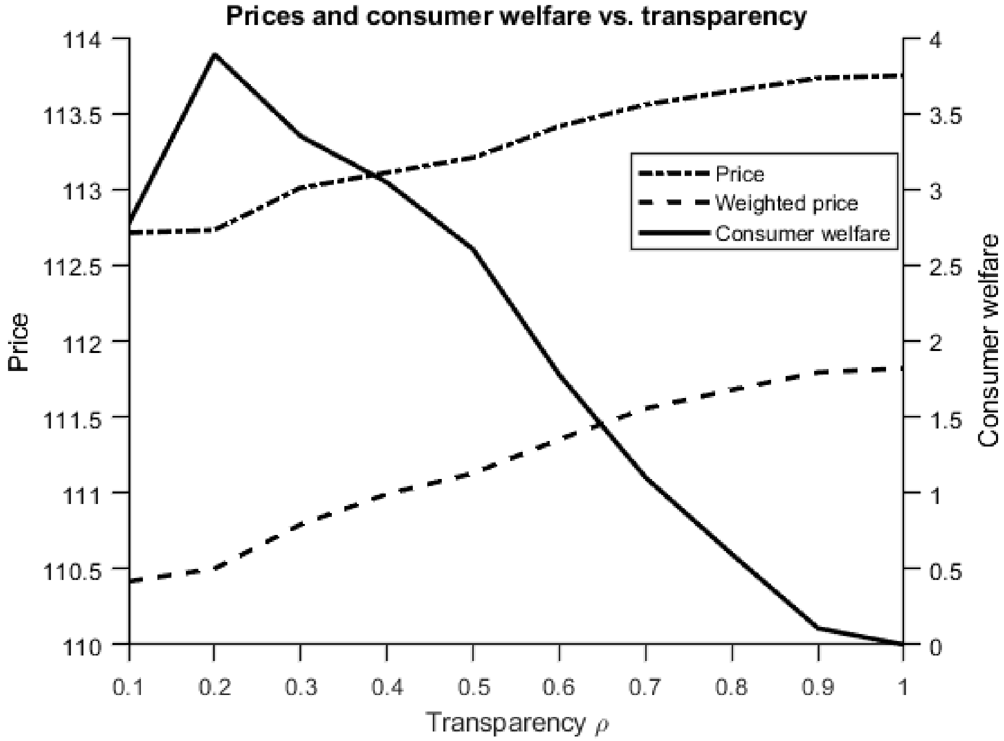


Figure 13: Counterfactual prices and consumer welfare vs. transparency

Notes: This figure shows counterfactual prices (on the left axis in CPL) and consumer welfare (on the right axis, relative to baseline margins) as a function of the market transparency regime ρ. In the baseline scenario with $\rho = 1$, prices are relatively high and consumer welfare is relatively low. Restricting transparency induces fierce competition for attention and hence reduces prices. Initially, consumer surplus increases (up to $\rho = 0.2$), but eventually consumer surplus decreases again due to inefficient matching.

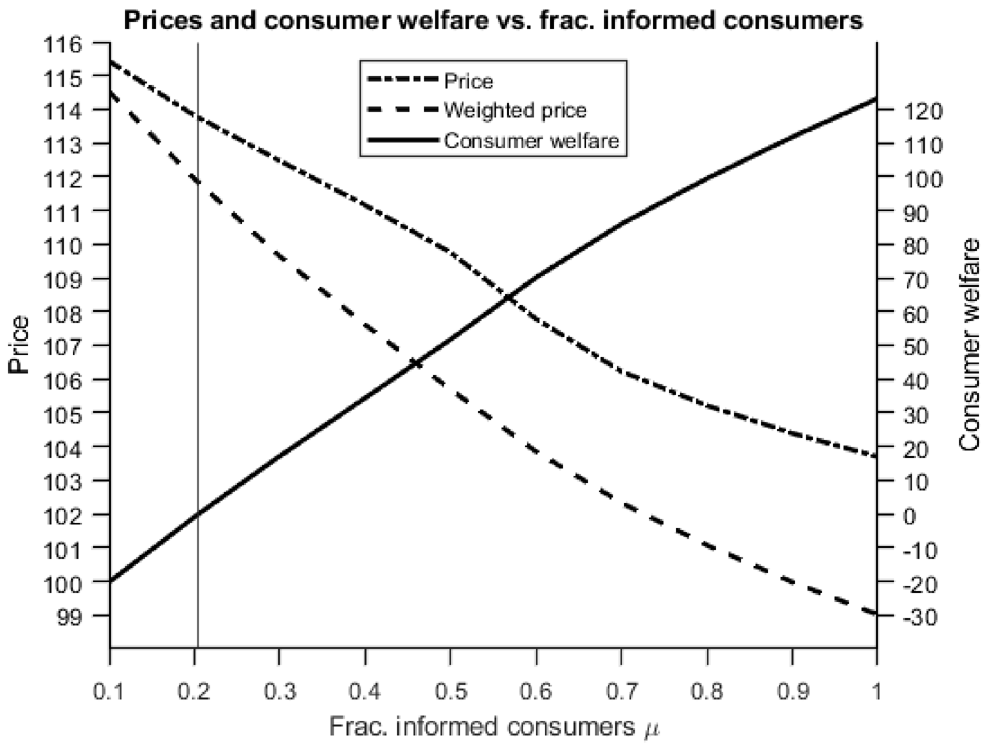


Figure 14: Counterfactual prices and consumer welfare vs. fraction of informed consumers

Notes: This figure shows counterfactual prices (on the left axis in CPL) and consumer welfare (on the right axis, relative to baseline margins) as a function of the fraction of informed consumers μ. In the baseline scenario, $\mu = 0.205$. Prices decrease monotonically and consumer welfare increases monotonically in μ. In order to reach an increase in consumer welfare by 4% (as achieved by a reduction in transparency to $\rho = 0.2$), a fraction of informed consumers of $\mu' \approx 0.25$ would be needed. This would require educating consumers toward using price comparison websites more frequently. Conversely, reducing transparency to facilitate competition for attention merely requires small changes to websites.

C Rank reversals

The rank reversal statistic r_{ij} is created for each station pair (i, j) as follows (Chandra and Tappata, 2011) :

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

where labeling stations i and j such that i tends to be more expensive ensures that $0 \leq r_{ij} \leq 0.5$. The standard deviation of price differences between stations i and j is defined as

$$\sigma_{ij} = \sqrt{\frac{1}{T_{ij}} \sum_{t=1}^{T_{ij}} (s_{ijt} - \overline{s}_{ijt})^2}$$

where $s_{ijt} = p_{jt} - p_{it}$ and $\overline{s}_{ijt} = \frac{1}{T_{ij}} \sum_{t=1}^{T_{ij}} s_{ijt}$.

D Measure for fraction of informed consumers

Although we are eventually interested in diesel refuelings, we base our calculations on total gasoline consumption because gasoline data is less distorted by truck refuelings. In Germany, total gasoline consumption in 2017 was 18.30m tons, implying a total of 976m refueling events at 25 liters each. Our measure of total annual visits is based on advertising impressions published by the German Audit Bureau of Circulation (IVW).[15] The three largest providers Clever Tanken (including website visits, listed and weighted as Autobild), Mehr Tanken, and T-Mobile Tanken App combined had 304.69m visits in 2017, out of which 199.6m can be attributed to gasoline car owners (proportional to the 65.5% of gasoline cars registered in Germany).[16] Thus, our measure of informed consumers is $\mu = 199.6/976 \approx 20.5\%$.[17]

Our calculations are imperfect for various reasons. On the one hand, we may overestimate the actual fraction of informed consumers, because not each website or app query may be associated with actual refueling intent. On the other hand, we may underestimate the actual fraction of informed consumers, because there are also various other sources of price information, e.g., teletext in Austria and other

[15] http://ausweisung.ivw-online.de/

[16] These calculations imply that, on average, in Germany there are 834.78k total visits per day and 10.1 visits per registered gasoline car per year.

[17] Cabral et al. (2018) base their calculation for the fraction of informed consumers on the number of app downloads and obtain 18%. Our approach takes into account actual app and website visits.

service providers (most noteworthy the ADAC, whose app is not listed by IVW since it does not include ads) in Germany.

Notwithstanding these drawbacks, we do believe that our measure approximates actual numbers very well.

E Intuition for identification

Suppose we only have cross-sectional price data from different markets, varying by number of firms n. In particular, we do not observe quantity data. How can we still simultaneously identify supply and demand parameters? In each market, firms compete in prices and demand is given by the usual logit expression. The structural parameters of interest $\theta = (c, v, \gamma)$ where c is marginal cost c (for simplicity no incomplete information here) and v and γ are demand parameters, giving rise to the following demand expression for firm i:

$$D_i(p_i, p_{-i}) = \frac{\exp(v - \gamma p_i)}{1 + \sum_{j=1}^{n} \exp(v - \gamma p_j)}$$

and profits

$$\pi_i(p_i) = (p_i - c - \varepsilon_i) D(p_i, p_{-i})$$

where ε_i here is a common knowledge cost shock. Importantly, there is no shock in the demand expression by assumption.

In a symmetric equilibrium, equilibrium prices satisfy

$$p = c + \frac{1 + \exp(v - \gamma p)}{1 + (n-1)\exp(v - \gamma p)} \frac{1}{\gamma}$$

which can be approximated by a second order Taylor polynomial as

$$p \approx c + \sum_{k=0}^{2} f_k(v, \gamma, n) p^k$$

and hence

$$p \approx \phi_0(\theta; n) + \phi_1(\theta) n + \phi_2(\theta) n^2$$

where f_k and ϕ_k are known non-linear functions in all parameters. In principle, regressing p on a constant, n and n^2 identifies the $(\phi_k)_{k=0}^{2}$, each of which are functions of θ. So we have 3 equations in 3 unknown parameters (c, v, γ), and hence purely variation in n identifies the supply- and demand parameters simultaneously.

We actually use non-linear GMM instead of approximation, so approximation error is less of a concern, but the basic intuition prevails.

Additionally, we have way more information available than in this simply setting. A similar argument shows that variation in observable input prices identifies θ when observing only a single market over time. We combine both sources of variation and moreover employ macro moments to enhance efficiency.

References

Abaluck, J. and Adams, A. (2020). What do consumers consider before they choose? identification from asymmetric demand responses. *Quarterly Journal of Economics*.

Albæk, S., Møllgaard, P., and Overgaard, P. B. (1997). Government-assisted oligopoly coordination? a concrete case. *The Journal of Industrial Economics*, 45(4):429–443.

Armstrong, M. (2017). Ordered consumer search. *Journal of the European Economic Association*, 15(5):989–1024.

Armstrong, M., Vickers, J., and Zhou, J. (2009). Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233.

Armstrong, M. and Zhou, J. (2011). Paying for prominence. *The Economic Journal*, 121(556).

Ater, I. and Rigbi, O. (2017). The effects of mandatory disclosure of supermarket prices.

Athey, S. and Haile, P. A. (2002). Identification of standard auction models. *Econometrica*, 70(6):2107–2140.

Athey, S. and Haile, P. A. (2007). Nonparametric approaches to auctions. *Handbook of econometrics*, 6:3847–3965.

Barseghyan, L., Molinari, F., and Thirkettle, M. (2020). Discrete choice under risk with limited consideration.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Borenstein, S., Cameron, A. C., and Gilbert, R. (1997). Do gasoline prices respond asymmetrically to crude oil price changes? *The Quarterly Journal of Economics*, 112(1):305–339.

Borenstein, S. and Shepard, A. (1996). Dynamic pricing in retail gasoline markets. *The RAND Journal of Economics*, 27(3):429–451.

Brenkers, R. and Verboven, F. (2006). Liberalizing a distribution system: the european car market. *Journal of the European Economic Association*, 4(1):216–251.

Bundeskartellamt (2011). Sektoruntersuchung Kraftstoffe: Abschlussbericht Mai 2011.

Bundesministerium für Wirtschaft und Energie (2018). Evaluierungsbericht zur Markttransparenzstelle für Kraftstoffe.

Bundeswettbewerbsbehörde (2011). Der österreichische Kraftstoffmarkt.

Burdett, K. and Judd, K. L. (1983). Equilibrium price dispersion. *Econometrica: Journal of the Econometric Society*, pages 955–969.

Butters, G. R. (1977). Equilibrium distributions of sales and advertising prices. *Review of Economic Studies*, 44(3).

Cabral, L., Schober, D., and Woll, O. (2018). Search and equilibrium prices: Theory and evidence from retail diesel.

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

Dana, J. D. (1994). Learning in an equilibrium search model. *International Economic Review*, pages 745–771.

Dewenter, R. and Heimeshoff, U. (2012). *Less pain at the pump? The effects of regulatory interventions in retail gasoline markets.* Number 51. DICE Discussion paper.

Dewenter, R., Heimeshoff, U., and Lüth, H. (2016). The impact of the market transparency unit for fuels on gasoline prices in germany. *Applied Economics Letters*, pages 1–4.

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2018). Consumer price search and platform design in internet commerce. *American Economic Review*, 108(7):1820–59.

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

Eibelshäuser, S. and Wilhelm, S. (2017). Markets take breaks: Dynamic price competition with opening hours.

Gourieroux, C. and Monfort, A. (1996). *Simulation-based econometric methods*. Oxford university press.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574.

Haucap, J., Heimeshoff, U., and Siekmann, M. (2016). *Selling gasoline as a by-product: The impact of market structure on local prices*. Number 240. DICE Discussion paper.

Haucap, J., Heimeshoff, U., Siekmann, M., et al. (2017). Fuel prices and station heterogeneity on retail gasoline markets. *The Energy Journal*, 38(6):81–103.

Heiss, F., McFadden, D., Winter, J., Wuppermann, A., and Zhou, B. (2016). Inattention and switching costs as sources of inertia in medicare part d. Technical report, National Bureau of Economic Research.

Hong, H. and Shum, M. (2006). Using price distributions to estimate search costs. *The RAND Journal of Economics*, 37(2):257–275.

Honka, E., Hortaçsu, A., and Vitorino, M. A. (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *The RAND Journal of Economics*, 48(3):611–646.

Horvath, M. (2019). Germany's market transparency unit for fuels: Fostering collusion or competition?

Houde, J.-F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, 102(5):2147–82.

Imbens, G. W. and Lancaster, T. (1994). Combining micro and macro data in microeconometric models. *The Review of Economic Studies*, 61(4):655–680.

Ivaldi, M., Jullien, B., Rey, P., Seabright, P., Tirole, J., et al. (2003). The economics of tacit collusion. *Final Report for DG Competition, European Commission*.

Janssen, M., Pichler, P., and Weidenholzer, S. (2011). Oligopolistic markets with sequential search and production cost uncertainty. *The RAND Journal of Economics*, 42(3):444–470.

Kühn, K.-U. and Vives, X. (1995). *Information exchanges among firms and their impact on competition*. European Communities.

Luco, F. (2019). Who benefits from information disclosure? the case of retail gasoline. *American Economic Journal: Microeconomics*, 11(2):277–305.

MacKay, A. and Miller, N. (2019). Estimating models of supply and demand: Instruments and covariance restrictions. *Available at SSRN 3025845*.

MacMinn, R. D. (1980). Search and market equilibrium. *Journal of Political Economy*, 88(2):308–327.

Manuszak, M. D. (2010). Predicting the impact of upstream mergers on downstream markets with an application to the retail gasoline industry. *International Journal of Industrial Organization*, 28(1):99–111.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026.

Montag, F. and Winter, C. (2020). Price transparency against market power. *Mimeo*.

Moraga-González, J. L., Sándor, Z., and Wildenbeest, M. R. (2017). Prices and heterogeneous search costs. *The RAND Journal of Economics*, 48(1):125–146.

Moraga-González, J. L., Sándor, Z., and Wildenbeest, M. R. (2018). Consumer search and prices in the automobile market.

Moraga-González, J. L., Sándor, Z., and Wildenbeest, M. R. (2020). Simultaneous search for differentiated products: The impact of search costs and firm prominence. *CEPR Discussion Paper No. DP14454*.

Moraga-González, J. L. and Wildenbeest, M. R. (2008). Maximum likelihood estimation of search costs. *European Economic Review*, 52(5):820–848.

Myśliwski, M., Sanches, F., Silva, D., and Srisuma, S. (2017). Identification and estimation of a search model: A procurement auction approach. *Available at SSRN 2968058*.

Nevo, A. (2000a). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, pages 395–421.

Nevo, A. (2000b). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of economics & management strategy*, 9(4):513–548.

Nishida, M. and Remer, M. (2018). The determinants and consequences of search cost heterogeneity: Evidence from local gasoline markets. *Journal of Marketing Research*, 55(3):305–320.

Noel, M. D. (2007a). Edgeworth price cycles, cost-based pricing, and sticky pricing in retail gasoline markets. *The Review of Economics and Statistics*, 89(2):324–334.

Noel, M. D. (2007b). Edgeworth price cycles: Evidence from the toronto retail gasoline market. *The Journal of Industrial Economics*, 55(1):69–92.

Obradovits, M. (2014). Austrian-style gasoline price regulation: How it may backfire. *International Journal of Industrial Organization*, 32:33–45.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., and Yontcheva, B. (2020). Information and price dispersion: Theory and evidence. *International Economic Review*, 61(2):871–899.

Pennerstorfer, D. and Weiss, C. (2013). Spatial clustering and market power: Evidence from the retail gasoline market. *Regional Science and Urban Economics*, 43(4):661–675.

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

Politis, D. N. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, pages 2031–2050.

Richardson, S. (2004). Integral equations. *MATHEMATICA JOURNAL.*, 9(2):460–482.

Rossi, F. and Chintagunta, P. K. (2016). Price transparency and retail prices: Evidence from fuel price signs in the italian highway system. *Journal of Marketing Research*, 53(3):407–423.

Ryan, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica*, 80(3):1019–1061.

Schultz, C. (2005). Transparency on the consumer side and tacit collusion. *European Economic Review*, 49(2):279–297.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, pages 105–130.

Sovinsky Goeree, M. (2008). Limited information and advertising in the us personal computer industry. *Econometrica*, 76(5):1017–1074.

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

Thomadsen, R. (2005). The effect of ownership structure on prices in geographically differentiated industries. *RAND Journal of Economics*, pages 908–929.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Varian, H. R. (1980). A model of sales. *The American Economic Review*, 70(4):651–659.

Wang, Z. (2009). (mixed) strategy in oligopoly pricing: evidence from gasoline price cycles before and under a timing regulation. *Journal of Political Economy*, 117(6):987–1030.

Wildenbeest, M. R. (2011). An empirical model of search with vertically differentiated products. *The RAND Journal of Economics*, 42(4):729–757.

Wolinsky, A. (1986). True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, pages 493–511.