Market Liberalization: Price Dispersion, Price Discrimination and Consumer Search in German Electricity Markets*

Klaus Gugler[†] Sven Heim[‡] Maarten Janssen[§]
Mario Liebensteiner[¶]

December 15, 2017

Abstract

How does the effect of market liberalization depend on consumer search? We answer this question by studying the German electricity retail market as an example. We first construct a simple theoretical model where an incumbent can price discriminate between searching and loyal consumers and consumers with idiosyncratic search costs have to decide whether or not to search after observing the baseline price of the incumbent. One result is that if there are relatively more consumers with very low search cost, price dispersion and price discrimination increase with the share of searching consumers. Using a unique panel dataset on spatially varying search queries at major online price comparison websites for the period 2011–2014 we empirically show that local incumbents increase their baseline rate while entrants decrease their tariffs as a reaction to more consumer search. Moreover, the incumbent increasingly price discriminates and sets its cheapest tariff much lower in reaction to consumer search. Price discrimination allows the incumbent to segment markets according to search intensity, and to simultaneously follow surplus appropriation for consumers in the non-competitive segment as well as business stealing strategies for consumers in the competitive segment.

Keywords: Search, Price Dispersion, Price Discrimination, Electricity Retail
JEL Classification: D43, D83, L11, L13, Q40

*We are deeply thankful to ene't for giving us access to their data on consumer search at price comparison websites. We thank Ulrich Laitenberger, Dieter Pennersdorfer, Steven Puller, Philipp Schmidt-Dengler, Mike Waterson, and Biliana Yontcheva as well as the audiences at the EARIE 2016 and CRESSE 2016 conferences for valuable comments. We also thank Acxiom for giving us data on household characteristics. Janssen acknowledges financial support from the Austrian Science Foundation FWF, under project number P 27995-G27.

[†]Vienna University of Economics and Business (WU), klaus.gugler@wu.ac.at

[‡]Mines ParisTech, Paris, and ZEW Centre for European Economic Research Mannheim, sven.heim@mines-paristech.fr

[§]University of Vienna, National Research University Higher School of Economics Moscow and CEPR., maarten.janssen@univie.ac.at

[¶]Vienna University of Economics and Business (WU), mario.liebensteiner@wu.ac.at

1 Introduction

After 20 or more years of market liberalization, an important question is why prices are highly dispersed and average prices paid by consumers remained high after liberalization. The success of market liberalization depends to a large extent on whether or not consumers are willing to search for and switch to alternative providers. Many markets where liberalization has taken place are characterized by an important asymmetry between the incumbent provider and new entrants in that consumers know the contract with their current provider, but have to pay a search cost to be informed of the contracts alternative providers offer. Once, they are informed they also have to decide whether or not to switch. In some of these markets, the incumbent has reacted to the presence of entrants by discriminating between loyal consumers who do not search and searching consumers who threaten to leave.

We analyze the interaction between the search behavior of consumers and the pricing behavior of the incumbent and entrant retailers both empirically and theoretically. In particular, we investigate how the extent of price discrimination and price dispersion depends on the fraction of consumers searching for alternative providers, taking into account that the consumer search behavior is endogenously determined by the pricing behavior of firms. Price discrimination is measured by the price difference the incumbent sets between loyal and searching consumers and price dispersion is measured by the highest price of the incumbent and the lowest price set by entrants.

The theoretical model is inspired by our application to the German electricity markets and has the following elements. Retail firms produce a homogeneous good so that consumers only start searching for better prices. Consumers differ in search cost and, to keep the model simple, switching cost are assumed to be proportional to search cost. Consumers observe the base price the incumbent sets to all consumers at no cost. Having observed the base price, consumers have to decide whether or not to search. Search is costly and allows the consumers to observe all other prices in the market, i.e., a search protocol known as "newspaper search" (see, e.g., Dana (1994) and Tappata (2009)). One can think of search as subscribing to a platform or visiting a price comparison website so that at a (emotional) cost consumers have access to all prices. At the platform, consumers choose between buying from the lowest-price entrant or staying with the incumbent at the price the incumbent announced on the platform. Consumer specific brand loyalty or switching cost imply the incumbent can charge higher prices on the platform than entrants. We show that by varying the search cost distribution, this simple model can accommodate a rich pattern of pricing and search behaviors. For example, if the fraction of consumers with very small search cost increases, price dispersion and price discrimination increase simultaneously with the fraction of consumers who search online. On the other hand, if the fraction of consumers with intermediate search cost increases at the expense of the fraction of consumers with high search cost, then price dispersion and price discrimination may decrease simultaneously with the fraction of consumers who search online.

The empirical part of our analysis employs a unique data set on retail electricity prices and consumer search at the German zip code level for the period 2011–2014. The German retail electricity market has been liberalized at the end of the previous millennium, where former local monopolies have been replaced by local retail competition. Since then, local incumbents compete with new entrants and all consumers that have initially stayed with their local incumbent supplier at the baseline tariff, have the freedom to engage in search and to switch to any alternative tariff offered for their local address either by new entrants or by the incumbent. However, some two decades after liberalization retail prices in Germany[1] remain high and even increased after liberalization. In 2015 still 76% of households stay with the incumbent, with 33% remaining at the expensive baseline tariff, while 43% switching to a cheaper incumbent tariff, and only 24% having switched to a new entrant (see the 2015 report of the German regulatory authority (Bundesnetzagentur, 2015)). In recent years, most consumers use online platforms to search for cheaper rates.[2] However, the practice of price discrimination of the incumbent supplier segments the market into consumers with higher and lower search and switching cost restricting the benefits of search to a part of the population. This may partly explain why retail prices on average do not fall, as figure 1 indicates (with cost increases an important other factor).

A key aspect of our dataset is the fact that we have data on search intensity at the zip code level. In particular, we have data on consumer activity (i.e., the actual number of search queries) at online comparison platforms, and given that most of the search for lower prices is via these platforms, we interpret this data as a direct measure of the search intensity at the local level. Such data are generally not observable and other empirical studies on consumer search markets often rely on indirect measures of consumer search activity.[3] Moreover, we are not aware of any paper having a direct measure of consumer search in a panel data context. In terms of prices, we have access to the baseline and online price of the incumbent and the lowest online price. Using these data, we empirically

[1]Retail prices remained not only high or even increased after liberalization in Germany but in many other European countries (see e.g. Cabral (2017)) and in the USA (see e.g. Hortacsu et al. (2015)).

[2]According to a 2011 survey 80% of the switchers searched online for alternative providers (`"http://www.atkearney.at/documents/3709812/3710656/BIP_Der_Strom_und_Gasvertrieb_im_Wandel.pdf"`, March 3, 2016). This number is likely to have increased in more recent years.

[3]E.g., Many papers proxy for lower search costs or consumer informedness by access to the internet or online versus offline purchases (Brown and Goolsbee, 2002; Brynjolfsson and Smith, 2014). Similarly, Pennersdorfer et al. (2014) use commuters versus non-commuters to distinguish between shoppers and non-shoppers.

show that local incumbents have higher baseline rates, while entrants have lower rates in municipalities with more consumer search. Moreover, the incumbent increases the extent of price discrimination and sets its cheapest, online tariff much lower in zip codes with more consumer search. Price discrimination allows the incumbent to segment markets according to search intensity, and to simultaneously follow surplus appropriation (from non-shoppers) as well as business stealing strategies (from shoppers).

At a methodological level, it is important to note that search intensity and pricing are determined simultaneously with search intensity being a function of (expected) price differences while prices are affected by the intensity with which consumers search. We circumvent the potential reverse causality of price patterns on search intensity by adopting an instrumental variables approach, allowing us to interpret the estimated empirical relationships as causal. The instruments we adopt are the availability of broadband internet and the share of households with a head below the age of 40. Both are thought to raise the likelihood of online shopping and are exogenous to pricing. Moreover, they have sufficient temporal and spatial variation to identify search intensity.

The German electricity market is just one example of the set of markets where over the last 20–30 years important market liberalizations have taken place. Many of these markets (including electricity markets in many states of the USA[4] and Canada, the UK, other EU Member States – following up the EU energy reform process –, other European countries (e.g. Norway) and other parts of the world) share important features with the features of the German electricity market. In all these markets, entrants have entered, but because consumers buy at a continuous basis from the incumbent, there is an important asymmetry as consumers know the base price of the incumbent whereas they have to incur a search cost to learn other prices. Moreover, incumbents may engage in price discrimination between consumers who somehow indicated they will terminate the baseline contract. There are many examples of other sectors, such as natural gas, telecommunications, health insurance, railways, postal services, and airlines, that have been liberalized where consumers may switch away from incumbents (or established firms) to alternative suppliers.[5] A key dividing line between these examples is whether or not consumers have ongoing contracts with their suppliers or not. Thus, like electricity markets, telecommunications and health insurance markets have the feature that consumers are naturally informed about their current supplier and will automatically continue their contract as long as they do not search for and switch to alternatives.[6] In most of the other markets

[4]Some states have (partly) abandoned the deregulation reforms as a consequence of the California electricity crisis in 2001 (followed shortly afterwards by the Ontario electricity crisis).

[5]There are other examples of markets that work in a similar fashion but are not subject to liberalization. For example, patents on drugs allow for a temporary monopoly of the patent holder. Once the patent expires, pharmaceutical companies face competition from suppliers of generic drugs.

[6]Retail banking shares this feature of firms enjoying incumbency advantages, although because of the fact that the banking sector is not liberalized, instead of there being one natural incumbent (the former

mentioned, the natural incumbency effect is much smaller as consumers do not have an ongoing contract for continuous delivery.

Given our focus on the role of consumer search in market liberalization processes and the impact of search on price dispersion and price discrimination, it is clear the paper is related to many different strands of literature.

By now, there is a large and varied, theoretical and empirical, literature on how consumer search affects price dispersion in homogeneous goods markets (see, e.g. De los Santos et al. (2012), Chandra and Tappata (2011), and Sorensen (2000) for empirical studies, and Stahl (1989) and Janssen and Moraga-González (2004) for theoretical contributions). Most of these models are based on random search, whereas in market liberalizations, it is important to take the incumbency effect into account that unless they engage in search, consumers buy from the baseline price of the incumbent.

There is also a growing literature dealing explicitly with search in electricity markets, but most of these papers focus on how consumers search (if at all). Giulietti et al. (2014) analyze the retail electricity market in the United Kingdom and find that roughly half the households had reasonably high search costs. Hortacsu et al. (2015) find that even though households generally switch to alternative retailers only rarely, they do switch more after a "bill shock" in the previous months. Moreover, they also find that households attach a brand advantage to the incumbent. In contrast to our paper both papers do not observe the actual searching behaviour of consumers.

In estimating the relationship between price dispersion and search intensity, the empirical literature has to deal with reverse causality as search intensity may also be a function of pricing patterns in the market. Thus, it is important to have direct and exogeneous measures for search intensity, something that the literature is lacking to date. Brown and Goolsbee (2002) use the variation in the share of consumers searching on the internet as a measure of consumer information. In the market for life insurance, they find that internet usage has a non-monotonic impact on price dispersion, but they also do not address the potential endogeneity of internet usage. More recently, Tang et al. (2010) examine the impact of changes in shopbot use on prices and price dispersion in online book retailing and find that an increase in shopbot use reduces average prices and price dispersion. Pennersdorfer et al. (2014) do find an inverted U-shaped relation between price dispersion and the share of informed consumers (as proxied by the share of commuters) in the Austrian gasoline retail market; Sengupta and Wiggins (2014) investigate whether online and offline prices for airline fares differ but do not find a significant difference.

The above literature does not explicitly deal with the incumbency effects (and the resulting asymmetry between firms that is important to understand liberalizing markets. More precisely, the literature generally assumes that the shares of informed and unin-

state monopoly), each (large) bank enjoys this incumbency feature.

formed consumers are randomly asisgned between firms which clearly does not hold true for markets with an incumbency effect. There is a small literature dealing with price discrimination and incumbency. For the UK retail electricity market, Davies et al. (2014) present evidence suggesting that firms deliberately differentiated their tariff structures, resulting in market segmentation according to consumers' usage. For the US mobile telephone sector, Seim and Viard (2011) show that entry induces firms to lower prices on average and to offer larger menus with more evenly spread usage plans benefiting high-valuation consumers in particular. For the same sector, Miravete (2011) concludes that incumbents – as a response to entry – temporarily increase the fogginess of their tariffs relative to the monopoly phase, while entrants use foggy pricing far less frequently, where foggy pricing refers to the practice of offering tariff options that are always more expensive than other offered tariff options regardless of usage profile. For the US airline industry, Goolsbee and Syverson (2008)'s results indicate that incumbents respond to the threat of entry by substantially reducing average fares on the directly threatened routes, but that they do not cut prices on routes to nearby airports in the same market. This bears some relationship to our result that the incumbent price discriminates more severely between searching consumers who can choose an alternative option and not searching consumers who cannot.

At a theoretical level, the idea that a firm would like to price discriminate against consumers with higher search cost is not new (see, e.g., Salop (1977)). Salop (1977)'s argument is casted, however, in a monopoly setting, and critically depends on the assumption that the monopolist is committed to charging prices according to a price distribution, while consumers can somehow react to changes in the price distribution (assuming they observe the distribution, but not the prices) by adopting a different search strategy. Cabral and Gilbukh (2017) also model firms engaging in price discrimination between active and passive searchers. Unless they pay a search cost, consumers buy from the high price of a firm. By paying the search cost and becoming active, they have access to all low prices of all firms. The focus of Cabral and Gilbukh (2017) is, however, very different from ours. They study a market where symmetric firms face cost shocks and they study the equilibrium price reactions to these cost shocks. In our paper, we want to understand how the asymmetric price behaviour between incumbents and entrants is affected by the presence of more searching consumers.

Our paper explains that incumbent's baseline price (and possibly average prices) may increase when there is more consumer search in the market. The success of market liberalization hinges on the possible strategies of the incumbent. Price discrimination allows the incumbent to segment consumers according to their search intensity, and to simultaneously follow a surplus appropriation strategy on consumers with high search cost as well as a business stealing strategy for consumers with low search cost. This strategy may prevent consumers from switching supplier and at the same time allow the incumbent to appropriate market power rents.

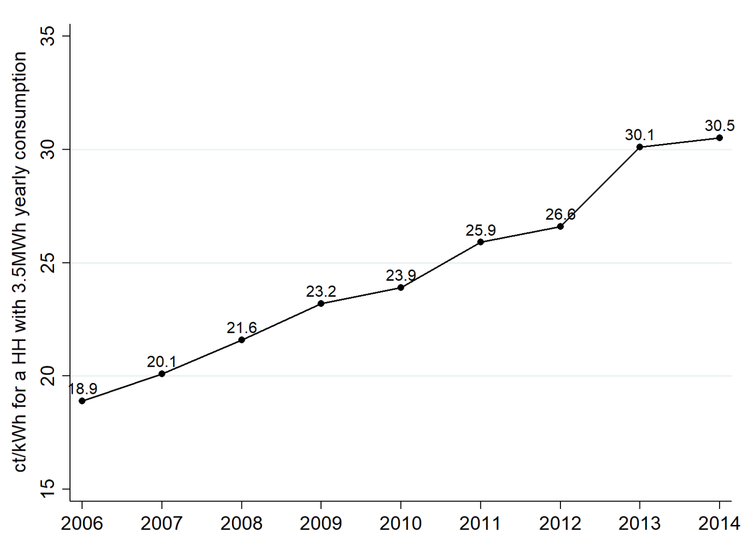
The rest of the paper is structured as follow. The next section describes the German retail electricity market in more detail. We then provide a theoretical model of the main features of this (and other) markets and show which pricing behaviors and search patterns may result from changes in the search cost distribution. We then put the theory to empirical scrutiny by investigating the search-pricing nexus in the German retail electricity market. Section 4 describes the empirical identification strategy, section 5 discusses the data, and section 6 presents the econometric results and provides robustness. Section 7 concludes.

2 Germany's Retail Electricity Market

In 1999 Germany's electricity liberalization brought about the end of local monopolies and opened the market for new entrants with free supplier choice for retail consumers. It was believed that increased competition and freedom of consumer choice would eventually result in lower retail margins with large economic benefits for consumers. However, price dispersion remained high after liberalization (see figure 1). Prior to market liberalization, the local incumbent served all local customers in its distribution grid area at a regulated tariff. After liberalization, the incumbents are legally obliged to sell electricity at a baseline tariff to all households which do not choose another supplier. Households who engage in search have the freedom to switch to any alternative tariff offered for their local address either by new entrants or by the incumbent.

Retail entry in the German electricity market is relatively easy, as also witnessed by the large number of active firms of on average 133 firms per zip code area. Entry is made easy by regulatory measures, such as non-discriminatory access to all customers in each regional market, and regulated and – for all firms in a given coverage area equal – fixed and variable grid charges, feed in tariffs for renewables, and concession fees. Costs for wholesale electricity can also be viewed as comparable across retailers, since wholesale electricity prices are determined centrally at the European Energy Exchange (EEX). EEX prices are representative since around 50% of electricity consumption is traded via the EEX, and they represent the opportunity costs of electricity even for vertically integrated suppliers. Thus, cost variation between firms within the same supply area is very small, the main costs that may vary across firms within a supply area are costs for the purchase of wholesale electricity due to individual hedging strategies as well as adminstrative costs. According to Bundesnetzagentur (2013, p. 152), purchase costs of wholesale electricity account for 21.3% and selling expenses (including retail margins) only for 7.5% of the total electricity bill of household consumers. Thus, electricity retail

Figure 1: Development of the Incumbent Base Tariffs

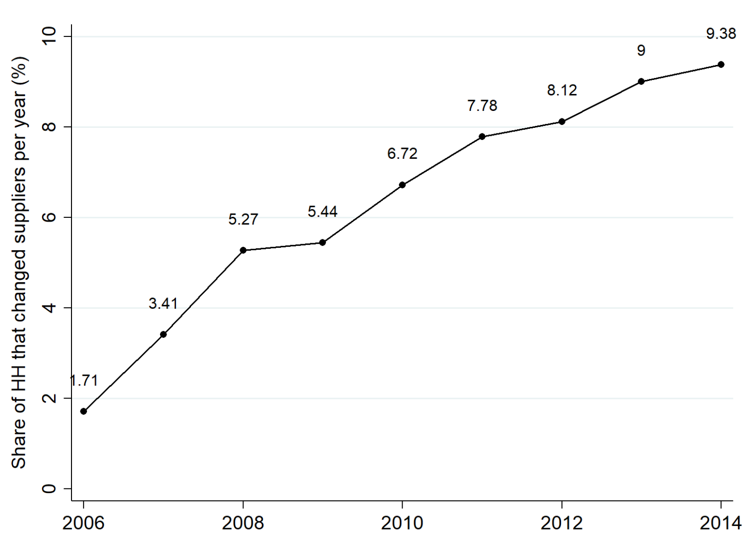


Note: The figure presents the average incumbent base tariffs for a household with yearly electricity consumption 3,500 kWh. Data are from the German Federal Network Agency (Bundesnetzagentur, 2015).

market entry in Germany can be characterized as fairly easy, since it follows a regulated, non-discriminatory procedure, and involves low sunk costs and risks.

Consumers generally have an annual contract with their supplier. Most of these contracts run for one calendar year and are automatically renewed if not cancelled by the consumer in due time. While the cancellation period for the incumbent base tariff is only two weeks by law, it is usually several months for the other tariffs and thus locked-in costs are higher once a consumer has decided to switch. While the most important characteristic of electricity contracts is the total annual electricity bill a household has to pay, many contracts include special rules that either provide security for the retailer or the customer. For example, according to a market report by the German regulatory authority Bundesnetzagentur (2013, p. 150), the average minimum contract period is 10 months (providing security for the retailer), while the average period of stable guaranteed prices is 13 months (providing security for the customer). In recent years, most households who consider changing their supplier visit an online price comparison platform. The largest platforms are Verivox, Toptarif, and Check24. Besides Toptarif, our database also covers all search activities conducted on several other well-known online price comparison platforms including Stromtipp.de (power hint), Energie-verbraucherportal.de (energy consumption portal) and mut-zum-wechseln.de (courage-to-change). Verivox started to

Figure 2: Switching rates of household consumers in electricity retail in Germany



Note: Data on supplier changes are obtained from Germany's regulatory authority (Bundesnetzagentur, 2015), data on the number of German households are from the German Federal Statistical Office.

provide search services in electricity in 2000, Toptarif in 2007, and Check24 in 2008. Despite these fairly recent starts of search services, according to a survey 80% of the switchers searched online for alternative providers in 2011.

Due to the creation of online platforms, the share of switching customers is growing in recent years (see Figure 2) as online price comparison platforms have significantly disburdened the costs of searching for cheaper providers. A comparison portal requires a consumer to enter all relevant details (zip code, how much electricity in KWh is consumed on a yearly basis and whether the contract is for private use or not). Then there are several options to choose from, such as whether or not to only consider "green" electricity, whether or not prices are guaranteed throughout the year[7] and whether or not there is a one-off bonus. The platform then lists the "personalized" price offers of all providers that are active in the indicated zip code with the cheapest price ranked first and immediately informs the consumer how much he saves over the year compared to the incumbent's baseline price. An example of a typical screenshot of an online platform is given in the Appendix (see Figure A1). Thus, the search process costs a little bit of time and effort,

[6]Switching rates include not only switching from the incumbent to an entrant but also to a large extent switching among entrant suppliers.

[7]If prices are not guaranteed consumers have an extraordinary termination right if their retailer increases prices.

but for all consumers who are used to online shopping, the search costs seem relatively small compared to the potential savings to be obtained (see Figure 3)). It is also clear that as the online price will only be observed by consumers who search, given the search technology it is relatively easy for the incumbent to screen consumers and set an online price that is different from the baseline price.

Electricity is a relatively homogeneous product, with "green" production technologies such as wind, solar or run-of-river plants, being the main differentiating feature. As most providers feature electricity from "green" and "non-green" generation sources, it should not matter for the end-consumer who delivers the product they chose. Moreover, the incumbent provider has the legal obligation to guarantee a continuous provision of electricity to the customer if an entrant fails to deliver at no additional fees to the consumers. Therefore, brand loyalty to the incumbent should, in principle, not play a role. Despite the obligation, not all German consumers may be aware, however, of this safety net and brand loyalty may play a role in practice.[8]

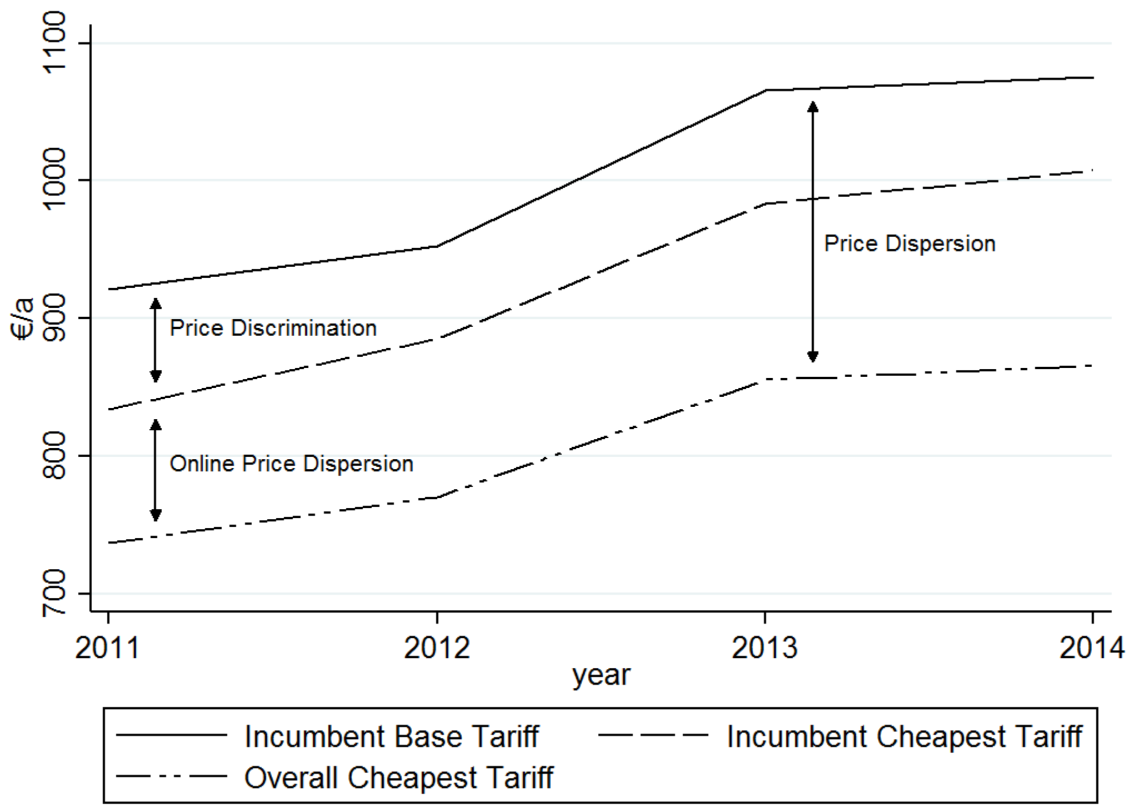
Not only have search costs declined over time, switching costs have also been significantly reduced as switching is now an automated process and conducted entirely by the new provider a consumer chooses. The new supplier automatically arranges all switching activities for new customers, such as unsubscribing from the old supplier, registration, at no switching fees. Moreover, by law, the cancellation period for the incumbents' base tariff is only two weeks.[9]

Entrants generally undercut the baseline tariff of the incumbent. On the other hand, the incumbent usually offers a cheaper tariff than its baseline tariff, yet offers a higher online price than the cheapest entrant tariff. Figure 3 shows that there are considerable price differences between the incumbent baseline tariff (p^{IB}), the incumbent cheapest tariff (p^{IC}), and the overall cheapest tariff provided by entrants (p^{OC})). As a result, we observe three forms of price dispersion: (i) Overall *price dispersion* ($p^{IB} - p^{OC}$), which is the overall range of the price distribution as the difference between the incumbent baseline tariff as the upper bound and the overall cheapest tariff as the lower bound; (ii) *Price discrimination* by the incumbent is measured by the difference between the incumbent baseline tariff and the incumbent cheaper online tariff ($p^{IB} - p^{IC}$), which displays the

[8]The German incumbent is one of the most reliable electricity providers worldwide. For example, the average SAIDI (System Average Interruption Duration Index) measuring the average supply interruption per household, was only 12.3 minutes in Germany in 2014, one of the best values worldwide. Up until now, there have been only two insolvencies of German incumbents (Stadtwerke Gera and Stadtwerke Wanzleben, both in 2014). Consumers were not affected, as electricity supply continued during the insolvency procedures. Hortacsu et al. (2015) mention the possibility that customers may believe that the incumbent supplier may exhibit a higher supply security although this is in fact not true.

[9]In many other states the switching process is comparable to the one in Germany's retail electricity. E.g. studying the UK market, (Giulietti et al., 2014, p. 561) argue that "search is perceived by consumers as being significantly more difficult than switching." MENTION HORTACU FOR TEXAS. OTHER COUNTRIES?

Figure 3: Average tariffs (€/year for 3,500 kWh)



Note: Calculation based on data from ene't.

maximum potential saving for consumers whose valuation of the incumbent brand exceeds the *loyalty premium*; And (iii) online price dispersion, which we call a *"loyalty premium"*, representing the difference between the incumbent cheaper tariff and the cheapest entrant tariff ($p^{IC} - p^{OC}$).[10]

3 A simple search theoretic model

We consider the following simple model of a liberalized market for a homogeneous product where an incumbent firm competes with entrants. The model closely follows the institutional details described above. All consumers observe the regular (baseline) price p^{IB} of the incumbent at no additional cost. There is an online price comparison website consumers can consult at a search cost s, which is distributed according to a distribution function $F(s; z)$, with support $[0, 1]$, where we use z to represent exogenous parameters that determine the shape of the search cost distribution. (In the empirical part, z are the instruments that are exogenous to the price differences across regions, but that do affect differences in search behaviour across regions). Comparing different municipalities, one can think of z as the fraction of consumers with high speed broadband access, the fraction of consumers who regularly buy online, etc. At the website, consumers will see potentially many prices, but (in line with the data we have) we are only interested in two of them: the price p^{OC} of the overall cheapest firm (usually an entrant) and the cheapest (online) price p^{IC} of the incumbent. We assume that once the consumer is on the website, he compares prices without additional search cost.

Apart from their search cost, consumers also have some brand loyality to the incumbent. Brand loyality will also differ between individuals. Not to complicate the analysis too much (and not to deal with two different distributions for search cost and brand loyality), we assume that brand loyality is proportional to search cost and say that the brand loyality of a consumer with search cost s is denoted by θs. One way to interpret this would be that consumers with higher search costs will be older and more wealthy consumers who do not want to risk their stable delivery of electricity by switching.[11] Once a consumer with search cost s is online and observes both prices p^{OC} and p^{IC} then he will continue to buy from the incumbent if $p^{IC} - \theta s < p^{OC}$.

The sequence of actions is as follows. In the first stage, the incumbent and entrant choose p^{IB}, p^{IC} and p^{OC} simultaneously. At the beginning of the second stage, consumers only observe p^{IB} and decide whether or not to search (given their expectations of the online prices). If they do not search they buy at p^{IB} from the incumbent. If they do

[10] An illustration of the longer-term development of the incumbents base tariff is provided in Figure 1 in the Appendix.

[11] Even though the entrant is also a stable supplier, there still may be a psychological element that the incumbent in Europe has always been a stable, high quality supplier.

search, then they buy where it is best for them. We use perfect Bayesian equilibrium with passive beliefs as our solution concept. In particular, if a consumer observes an unexpected price p^{IB} (different from the equilibrium level), then he wil continue to believe that p^{IC} and p^{OC} are at their equilibrium levels.

We will be looking for a perfect Bayesian equilibrium where the low search cost consumers search online and the high search cost consumers stay with the baseline price of the incumbent. Of the consumers that search online, the ones with very low brand loyality (and thus also low search cost) buy from the entrant, while other online consumers buy from the incumbent at his cheapest (online) price. The *ex ante* utility of a consumer with search cost s to buy at the three different prices is given by $v - p^{OC} - s, v - p^{IC} + (\theta - 1)s$ and $v - p^{IB} + \theta s$ if he buys from the entrant, the online price of the incumbent and the baseline price of the incumbent, respectively. In such an equilibrium, the cut-off values for the search costs are $\widehat{s}_1 = \left(p^{IC} - p^{OC}\right)/\theta$ and $\widehat{s}_2 = \left(p^{IB} - p^{IC^e}\right)$ such that all consumers with $s < \widehat{s}_1$ buy from the entrant, all consumers with $\widehat{s}_1 < s < \widehat{s}_2$ buy at the web price of the incumbent and all consumers with $s > \widehat{s}_2$ buy at the regular price of the incumbent. Note that in the definition of $\widehat{s}_2$ we have p^{IC^e}, the online price of the incumbent consumers expect to find if they search, but before actually engaging in search, and not the realized online price of the incumbent: when deciding whether or not to search, the consumer does not know yet which online price he will observe and therefore the expectation is relevant.[12] Note also that we do not have to deal with expected prices in the definition of $\widehat{s}_1$ because these are consumers that are indifferent between buying at one or the other online prices, which implies they are already at the website, incurred the search cost, and observe both prices.

In the theoretical part we assume, without loss of generality, that the firms have no cost of supplying. The equilibrium prices we derive can therefore be interpreted as the margins firms make. Given this division of consumers, the respective profits of the entrant and incumbent are as follows:

$$\pi_E = F(\widehat{s}_1; z)p^{OC} = F\left(\frac{p^{IC} - p^{OC}}{\theta}; z\right)p^{OC}$$

and

$$\begin{aligned}\pi_I &= \left[F(\widehat{s}_2; z) - F(\widehat{s}_1; z)\right]p^{IC} + (1 - F(\widehat{s}_2; z))p^{IB} \\ &= \left[F(p^{IB} - p^{IC^e}; z) - F\left(\frac{p^{IC} - p^{OC}}{\theta}; z\right)\right]p^{IC} + (1 - F(p^{IB} - p^{IC^e}); z)p^{IB}.\end{aligned}$$

[12]To determine the equilibrium value of p^{IC} and p^{IB} it is important that a consumer will not observe a deviation from p^{IC} before the decision whether or not to search is made.

This yields the following F.O.C.´s (evaluated at the equilibrium where $p^{IC^e} = p^{IC}$) for the entrant and the incumbent, respectively:

$$F\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)-f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)\frac{p^{OC}}{\theta}=0, \tag{1}$$

$$F(p^{IB}-p^{IC};z)-F\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)-f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)\frac{p^{IC}}{\theta}=0, \tag{2}$$

and

$$-f(p^{IB}-p^{IC};z)(p^{IB}-p^{IC})+(1-F(p^{IB}-p^{IC};z))=0, \tag{3}$$

where $f(.)$ is the density function that is associated with $F(.)$.

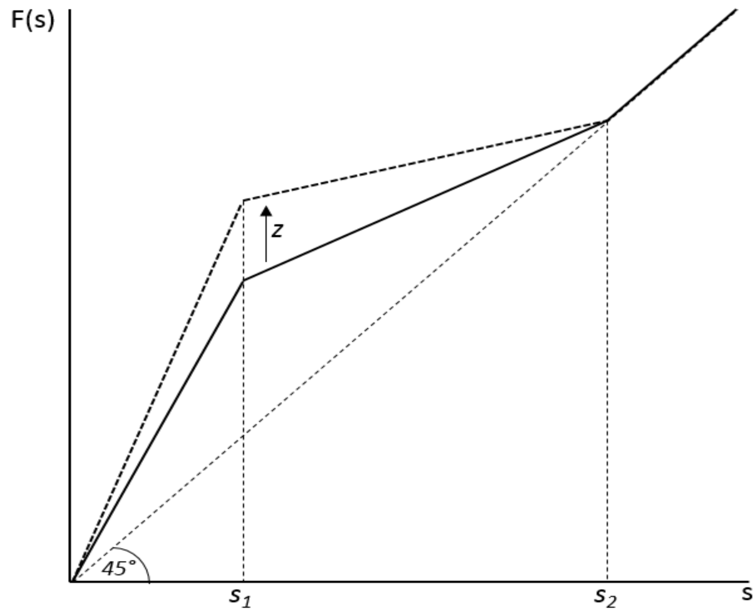
In this model, price discrimination by the incumbent is measured by $p^{IB}-p^{IC}$, while (online) price dispersion (i.e. the "loyalty premium") is modelled by $p^{IB}-p^{OC}$ and $p^{IC}-p^{OC}$, respectively. The fraction of consumers that actively searches is given by $F(p^{IB}-p^{IC};z)$. For a given x, these three FOCs determine the equilibrium values of p^{IB}, p^{IC} and p^{OC} and the associated levels of price discrimination and price dispersion and the faction of active searchers. To explain differences in the observed level of price discrimination and price dispersion and the faction of active searchers between regions we have to see how the equilibrium level of prices change with changes in z.

It is clear from (1)–(3) that depending on how the search cost distribution changes in z, different patterns are possible. The next result provides a general statement on the conditions affecting price discrimination and online price dispersion

Proposition 1 *Price discrimination increases if, and only if, the inverse hazard rate evaluated at the equilibrium values* $\frac{1-F(p^{IB^*}-p^{IC^*};z)}{f(p^{IB^*}-p^{IC^*};z)}$ *is increasing in* z. *Moreover,* p^{OC} *and online price dispersion are positively related to* p^{IC} *if the density functions are non-increasing, i.e.,* $\partial f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)/\partial(p^{IC}-p^{OC})\leq 0$. *Finally, online dispersion and price discrimination are linked by* $1+f(p^{IB}-p^{IC};z)(p^{IB}-p^{IC})=f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)\left(\frac{p^{IC}+p^{OC}}{\theta}\right)$.

The economic intuition behind the result on price discrimination is clear: for a given value of p^{IC} the incumbent faces a trade-off in its decision whether or not to increase p^{IB}. Raising p^{IB} increases the profits over all consumers $1-F(p^{IB^*}-p^{IC^*};z)$ who will stay on the baseline tariff; on the other hand, a fraction proportional to the density $f(p^{IB^*}-p^{IC^*};z)$ will decide to search. At the margin, the ones that decide to search are the ones that will eventually buy at the incumbent's online price p^{IC^*} as the marginal consumer is the one with a higher search and switching cost. The incumbent will loose $p^{IB^*}-p^{IC^*}$ per (marginal) consumer who searches. If the inverse hazard rate is increasing in z, relatively more consumers will stay on the base line tariff if z increases making price discrimination more profitable.

Figure 4: A piece-wise linear search cost distribution



Note: An increase in x shifts the piece-wise linear search cost distribution in a way that there is more mass of consumers with lower search costs.

Also, to understand online price dispersion, if p^{IC} increases, then there is a larger potential demand for the entrant and, under "normal" demand conditions, he should increase his price, but not to the full extent (thereby also increasing sales). Online price dispersion and price discrimination are linked by the decision of the incumbent how to set its online price.

To investigate the determinants of price dispersion and the fraction of active searchers, it is useful to analyze a specific form of a search cost distribution. There are two families of search cost distributions that povide reasonably simple expressions for the FOCs: an exponential search cost distribution according to which $F(s; x) = 1 - e^{-zs}$ on $s \in [0, \infty)$ or a piece-wise linear search cost distribution

$$F(s) = \begin{cases} zs & \text{for } s < \widetilde{s}_1 \\ \alpha + \beta s & \text{for } \widetilde{s}_1 \leq \ s < \widetilde{s}_2 \\ s & \text{for } s \geq \widetilde{s}_2, \end{cases}$$

where to have a proper piece-wise linear distribution function, we should have $\alpha = \frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}$, $\beta = \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}$ and $\widetilde{s}_2 > \widetilde{s}_1$ and $z > 0$. If $z = 1$, we have the uniform distribution. Straightforward calculations show that for the exponential distribution, $p^{IB} - p^{IC} = 1/z$

and that $F(p^{IB} - p^{IC}) = 1 - e^{-1}$, which is independent of z. Thus, with an exponential distribution we cannot explain changes in the fraction of searchers between regions. That is why we adopt the piece-wise linear distribution in the rest of this section and we focus on parameter values such that $\widehat{s}_1 < \widetilde{s}_1 < \widehat{s}_2 < \widetilde{s}_2$, i.e., the consumer that is indifferent between two online offers is in the first interval of the search cost distribution, while the consumer that is indifferent between searching and not searching is in the second interval of the search cost distribution.[13]

The piece-wise linear formulation allows us to have different ways in which the search cost distribution may tend to have more consumers with lower search cost. The simplest formulation is in terms of z: an increase in z unambiguously leads the search cost distribution to have more mass on consumers with lower search cost at the expense of consumers with intermediate search cost. An increase in z is the only way to model that the fraction of consumers with very low search cost increases. An increase in $\widetilde{s}_1$ and/or in $\widetilde{s}_2$ also leads the search cost distribution to have more mass on consumers with lower search cost, but (if and) only if $z > 1$. In addition, increases in $\widetilde{s}_1$ or in $\widetilde{s}_2$ leave the fraction of consumers with the very lowest search cost unchanged and only increases the fraction of consumers with more intermediate search cost.

In the empirical part of the paper, we use an instrumental variable approach and use the fraction of people with fast broadband access and the fraction of households where the household head is below 40 as instruments for the (potentially endogenous) fraction of households that search for lower prices using one of the online platforms. In our theoretical model, we interpret these instruments in terms of a larger fraction of households having very low search cost. Therefore, we will focus here on the comparative statics with respect to z and discuss the other comparative static results in the Appendix.

It is important to note that having relatively more consumers with lower search cost does not automatically imply that there will be more active searchers. The number of active searchers $F(p^{IB^*} - p^{IC^e})$ is also endogeneously determined by the equilibrium prices (and the expected prices online). To determine the number of active searchers, we first determine the level of price discrimination $p^{IB^*} - p^{IC^e}$. Using (3) it is easy to see that for the case where the search cost distribution is piece-wise linear

$$p^{IB} - p^{IC} = \frac{\widetilde{s}_2 - \widetilde{s}_1 - (z-1)\widetilde{s}_1\widetilde{s}_2 - (\widetilde{s}_2 - z\widetilde{s}_1)\left(p^{IB} - p^{IC}\right)}{(\widetilde{s}_2 - z\widetilde{s}_1)},$$

so that the equilibrium level of price discrimination equals

$$p^{IB^*} - p^{IC^*} = \frac{\widetilde{s}_2 - \widetilde{s}_1 - (z-1)\widetilde{s}_1\widetilde{s}_2}{2\left(\widetilde{s}_2 - z\widetilde{s}_1\right)} \tag{4}$$

[13] The qualitative results continue to hold if $p^{IB} - p^{IC} > \widetilde{s}_2$ although the specfic formulaes will be different.

and thus that the equilibrium fraction of online searchers equals

$$F(p^{IB^*} - p^{IC^*}) = \frac{(z-1)\widetilde{s}_1\widetilde{s}_2 + \widetilde{s}_2 - \widetilde{s}_1}{2\left(\widetilde{s}_2 - \widetilde{s}_1\right)}.$$

Finally, applying the piece-wise linear search cost distribution to (1) and (2), it is easy to see that that the relation between the equilibrium online prices is given by

$$p^{OC^*} = \frac{1}{2} p^{IC^*}$$

so that

$$p^{IC^*} = \frac{\widetilde{s}_2 - \widetilde{s}_1 + (x-1)\widetilde{s}_1\widetilde{s}_2}{3x\left(\widetilde{s}_2 - \widetilde{s}_1\right)}\theta,$$

which implies that

$$p^{OC^*} = \frac{\widetilde{s}_2 - \widetilde{s}_1 + (x-1)\widetilde{s}_1\widetilde{s}_2}{6x\left(\widetilde{s}_2 - \widetilde{s}_1\right)}\theta$$

and

$$p^{IB^*} = \frac{\widetilde{s}_2 - \widetilde{s}_1 + (x-1)\widetilde{s}_1\widetilde{s}_2}{3x\left(\widetilde{s}_2 - \widetilde{s}_1\right)}\theta + \frac{\widetilde{s}_2 - \widetilde{s}_1 - (x-1)\widetilde{s}_1\widetilde{s}_2}{2\left(\widetilde{s}_2 - x\widetilde{s}_1\right)}.$$

Using the previous Proposition and the fact that for a piece-wise linear distribution $f' = 0$ in the interior of the intervals, online equilibrium prices always change in the same direction and the level of online price dispersion $p^{IC} - p^{OC}$ positively correlates with both prices.

The above expressions hold true as long as $\widehat{s}_1 < \widetilde{s}_1 < \widehat{s}_2 < \widetilde{s}_2$. Using the expressions for the different prices, and therefore for $\widehat{s}_1$ and $\widehat{s}_2$, this implies that the parameters should satisfy

$$\frac{\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_1(5\widetilde{s}_2 - 6\widetilde{s}_1)} < z < \frac{\widetilde{s}_2^2 - (\widetilde{s}_2 - \widetilde{s}_1)(1 - \widetilde{s}_2)}{\widetilde{s}_1\widetilde{s}_2}. \quad (5)$$[14]

In the two propositions below we formulate the comparative statics properties of our model using the piece-wise linear search cost distribution. The first Proposition that follows states the results in terms of absolute price levels, while the next Proposition states the results in terms of price differences (price discrimination and dispersion).

Proposition 2 *(price levels) If (5) holds, then an increase in the fraction of online searchers* $F(p^{IB^*} - p^{IC^*})$*, initiated by an increase in* z*, coincides with a decrease in online prices* p^{OC^*} *and* p^{IC^*} *if, and only if* $\widetilde{s}_2 - \widetilde{s}_1 > \widetilde{s}_2\widetilde{s}_1$*, while it coincides with an increase in the baseline price* p^{IB^*} *if* θ *is small enough,* z *is large enough, or* $\widetilde{s}_2 - \widetilde{s}_1$ *is small enough.*

[14]Note that these conditions are independent of θ. As an example, if $z = 1$ these inequalities reduce to $\frac{1}{6} < \widetilde{s}_1 < \frac{1}{2}$ and $\widetilde{s}_2 > \frac{1}{2}$. Or, when $\widetilde{s}_1 = \frac{1}{5}$ and $\widetilde{s}_2 = \frac{3}{5}$, $\frac{7}{9} < z < \frac{5}{3}$ has to hold.

The conditions mentioned in the Proposition should be interpreted as follows. First, if $\widetilde{s}_2 - \widetilde{s}_1 < \widetilde{s}_2\widetilde{s}_1$ can be interpreted as that $\widetilde{s}_1$ is relatively close to $\widetilde{s}_2$. In this case, if z increases there are relatively many consumers that will search online that are relatively loyal to the incumbent (have higher switching cost) and they will continue to buy from the incumbent (but at its online price). This gives the incumbent more market power online, resulting in higher prices. Second, the result for the incumbent's baseline price can be understaood as follows. If θ is relatively small, then there is fierce competition online and the more consumers search online, the more the incumbent wants to extract surplus from the consumers with high search and switching cost. On the other hand, for the piece-wise linear search cost distribution we specified, if z is relatively large, or $\widetilde{s}_2 - \widetilde{s}_1$ is relatively small, then there are few consumers that are indifferent between staying with the incumbent or searching online, which again gives the incumbent an incentive to increase its baseline price.[15]

In line with our presentation of the empirical results, we present the results on price differences separately.

Proposition 3 *(Price discrimination and dispersion) If (5) holds, then an increase in the fraction of online searchers* $F(p^{IB^*} - p^{IC^*})$*, initiated by an increase in* z*, coincides with (i) an increase in price discrimination* $p^{IB^*} - p^{IC^*}$ *and (ii) a decrease in online price dispersion* $p^{IC^*} - p^{OC^*}$*, if and only if* $\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2 > 0$ *and (iii) an increase in price dispersion* $p^{IB^*} - p^{OC^*}$ *if* θ *is small enough,* x *is large enough, or* $\widetilde{s}_2 - \widetilde{s}_1$ *is small enough.*

These results can be understood along similar lines as above. Proposiion 1 already stated under general conditions that online price dispersion is correlated to the incumbent's online price. That Proposition also indicated that price discrimination may increase if the inverse hazard condition is satisfied, which is the case for the piece-wise linear distribution. Finally, overall price dispersion is closely related to the incumbent's baseline price and in Proposition 2 we have explained the conditions under which that price is increasing.

Figure 4 depicts how the different prices change as a function of x when $\widetilde{s}_2 = 3/5$ and $\widetilde{s}_1 = 1/5$ and $\theta = 2/5$. As $\frac{\partial F(p^{IB^*} - p^{IC^*})}{\partial z}$ is a constant positive number, this figure can also be interpreted how prices are linked to the fraction of searchers. One can see that for these parameter values, the incumbent's baseline price is increasing in the fraction of searchers, whereas the other two prices are decreasing, resulting in more price discrimination and price dispersion, while online price dispersion is decreasing.

INSERT FIGURE 4

[15] One can also inquire into how the average price depends on search intensity. The weighted average price is given by $(1 - (F(\widehat{s}_2))p^{IB^*} + (F(\widehat{s}_2) - F(\widehat{s}_1))p^{IC^*} + F(\widehat{s}_1)p^{OC^*} = p^{IB^*} - F(\widehat{s}_2)(p^{IB^*} - p^{IC^*}) - F(\widehat{s}_1)(p^{IC^*} - p^{OC^*})$. One can find parameter regions where this average price is increasing in z (and thus, increasing in the fraction of consumers searching).

Thus, if markets (regions) mainly differ in the fraction of low search cost consumers (measured here by a difference in z), then an increase in online search is accompanied by price discrimination by the incumbent, and if the search costs in the population are sufficiently spread ($\widetilde{s}_2 - \widetilde{s}_1$ is large enough compared to $\widetilde{s}_1 \widetilde{s}_2$) by less online price dispersion and more overall price dispersion. Of course, other pricing patterns are also possible for different parameter values.

4 Empirical Identification of the Effects of Consumer Search on Pricing Strategies

Before we present the data and our empirical results it is important to explain our identification strategy. The relationship we are interested in can be described by a model of the form:

$$\ln(y_{it}) = \beta \ln(\mu_{it}) + \gamma \mathbf{x}_{it} + \delta_i + \eta_t + \epsilon_{it}, \tag{6}$$

where the subscripts i and t, therefore, indicate ZIP codes and years, respectively. y either denotes an electricity tariff (p^{IB}, p^{IC}, p^{OC}) or a dispersion measure ($p^{IB} - p^{OC}$, $p^{IB} - p^{IC}$, $p^{IC} - p^{OC}$) and is a function of consumer search intensity (μ), a set of control variables ($\mathbf{x}$) including (logged) regional electricity costs, the number of regional electricity retailers and some regional household characteristics like income and average household size. Both, search intensity and prices (but also costs), exhibit substantial spatial and temporal variation, which enables us to control for unobserved heterogeneity by ZIP-code (δ_i) and year (η_t) fixed effects that may co-influence pricing strategies. ϵ is the random error term.

Since we only observe consumer search at the online portals represented in our sample, but all consumer search activity (i.e., we do not observe search activity on all existing platforms), we estimate constant elasticities in a log-log relationship. That is, we include the dependent variables (i.e. tariffs and dispersion measures) as well as search intensity in logs indicating by how much a percentage change in search impacts the dependent variable in percentage terms. Assuming that search patterns at other major comparison websites is not different from search at the platforms that we observe in our data, the elasticity estimate allows us to make inferences about the whole market.[16] Hence, our parameter

[16] The assumption that the searching patterns we observe are similar across platforms is reasonable: We also have cross-sectional information on consumer search at Verivox – another large price comparison platform in Germany – for the year 2014. For that year we checked the spatial correlation of search intensity between Verivox and the platforms we use here and find a correlation coefficient of 85%, indicating that search at Verivox does not seem to be very different from search observed in our dataset.

of interest β reflects the percentage change in pricing behavior for a one percent change in search activity.[17]

In order to get causal inference for β we have to deal with the potentially simultaneous relationship between prices and search intensity: consumers may search more if they expect more dispersed electricity prices. To circumvent the endogeneity issue we instrument for search ($\ln(\mu)$) using two instruments: i) The share of young households in a zip code measured as the share of households with a household head below the age of 40 (*U40*) and ii) the regional coverage of broadband internet (*BBC*) measured as the share of households in a zip code with at least 16Mbit/s internet speed. Both instruments are assumed to affect search costs but not prices (other than through search): younger people are probably more familiar with the internet in general and with online shopping in particular. Fast internet makes onine shopping more convenient. Using these instruments we estimate the following linear projection of μ

$$\ln(\mu_{it}) = \alpha \mathbf{z}_{it} + \gamma^{FS}\mathbf{x}_{it} + \delta_i^{FS} + \eta_t^{FS} + u_{it}, \tag{7}$$

where vector $\mathbf{z}$ consists of our two instruments *U40* and *BBC*. The superscript *FS* indicates that the parameters are from the first-stage estimation.

The critical assumption is that both instruments affect pricing strategies only through the search channel. However, as young people may also have less income which in turn may affect pricing decisions, e.g. because the retailer has a higher risk that consumers may not be able to pay for their electricity bill, we also control for income characteristics. As income is included in $\mathbf{x}$ we assume that the instrument vector $\mathbf{z}$ is uncorrelated with the error term ϵ such that

$$E(\epsilon_{it}\mathbf{z}_{it}|\delta_i, \eta_t, \mathbf{x}_{it}) = 0.$$

Plugging the first-stage prediction of search (i.e. $\widehat{\ln(\mu)}$) into Equation (6) should yield unbiased estimates for β.

5 Data and Variables

We utilize novel and unique panel data at the German zip code level (8,224 zip codes) for the annual period 2011–2014, that allow for an empirical investigation of the effects of consumer search on retail pricing decisions. A key asset in the data is the information about online search queries at major price comparison portals which enables us to construct a direct measure of consumer search. Notably, to the best of our knowledge, this is the first paper using a direct measure of search in a panel data context. Another

[17] However, the estimations are fully robust if estimated in levels as we show in section 6.

particular advantage over other empirical studies is that we do not have to make assumptions about market delineation, as consumers at retail electricity markets can only choose among electricity suppliers providing for their available local address.

Our data stem from four sources. From *ene't*, a German software and data provider for the electricity industry, we received detailed data on consumer search activities, retail electricity tariffs and cost components. The marketing company *Acxiom* provided data on structural household characteristics in Germany. While the ene't data vary monthly (the incumbent cheapest tariff varies yearly), Acxiom data are given yearly. For matching purposes, we aggregate all data to an annual frequency. This corresponds well with the length of a typical electricity contract. We also use data from the *European Energy Exchange (EEX)* to get a proxy for the purchase costs of wholesale electricity (as a component of utilities' total costs). Moreover, we use data on regional broadband internet coverage from *breitbandatlas.de*. Table 1 provides the summary statistics of the variables in our regressions.

Price Data

ene't provided all tariff data (including 19% VAT) for p^{IB}, p^{IC} and p^{OC}. In the empirical application we focus on a typical two-person household with an annual consumption level of 3,500 kWh. This is the default consumption level suggested by all major price comparison sites for a two-person household.[18]

Table 1 shows that, on average, a sample household with 3,500 kWh annual consumption of electricity pays around 1,004 EUR per year if it stays with the default tariff by the incumbent. The incumbent's cheaper tariff is around 8% lower at 928 EUR, while the overall cheapest entrant tariff is around 807 EUR (which is 20% cheaper than the incumbent default tariff). All prices exhibit substantial variation in our observation period regarding their spacial and time dimensions. Figure A2 in the Appendix exemplarily shows the spatial distribution of prices for the year 2012. From the figure it can be seen that prices are generally higher in eastern and northern Germany because of the many windplants in these areas which cause an increase of grid charges. Figure A4 in the Appendix shows the spatial cost distribution. We consider these patterns in the empirical analysis by controlling for variations in costs at the zip code level (the data we use to compute costs are described in the subsection on the control variables below) and we also add zip cost fixed effects. Notably, the price data also provide substantial temporal variation. This can be seen examplary for the incumbents tariffs in Figure A3 in the

[18]3,500 kWh is also the household consumption level that is typically applied by other agencies (e.g. (Bundesnetzagentur, 2015) for comparing retail tariffs. ene't also provided tariff data for other annual consumption levels (2,000 kWh and 4,000 kWh), however only for p^{IB} and p^{OC} (not for p^{IC}). Regression estimates using p^{IB} and p^{OC} as well as $p^{IB} - p^{OC}$ for alternative consumption levels yield robust results.

Appendix.

Consumer Search Activity

ene't provided data on individual online search queries enabling us to construct a dataset on regional consumer search intensity. The database covers detailed information on all search queries conducted on several well-known online price comparison platforms including Toptarif.de (top-tariff), Stromtipp.de (power-hint), Energie-verbraucherportal.de (energy-consumer-portal), and mut-zum-wechseln.de (courage-to-change), of which Toptarif.de is by far the biggest platform.[19] For each query we observe a timestamp, the zip code for which the offered electricity tariffs are requested, the (expected) yearly consumption entered into the search mask, the type of search query (household or industrial customer), consumer preferences (e.g. whether a consumer only wants to get "green" energy tariffs displayed), as well as a search session ID indicating the order of the queries of the same consumer.[20] In sum, we have information on 35,855,071 search queries from 17,302,530 search sessions of which 96.7% (i.e. 16,778,214 sessions) are conducted by households and the remaining 3.3% (i.e. 524,316 sessions) by industrial customers. As many searchers conduct several search queries within a search session (e.g. comparing prices for different levels of consumption) we focus on the number of search sessions per year and zip code rather than on the absolute number of search queries and refer to a consumer conducting a search session as being fully informed regardless of the depth of the search activity.[21] Since our focus is on household consumers, we disregard search from industrial consumers. Furthermore, we exclude 551,256 search sessions, which exclusively consider eco-label (i.e. "green") certified tariffs.[22] Those searches are most likely not predominantly price driven but related to product differentiation and, on average, €152 more expensive than the cheapest tariff.

Using this database, we construct a measure of consumer information as the number of search sessions within a ZIP code per year divided by the number of households:

$$\mu_{it} = \frac{SearchSessions_{it}}{Households_{it}}.$$

[19] Toptarif is one of the three major electricity and gas price comparison websites along with Verivox and Check24. It was acquired by Verivox in July 2014 but continues to operate as Toptarif.

[20] We are not able to observe actual switching, because clicking on a certain supplier tariff at the online comparison website redirects the searcher to a website where the switch may be finalized. This limitation is common to online data (see Koulayev, 2014). Yet, switching requires searching, so the impact of consumer search on price strategies seems to be consistently estimable. Brynjolfsson and Smith (2001) confirm this and find that factors that drive clicks are reasonable and unbiased indicators of sales in their study on online book purchases.

[21] It should also be noted that a search session only contains the current search activity of an individual household and we cannot distinguish whether the same household starts a new search session on another day. Therefore, we treat each search session as conducted by an individual household.

[22] Nevertheless, our results are fully robust to the inclusion of eco-label searches.

Since we observe some extreme outliers in some zip codes, apparently resulting from price comparing software botsEor data crawling researchers, we truncate 2% of the upper bound of the sample distribution of our consumer information measure.[23]

At the mean, 9.2% of households within a zip code search for retail tariffs at one of our sample comparison websites, whereas there is sufficient variation ranging from 0% to 35.4%. An illustration of the spatial variation in search intensity is given in figure A5 in the Appendix.

Instrumental Variables

The data on the share of young households, i.e. households with a household head younger than 40 years (*U40*), are obtained from *Acxiom* at a zip code-year resolution. Data on local broadband internet coverage (*BBC*), i.e. the share of households with access to internet speed of 16 MBit/s or higher, are obtained from textitBreitbandatlas.de with the same resolution. The *BBC* data are on the municipal level and we aggregate them to the zip code level to match them with our other data.

Control Variables

We compute a variable reflecting the retailers' costs in order to control for regional cost differences. Detailed data on cost components are primarily obtained from textitene't and include, for example, grid charges, concession fees, renewable energy surcharges ("EEG Umlage"), CHP surcharges ("KWK Umlage") and electricity taxes. Grid charges are paid by the electricity provider to the respective system operator and, thus, vary across grid areas (i.e. clusters of zip codes) and time as they are adjusted annually. The concession fee has to be paid by the system operator to the respective municipality for the right to install and operate electric cables on public roads. Hence, the concession fees vary at the municipality level and also over years. The remaining cost components only vary over time but not spatially. Moreover, we also add the one-year ahead future prices at the EEX spot market to our cost variable to proxy for the costs of wholesale electricity as this presents the standard purchasing strategy for retailers.[24]

To measure spatial competition we use the number of electricity retail suppliers within a zip code, as provided by *ene't*. The number of competitors in a zip code varies between 55 and 198 in our observation period.[25]

[23] Figure A6 in the Appendix provides a histogram on consumer information before and after trimming the data.

[24] Even if retailers purchase electricity from other channels than via the power exchange (e.g. bilateral contracts, OTC, etc.), the spot price still represents the opportunity cost.

[25] These numbers may seem high but correspond well with Duso and Szücs (2017) and Bundesnetzagentur (2015) and may be the consequence of low entry barriers in the electricity retail market in Germany.

Table 1: Summary statistics

Variable	Unit, data source	Mean	SD	Min	Max	Obs
Dependent variables						
Incumbent Base (p^{IB})	€/a, *ene't*	1.004	78.1	791	1,204	29,090
Incumbent Cheapest (p^{IC})	€/a, *ene't*	928	86.4	697	1,136	29,090
Overall Cheapest (p^{OC})	€/a, *ene't*	808	59.2	657	903	29,090
Price Dispersion ($p^{IB} - p^{OC}$)	€/a, *ene't*	197	38.8	53.4	354	29,090
Price Discrimination ($p^{IB} - p^{IC}$)	€/a, *ene't*	76.4	41.9	0	282	29,090
Onlince price dispersion ($p^{IC} - p^{OC}$)	€/a, *ene't*	121	46.3	0	259	29,090
Endogenous variable						
Consumer search intensity (μ)	%, *ene't*	9.20	6.4	0	35.4	29,090
Instruments						
Head of HH below age of 40 ($U40$)	%, *Acxiom*	24.7	5.1	7.69	55.1	29,090
Broadband Internet Coverage (BBC)	%, *breitbandatlas.de*	64.0	33.5	0	100	29,090
Control variables						
Costs (C)	€/a, *ene't & EEX*	682	42.4	551	845	29,090
Competitors	#, *ene't*	133	24.8	54.6	198	29,090
Average HH size	%, *Acxiom*	2.11	0.2	1.52	2.54	29,090
Income<25k €/year	%, *Axciom*	0.39	0.1	0.02	0.83	29,090

Note: "Obs" are zip code-year observations. €/a refers to an annual electricity consumption of 3.5 MWh.

Other control variables refer to structural household characteristics which we obtained from *Acxiom*. The average household size (HH size) may represent an indication of the composition of households in a zip code. The share of households with an annual income of less than €25,000 (Income<25k€/year) may be relevant as low incomes may affect price elasticities.

6 Results

In Table 2 we present the results from our Panel IV estimations for the three retail prices of interest. The reported results are those from log-log estimations to allow interpretation as elasticities.[26] The instruments are sufficiently strongly correlated with the endogenous variable in order to identify it, as shown by the high values taken by the F-statistic of the excluded instruments. Also, the Durbin-Wu-Hausman test suggests that consumer search intensity ($\ln(\mu)$) is indeed endogeneous by rejecting the null hypothesis of consumer search being an exogenous regressor. The first stage results for all IV estimations in the paper (i.e. log-log, level-level, with and without control variables) are reported in Table A1 of the Appendix. For information purposes, we also report ordinary least squares panel estimates for the subsequent tables in the Appendix in Table A2 and Table A3.

[26] However, the results from level-level regressions are very similar and reported in Tables A4 and A5 in the Appendix.

The direction of the sign and the significance are always as in the IV estimations but the magnitude is much lower suggesting that neglecting endogeneity leads to underestimation of the impact of consumer search on pricing.

Coming to the results, column 1 of Table A4 provides evidence that the incumbent reacts with a slight increase in its baseline tariff to more search intensity in a zip code. For a change in consumer search intensity by 10% the incumbent raises its tariff by 0.38%. Thus, while we theoretically showed in Section 3 that the effect of search intensity on the incumbent base price can go either way depending on exactly how the search cost distribution changes and on the loyalty of consumers, empirically we find a positive effect. In principle all prices would go down if the share of searchers increase since setting a lower price attracts more consumers, that is also the incumbent would have an incentive to reduce its base tariff to keep consumers from searching. However, if our model provides a reasonable description of the main effects it is apparently the case that the search cost distribution shifts in a way that the base price goes up [REF. TO PROPOSITION 2]. Since a (slightly) lower price would not keep many consumers from searching, the incumbents' incentives to decrease the base tariff is muted. Furthermore, catering to higher search cost consumers allows the incumbent to siphon off larger loyalty rents.

Column 2 shows in contrast that the incumbent reacts to more search activity in its zip code by reducing its cheapest tariff considerably. For a 10% increase in the search activity the incumbent decreases its cheapest tariff by 1.54%. Moreover, column 3 reveals that the overall cheapest tariff in the market provided by an entrant supplier also decreases with more consumer search, whereas the effect is less pronounced as with the incumbent cheapest tariff. For every 10% increase in the search intensity in a zip code the overall cheapest tariff in the market decreases by 0.40%. Consistently with our theoretical model, the incumbent cheapest tariff reacts more strongly to increases in the share of searching consumers than the overall cheapest tariff. The incumbent reacts more sensitively to increased search pressure than an entrant in order to prevent consumers from switching to alternative suppliers. The incumbent would lose a larger mark up when losing a customer as the incumbent cheapest price is higher than the overall cheapest price in the market (and costs are almost the same besides of potential differences in hedging strategies and marketing expenditures which we do not observe). [REF. TO PROPOSITION 2]

We now briefly explain the impact of the other included variables. We estimate the costs pass-through to the end-user retail tariffs, which is much higher in the competitive segments of the electricity retail market. For the incumbent baseline tariff we estimate a pass-through of only around 23% whereas 45% of cost increases are passed on to consumers of the incumbent cheapest tariff and 49% for the cheapest entrant tariff. These different pass-through patterns are in line with Duso and Szücs (2017) who investigate pass-through in the German electricity retail markets and also find that incumbents pass-through to costs to a lesser extent.

The average household size increases the baseline tariff of the incumbent, whereas there is no statistically significant effect on the other tariffs. With regard to its baseline tariff the reaction of the incumbent to a larger share of poor households ($Income < 25k$ €$/year$) is positive but statistically insignificant. However, with regard to its cheapest tariff the incumbent reacts with a significant price decrease the higher the share of low income households. To a lower extent the same is true for the cheapest entrants. This may imply that price sensitive consumers may not be willing to pay a high loyalty premium to stay with the incumbent at a higher tariff compared to the overall cheapest tariff, so that the incumbent has to approach the overall cheapest tariff with its competitive price.

Possibly surprisingly the number of competitors positively affects all prices, whereas the estimated effect is close to zero. One explanation explanation may be that once they have attracted some customer base, entrants follow a similar strategy as the one the incumbent chooses in our theoretical model. That is, after consumers have switched to them, entrants may price discriminate themselves. Our results imply that the number of firms in the markets may not be good competition indicator in regulated markets characterized by incumbency advantages. As we do not have data on all entrants' prices (or on market shares) we cannot analyze these issues in further detail. In line with this explanation, but using symmetric firms, Stahl (1989) shows that in search markets prices may be increasing in the number of firms in the market as the chance to sell to informed consumers becomes very small, destroying the competitive impact of more firms. These coefficients have to be interpreted with caution, however, as they do not reflect causal effects: prices and number of firms are likely to influence to each other. This endogeneity issue does not, however, affect our estimates for the impact of consumer search on pricing strategies and is (therefore) beyond the focus of this paper.

Table 3 presents estimations on the impact of search on the three price dispersion measures, which naturally reflect the results of Table 2. Column 1 focuses on *price dispersion*, measured as the incumbent baseline tariff (P^{IB}) as the upper bound of the tariff distribution minus the overall cheapest tariff by an entrant (P^{OC}) as the lower bound. Evidently, price dispersion goes up if more consumers search, technically since the incumbent slightly increases its base tariff and at the same time the overall cheapest price declines with search. Thus, as Tables 2 and 3 indicate, price dispersion is caused by contrasting pricing strategies of the incumbent and the entrants rather than by different intensities of the same strategies. For every 10% increase in the search intensity, the extent of price dispersion goes up by 3.4% suggesting that consumers' gain from searching increases with the share of searching consumers.

Incumbents react to increased price pressure from consumer search via *price discrimination*, as they offer a cheaper tariff for consumers who search (shoppers), which is still

Table 2: IV estimates of the impact of consumer search on prices

	(1) Incumbent Base: $\ln(p^{IB})$	(2) Incumbent Cheapest: $\ln(p^{IC})$	(3) Overall Cheapest: $\ln(p^{OC})$
Search: $\ln(\mu)$	0.038***	-0.154***	-0.040***
	(0.008)	(0.029)	(0.008)
Costs: $\ln(C)$	0.233***	0.450***	0.491***
	(0.009)	(0.028)	(0.009)
#Competitors	0.000***	0.005***	0.000***
	(0.000)	(0.000)	(0.000)
Average HH size	0.020**	0.086***	0.014*
	(0.008)	(0.027)	(0.007)
Income <25k €/a	0.010	-0.088***	-0.012*
	(0.008)	(0.020)	(0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	23.66	23.66	23.66
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for $\ln(\mu)$ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

above the overall cheapest tariff in the market, and a high incumbent baseline tariff for non-shoppers. Price discrimination becomes more pronounced with increasing search intensity. An increase in the share of informed consumers by 10% widens the gap between the incumbents baseline tariff and its cheaper tariff by 18.8%. The extent of price discrimination unambiguously increases if a larger share of consumers searches, predominantly because the incumbent decreases its cheapest tariff significantly as a reaction to consumer search to aggressively prevent existing customers from switching to competitors.

We also see that the *loyalty premium*, measured as the difference between the incumbents cheapest tariff and the overall cheapest tariff in the market, narrows considerably with search intensity. The more consumers search in a market, the more is the incumbent forced to approach the overall cheapest price. For a 10% increase in search intensity, the loyalty premium narrows by 19.3%. The premium the incumbent can charge over and above the cheapest tariff in the market is larger the more consumers are willing to pay for the services of the incumbent. The larger the share of shoppers, the smaller the loyalty premium. If the share of searchers increases, the incumbent reacts more sensitively to this increased competition than the entrants, since the incumbent would lose a larger mark-up when losing a customer.

Overall, we find that the share of non-shoppers staying with the incumbent base tariff gets milked (i.e. surplus appropriation) with more search. In contrast, those who are

Table 3: IV estimates of the impact of consumer search on dispersion

	(1) Price Dispersion: $\ln(p^{IB} - p^{OC})$	(2) Price Discrimination: $\ln(p^{IB} - p^{IC})$	(3) Online Price Dispersion: $\ln(p^{IC} - p^{OC})$
Search: $\ln(\mu)$	0.336***	1.880***	-1.934***
	(0.063)	(0.382)	(0.405)
Costs: $\ln(C)$	-0.844***	-1.705***	0.342
	(0.065)	(0.351)	(0.381)
#Competitors	0.001***	-0.045***	0.069***
	(0.000)	(0.003)	(0.003)
Average HH size	0.036	-0.317	2.083***
	(0.059)	(0.343)	(0.377)
Income <25k €/a	0.113*	0.969***	-1.170***
	(0.065)	(0.266)	(0.292)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	23.66	23.66	23.66
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

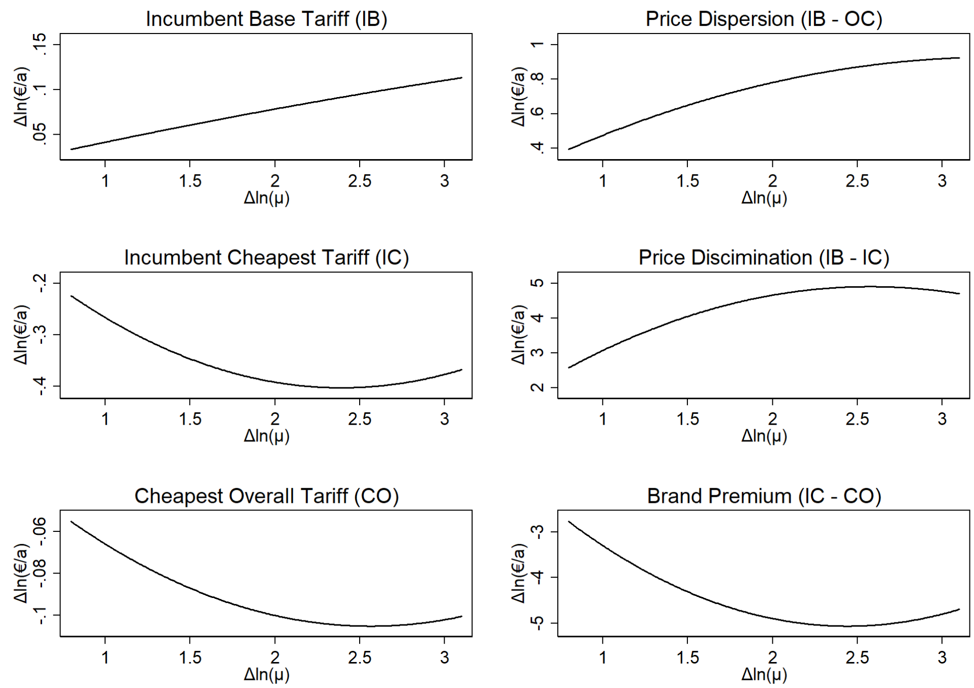
Note: Standard errors clustered at the zip-code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

willing to search either get a cheaper incumbent tariff, which includes a loyalty premium compared to the overall cheapest tariff in the market, or switch to a cheaper entrant tariff. The incumbent reacts to more consumer search with price discrimination, i.e. slightly increasing its base tariff and at the same time drastically reducing its cheaper tariff (i.e. business stealing). Entrants react to more informed consumers with somewhat lower prices. This gives rise to increasing price dispersion and discrimination if there is a larger share of informed consumers in the market, as well as an increasing alignment of incumbent and overall cheapest prices in the competitive segment.

Robustness

The above results are robust for a large variety of alternative specifications. The estimates only differ slightly if we drop the covariates (Table A12 and A13 in the Appendix) and they are are also fully robust to a level-level specification (instead of a log-log specification) as shown in Tables A4 and A5 of the Appendix. Also, as the first stage statistics from Table A1 suggest that broadband coverage is not always significant, we have re-estimated the models with *U40* as the only instrument and the results only change marginally (Tables A14 and A15 in the Appendix). Our findings are also robust against a potentially more general form of non-linear relationship between search and prices. To test this we have added a quadratic term of $\ln(\mu)$ into Equation 6. We instrument for $\ln(\mu)^2$ by using the

Figure 5: Quadratic impact of consumer search on prices and dispersion



Note: Figure presents the quadratic relation of prices and dispersion measures within the 5% and 95% intervals of the logged consumer search.

square of the first stage estimate of $\ln(\mu)$ from Equation 7 as the instrument for $\ln(\mu)^2$.[27] The results are reported in Table A10 and Table A11 in the Appendix. The graphical illustration is shown in Figure 5.

Finally, we have estimated models with different outcome variables. In the first set of these models (Tables A6 and A7) we estimate the impact of search on retailers' markups for different tariffs and differences in these markups. We compute markups as the differences between (net) prices and costs. In the second set (Tables A8 and A9) we use Lerner Indices as the dependent variables. They are computed as the ratio of markups to prices. The results for markups and Lerner Indices are as one would expect from the results of the price estimations.

[27] Wooldridge (2010, p. 262) on this approach

7 Conclusion

After 20 years of market liberalization and the entry of many companies, incumbents still serve many consumers at higher prices than their competitors and prices remain highly dispersed. We provide a theoretical model and an empirical analysis addressing these issues. Using a search theoretic model incorporating incumbency advantages, we show that the incumbent baseline price may increase when consumers search more. Unless they make an active move to search and switch at a cost, consumers stay with the incumbent contract. The incentive to increase the baseline tariff arises if a lower price would not keep many consumers from searching and catering to high search cost consumers allows the incumbent to siphon off larger loyalty rents. In contrast, once consumers have shown a willingness to search, the incumbent has all the incentive to prevent consumers from switching to an entrant by setting a low online price. That is, incumbents engage in price discrimination with a high baseline price and a low online price. The incumbent's cheapest online tariff is still higher, however, than the overall cheapest tariff in the market, as consumers attach a loyalty premium to the incumbent. By engaging in price discrimination, the incumbent segments the markets into a segment of consumers with higher search cost and those with lower search cost.

Using data on the three tariffs (the incumbent's baseline price, the incumbent's cheapest (online) price and the overall cheapest price) for German electricity retail areas (ZIP codes) in the years 2011–2014, we empirically find that on average, the incumbent's baseline tariff is larger in areas where consumers search more, whereas the incumbent's online tariff is aggressively reduced in areas with intensified consumer search. The overall cheapest tariff set by an entrant also decreases if there is more search. This pattern of price effects give rise to increased price dispersion and price discrimination as well as a reduced online price dispersion if consumers search more. The empirical analysis uses an instrumental variable approach to identify causal effects: we instrument for search intensity with the share of young households and the local broadband coverage per ZIP code area, factors that are positively related to online search, but independent of the electricity prices.

The success of market liberalization hinges on the possible strategies the incumbent is able to choose. Price discrimination allows the incumbent to segment markets according to search behavior, and to simultaneously follow a surplus appropriation strategy for consumers with high search and switching cost as well as a business stealing strategy for those consumers who indicated they are willing to engage in search. In this way, few consumers may actually switch suppliers, while appropriating an important share of the market rents. This strategy may explain why even after 20 years of market liberalization, incumbent's prices (and possibly average prices) are still relatively high, and why a large fraction of consumers remain buying from the incumbent.

References

Brown, J. R., Goolsbee, A., 2002. Does the internet make markets more competitive? evidence from the life insurance industry. Journal of the Political Economy 110 (3), 481–507.

Brynjolfsson, E., Smith, M., 2001. Consumer decision-making at an internet shopbot: Brand still matters. Journal of Industrial Economics 49 (4), 541–558.

Brynjolfsson, E., Smith, M., 2014. Frictionless commerce? a comparison of internet and conventional retailers. Management Science 46, 563–585.

Bundesnetzagentur, 2013. Monitoring report 2013.

Bundesnetzagentur, 2015. Monitoring report 2015.

Cabral, L., 2017. Switching costs, search costs, and market power: Theory and application to energy markets, Closing Address, V International Academic Symposium, Challenges for the Energy Sector, Barcelona, February 7.

Chandra, A., Tappata, M., 2011. Consumer search and dynamic price dispersion: An application to gasoline markets. RAND Journal of Economics 42 (2), 681–704.

Dana, J., 1994. Learning in an equilibrium search model. International Economic Review 35, 745–771.

Davies, S., Price, C. W., Wilson, C. M., 2014. Nonlinear pricing and tariff differentiation: Evidence from the british electricity market. The Energy Journal 35 (1), 57–77.

De los Santos, B., Hortacsu, A., Wildenbeest, M., 2012. Testing models of consumer search using data on web browsing and purchasing behavior. American Economic Review 102, 29552980.

Duso, T., Szücs, F., 2017. Market power and asymmetric pass-through in German electricity retail. European Economic Review 98, 354–372.

Giulietti, M., Waterson, M., Wildenbeest, M., 2014. Estimation of search frictions in the british electricity market. The Journal of Industrial Economics 62, 555–590.

Goolsbee, A., Syverson, C., 2008. How incumbents respond to the threat of entry? evidence from the major airlines. Quarterly Journal of Economics 123 (4), 1611–1633.

Hortacsu, A., Madanizadeh, S., Puller, S., 2015. Power to choose? an analysis of consumer inertia in the residential electricity market, NBER Working Paper No. 20988.

Janssen, C. W., Moraga-González, J. L., 2004. Strategic pricing, consumer search and the number of firms. Review of Economic Studies 71 (4), 1089–1118.

Koulayev, S., 2014. Search for differentiated products: Identification and estimation. RAND Journal of Economics 45 (3), 553–575.

Miravete, E. J., 2011. Competition and the use of foggy pricing. American Economic Journal: Microeconomics 5 (1), 194–216.

Pennersdorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., Yontcheva, B., 2014. Information and price dispersion: Evidence from retail gasoline, University of Vienna, Working Paper.

Salop, S., 1977. The noisy monopolist: Imperfect information, price dispersion and price discrimination. Review of Economic Studies 44, 393–406.

Seim, K., Viard, V. B., 2011. The effect of market structure on cellular technology adoption and pricing. American Economic Journal: Microeconomics 3 (2), 221–251.

Sengupta, A., Wiggins, S. N., 2014. Airline pricing, price dispersion, and ticket characteristics on and off the internet. American Economic Journal: Economic Policy 6 (1), 272–307.

Sorensen, A. T., 2000. Equilibrium price dispersion in retail markets for prescription drugs. Journal of Political Economy 108 (4), 833–850.

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

Tang, Z., Smith, M. D., Montgomery, A., 2010. The impact of shopbot use on prices and price dispersion: Evidence from online book retailing. International Journal of Industrial Organization 28, 579–590.

Tappata, M., 2009. Rockets and feathers: Understanding asymmetric pricing. RAND Journal of Economics 40, 673687.

Wooldridge, J. M., 2010. Econometric Analysis of Cross Section and Panel Data. The MIT Press.

Appendix

A Proofs of Propositions

Proposition 1. Price discrimination increases if, and only if, the inverse hazard rate evaluated at the equilibrium values $\frac{1-F(p^{IB^*}-p^{IC^*};z)}{f(p^{IB^*}-p^{IC^*};z)}$ is increasing in z. Moreover, online price dispersion is positively related to p^{IC} if the density functions are non-increasing, i.e., $\partial f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)/\partial(p^{IC}-p^{OC}) \leq 0$. Finally, online dispersion and price discrimination are linked by $1+f(p^{IB}-p^{IC};z)(p^{IB}-p^{IC}) = f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)\left(\frac{p^{IC}+p^{OC}}{\theta}\right)$.

Proof. For the result on price discrimination, we take the total differential of (3) with respect to $p^{IB}-p^{IC}$ and z, yielding

$$\left[-2f(p^{IB}-p^{IC};z)-\frac{\partial f(p^{IB}-p^{IC};z)}{\partial(p^{IB}-p^{IC})}(p^{IB}-p^{IC})\right]d(p^{IB}-p^{IC}) \tag{8}$$
$$=\left[\frac{\partial f(p^{IB}-p^{IC};z)}{\partial z}(p^{IB}-p^{IC})+\frac{F(p^{IB}-p^{IC};z)}{\partial z}\right]dz$$

As profit maximization implies that the second-order condition of (3) with respect to $p^{IB}-p^{IC}$ is negative, it should be that in an equilibrium,

$$-2f(p^{IB}-p^{IC};z)-\frac{\partial f(p^{IB}-p^{IC};z)}{\partial(p^{IB}-p^{IC})}(p^{IB}-p^{IC})<0.$$

On the other hand, the inverse hazard rate $\frac{1-F(p^{IB^*}-p^{IC^*};z)}{f(p^{IB^*}-p^{IC^*};z)}$ is increasing in z if, and only if,

$$-\frac{\partial f(p^{IB}-p^{IC};z)}{\partial x}(1-F(p^{IB}-p^{IC};z))-\frac{F(p^{IB}-p^{IC};z)}{\partial z}f(p^{IB}-p^{IC})>0,$$

which using the (3) can be rewritten as

$$-f(p^{IB}-p^{IC};z)\left[\frac{\partial f(p^{IB}-p^{IC};z)}{\partial z}(p^{IB}-p^{IC})+\frac{F(p^{IB}-p^{IC};z)}{\partial z}\right]>0.$$

Thus, if the inverse hazard rate is increasing in z, then in any equilibrium both square bracket terms in (8) are negative, implying $\frac{d(p^{IB}-p^{IC})}{dz}>0$.

To investigate online price dispersion, we take the total differential of (1) with respect to p^{IC} and p^{OC} to get

$$0 = \frac{1}{\theta}\left[f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right) - f^{'}\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)\frac{p^{OC}}{\theta}\right]dp^{IC}$$
$$+\frac{1}{\theta}\left[-2f\left(\frac{p^{IC}-p^{OC}}{\theta};z\right) + f^{'}\left(\frac{p^{IC}-p^{OC}}{\theta};z\right)\frac{p^{OC}}{\theta}\right]dp^{OC},$$

where $f^{'}$ is the derivative of the density function with respect to prices. From the second-order condition for profit maximization by the entrant we know that he second term in square brackets must be negative. If $f^{'}\left(\frac{p^{IC}-p^{OC}}{\theta};z\right) \leq 0$, then the first term in square brackets is positive, and in absolute value smaller than the first term in square brackets. Thus, $0 < dp^{OC}/dp^{IC} < 1$. Therefore, $0 < (dp^{OC} - dp^{OC})/dp^{IC} < 1$.

Finally, to understand how price discrimination and online price dispersion are related, we substitute (1) and (3) into (2) to get the condition stated in the Proposition. Q.E.D.

Proposition 2 (price levels). If (5) holds, then an increase in the fraction of online searchers $F(p^{IB^*} - p^{IC^*})$, initiated by an increase in z, coincides with a decrease in online prices $p^{OC^*}$and p^{IC^*}if, and only if $\widetilde{s}_2 - \widetilde{s}_1 > \widetilde{s}_2\widetilde{s}_1$, while it coincides with an increase in the baseline price p^{IB^*} if θ is small enough, z is large enough, or $\widetilde{s}_2 - \widetilde{s}_1$ is small enough.

Proof. It is clear that

$$\frac{\partial(p^{IB^*} - p^{IC^*})}{\partial z} = \frac{(\widetilde{s}_2 - \widetilde{s}_1)\widetilde{s}_1(1-\widetilde{s}_2)}{2\left(\widetilde{s}_2 - z\widetilde{s}_1\right)^2} > 0.$$

From the expressions determining equilibrium prices, it follows that

$$2\frac{\partial p^{OC^*}}{\partial z} = \frac{\partial p^{IC^*}}{\partial z} = -\frac{\theta}{3z^2}\left(\frac{\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}\right),$$

which is clearly negative if, and only if $\widetilde{s}_2 - \widetilde{s}_1 > \widetilde{s}_2\widetilde{s}_1$. Also,

$$\frac{\partial p^{IB^*}}{\partial z} = \frac{\theta}{3z^2}\left(-1 + \frac{\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}\right) + \frac{(\widetilde{s}_2 - \widetilde{s}_1)\widetilde{s}_1(1-\widetilde{s}_2)}{2\left(\widetilde{s}_2 - z\widetilde{s}_1\right)^2}.$$

As the second term is positive, this is clearly positive if either the first term is small enough (θ is small enough or z is large enough), or the first term is positive ($\widetilde{s}_2 - \widetilde{s}_1$ is small enough). *Q.E.D.*

Proposition 3 (Price discrimination and dispersion). If (5) holds, then an increase in the fraction of online searchers $F(p^{IB^*} - p^{IC^*})$, initiated by an increase in z, coincides with (i) an increase in price discrimination $p^{IB^*} - p^{IC^*}$ and (ii) a decrease in online price dispersion $p^{IC^*} - p^{OC^*}$, if and only if $\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2 > 0$ and (iii) an increase in price dispersion $p^{IB^*} - p^{OC^*}$ if θ is small enough, x is large enough, or $\widetilde{s}_2 - \widetilde{s}_1$ is small enough.

Proof. The proof simply follows from calculating the different partial derivatives. As

$$\frac{\partial(p^{IB^*} - p^{IC^*})}{\partial z} = \frac{(\widetilde{s}_2 - \widetilde{s}_1)\widetilde{s}_1(1 - \widetilde{s}_2)}{2\left(\widetilde{s}_2 - z\widetilde{s}_1\right)^2} > 0$$

and

$$\frac{\partial F(p^{IB^*} - p^{IC^*})}{\partial z} = \frac{\widetilde{s}_1\widetilde{s}_2}{2\left(\widetilde{s}_2 - \widetilde{s}_1\right)} > 0,$$

an increase in the fraction of online searchers, initiated by an increase in z, certainly leads to an increase in price discrimination $p^{IB^*} - p^{IC^*}$. As

$$\frac{\partial(p^{IC^*} - p^{OC^*})}{\partial z} = -\frac{\theta}{6z^2}\left(\frac{\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}\right)$$

it leads to a decrease in online price dispersion if $\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2 > 0$. Finally, as

$$\frac{\partial(p^{IB^*} - p^{OC^*})}{\partial z} = \frac{\theta}{6z^2}\left(-1 + \frac{\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}\right) + \frac{(\widetilde{s}_2 - \widetilde{s}_1)\widetilde{s}_1(1 - \widetilde{s}_2)}{2\left(\widetilde{s}_2 - z\widetilde{s}_1\right)^2},$$

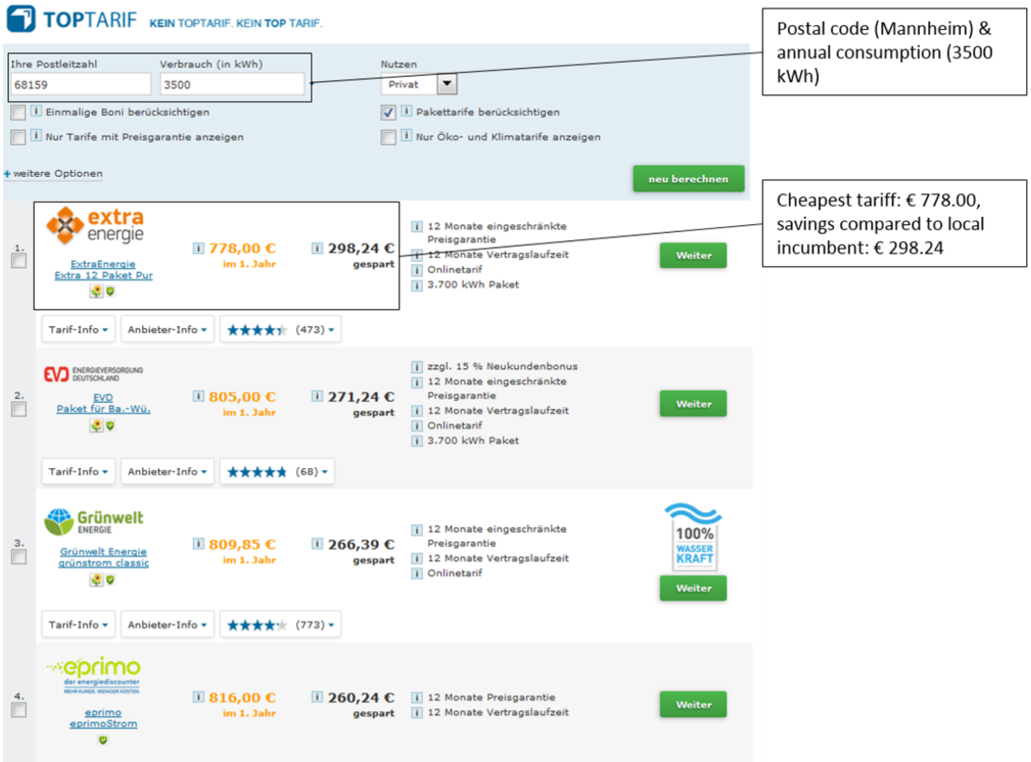
and the second term is positive, it leads to an increase in price discrimination if either the first term is small enough (θ is small enough or z is large enough), or the first term is positive ($\widetilde{s}_2 - \widetilde{s}_1$ is small enough). *Q.E.D.*

B Additional Comparative Statics

Even under a piece-wise linear specification of the search cost distribution, the theoretical model allows for a rich set of comparative statics results, depending on how the shift of the search cost distribution is modelled. Interestingly, even if the search cost distribution shifts probability mass towards lower search costs, it is not necessary that there will be more online search. In the main body of the paper we have seen that if x is increasing, then lower search costs coincide with more online search, but a shift in the search cost distribution towards lower search costs, represented by a change in $\widetilde{s}_2$ coincides with less online search. The reason is that a different part of the search cost distribution is shifted. An increase in x represents more consumers with the lowest search cost levels, whereas an increase in $\widetilde{s}_2$ (for $x > 1$) – or a decrease in $\widetilde{s}_2$ (for $x < 1$) implies that probability mass is shifted from very high search cost levels to intermediate search cost levels (where the fraction of consumers with the lowest search costs is not affected). The next result states this more formally for

Claim 4 *(a) An increase in $\widetilde{s}_2$ leads to an increase in the fraction of online searchers $F(p^{IB^*} - p^{IC})$, an increase in price discrimation $p^{IB^*} - p^{IC}$ and an increase in online* *prices, if and only if If* $x < 1$. *(b) If* $x < 1$ *an increase in* $\widetilde{s}_1$ *leads to a decrease in the fraction of online searchers* $F(p^{IB^*} - p^{IC})$, *an increase in price discrimation* $p^{IB^*} - p^{IC}$ *and a decrease in online prices. If* $x > 1$, *the impact on the fraction of online searchers is ambiguous, whereas the other effects are reversed.*

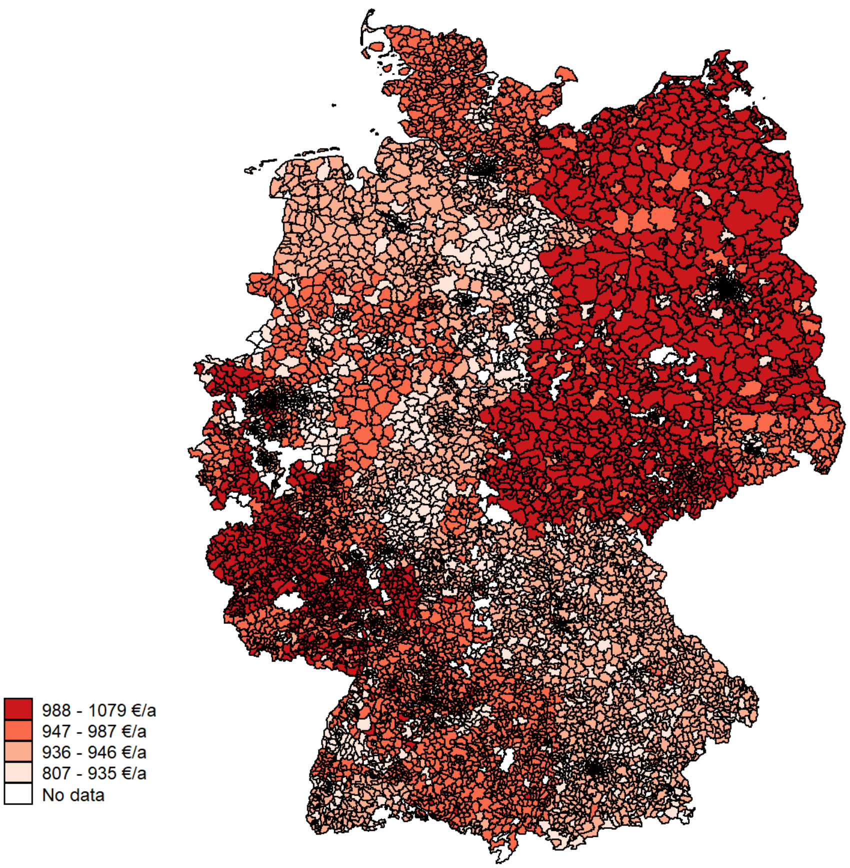
Figure A1: Screenshot of a typical online comparison platform



Note: the comparison platform lists all available tariffs for a consumer given his/her annual consumption level for his/her local zip code, starting with the cheapest available tariff (including annual savings compared to the default incumbent baseline tariff). Site access on July 25, 2015.

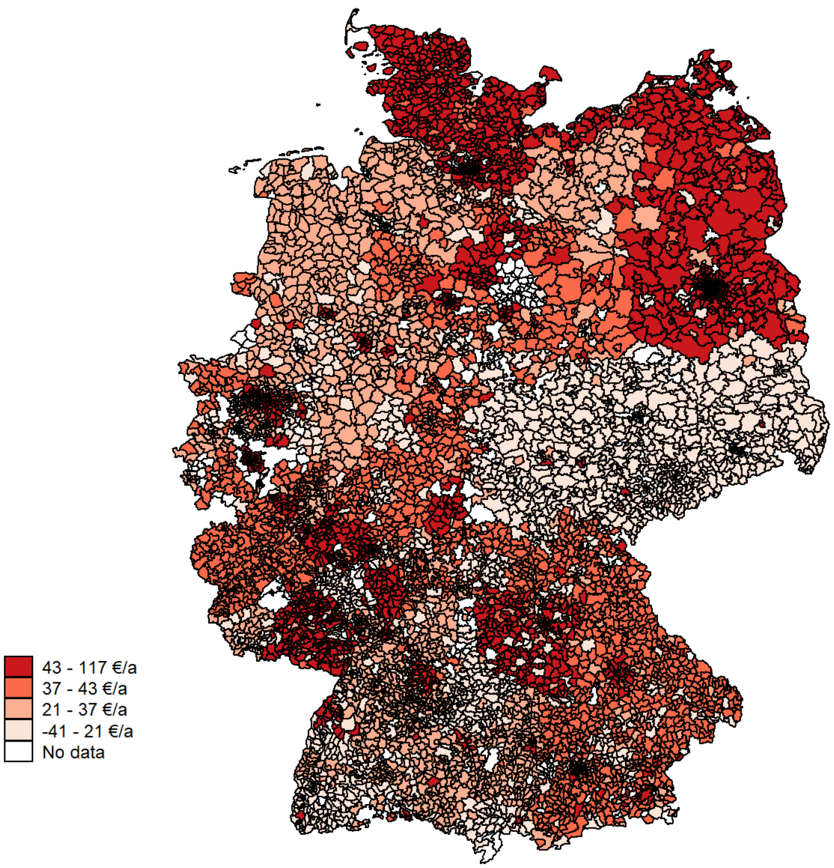
C Additional Figures and Tables

Figure A2: Incumbent Base Tariffs (2012)



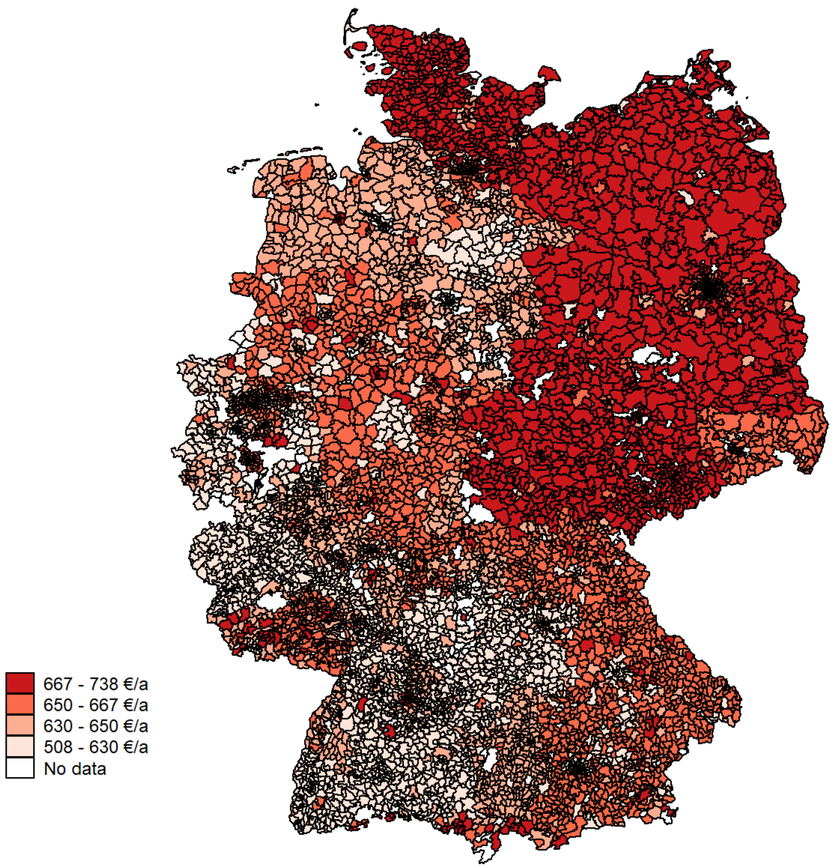
Note: The figure presents the spatial distribution of the incumbent base tariffs for the year 2012.

Figure A3: Change of Incumbent Base Tariffs from 2011 to 2012



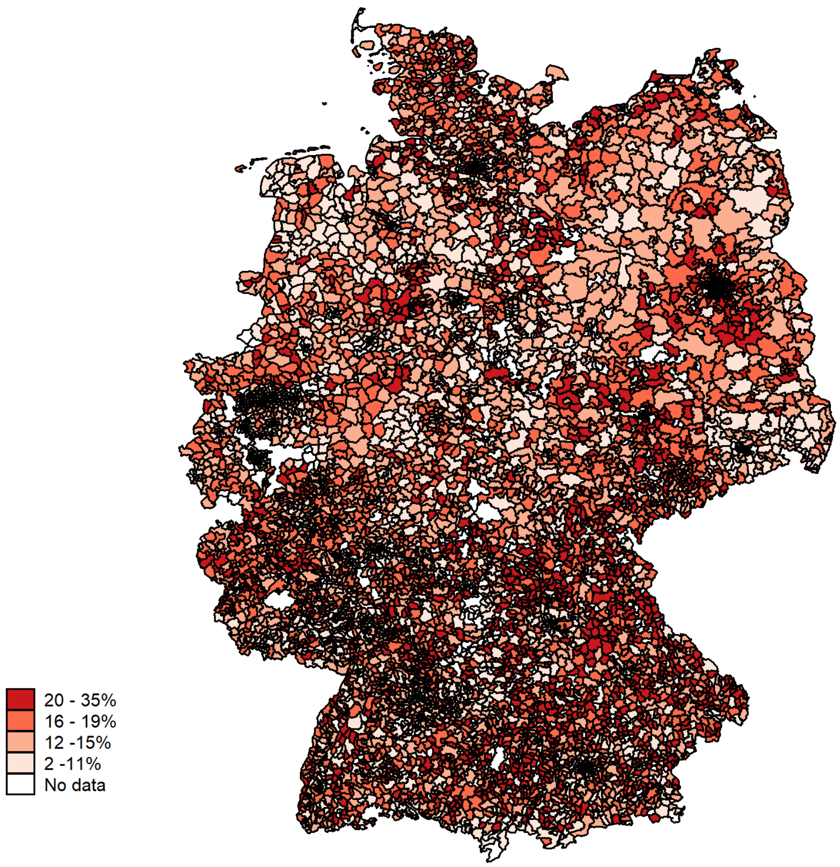
Note: The figure presents the spatial distribution of changes in the incumbent base tariffs from 2011 to 2012.

Figure A4: (Net) Costs (2012)



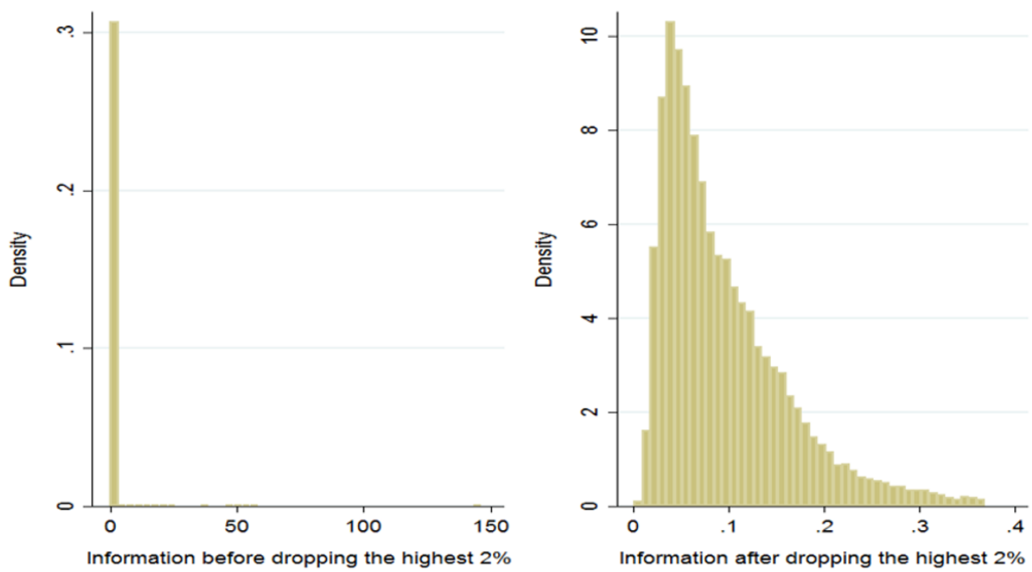
Note: The figure presents the spatial distribution of the (net) costs for electricity retailers in 2012.

Figure A5: Consumer Search Activity (2012)



Note: Figure presents the spatial distribution of consumer search intensity in 2012.

Figure A6: Distribution of consumer search intensity before and after trimming the highest 2% of the observations.



Note: We trim the highest 2% of the observations on consumer search, as they may be the result of data crawling "shop bots" or data research but not from actual consumers searching for electricity tariffs.

Table A1: First-Stage regressions of consumer search (μ)

	(1) $\ln(\mu)$	(2) $\ln(\mu)$	(3) μ	(4) μ
U40	0.0185***	0.0152***	0.1081***	0.1036***
	(0.0029)	(0.0028)	(0.0273)	(0.0271)
BBC	0.0001	0.0002	0.0024	0.0039**
	(0.0002)	(0.0002)	(0.0017)	(0.0018)
Costs: $\ln(C)$	0.4711***			
	(0.1247)			
Costs: (C)			0.0270***	
			(0.0019)	
# Competitors	0.0066***		0.0844***	
	(0.0006)		(0.0048)	
Average HH size	0.8072***		4.1732***	
	(0.1009)		(1.0051)	
Income <25k €/year	-0.1078		-4.5401***	
	(0.1063)		(1.0845)	
Zip code fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	29,090	29,090	29,090	29,090

Notes: Standard errors clustered at the zip code level in parentheses. Instruments are reported in bold. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A2: OLS estimates of the impact of consumer search on prices

	(1) Incumbent Base: $\ln(p^{IB})$	(2) Incumbent Cheapest: $\ln(p^{IC})$	(3) Overall Cheapest: $\ln(p^{OC})$
Search: $\ln(\mu)$	0.003***	-0.004***	-0.001***
	(0.000)	(0.001)	(0.000)
Costs: $\ln(C)$	0.250***	0.379***	0.471***
	(0.008)	(0.016)	(0.007)
#Competitors	0.001***	0.004***	-0.000
	(0.000)	(0.000)	(0.000)
Average HH size	0.043***	-0.014	-0.012***
	(0.005)	(0.013)	(0.004)
Income <25k €/a	0.006	-0.074***	-0.008*
	(0.007)	(0.013)	(0.004)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	29,090	29,090	29,090

Notes: Robust standard errors clustered at the zip code level in parentheses. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A3: OLS estimates of the impact of consumer search on dispersion

	(1) Price Dispersion: $\ln(p^{IB} - p^{OC})$	(2) Price Discrimination: $\ln(p^{IB} - p^{IC})$	(3) Loyalty Premium: $\ln(p^{IC} - p^{OC})$
Search: $\ln(\mu)$	0.014***	0.053***	-0.031*
	(0.002)	(0.015)	(0.018)
Costs: $\ln(C)$	-0.686***	-0.835***	-0.592***
	(0.045)	(0.206)	(0.229)
#Competitors	0.004***	-0.032***	0.056***
	(0.000)	(0.001)	(0.001)
Average HH size	0.253***	0.890***	0.831***
	(0.030)	(0.183)	(0.199)
Income <25k €/a	0.083	0.795***	-0.980***
	(0.052)	(0.182)	(0.218)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A4: IV estimates of the impact of consumer search on prices – level-level estimates

	(1) Incumbent Base: p^{IB}	(2) Incumbent Cheapest: p^{IC}	(3) Overall Cheapest: p^{OC}
Search: μ	3.823***	-18.799***	-3.977***
	(1.384)	(5.481)	(1.199)
Costs: C	0.263***	1.030***	0.651***
	(0.040)	(0.154)	(0.034)
#Competitors	0.283**	4.922***	0.342***
	(0.124)	(0.482)	(0.106)
Average HH size	19.923***	39.003	4.698
	(7.090)	(27.959)	(6.240)
Income <25k €/a	20.378**	-146.639***	-22.987***
	(9.581)	(33.482)	(7.725)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	7.59	7.59	7.59
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A5: IV estimates of the impact of consumer search on dispersion – level-level estimates

	(1) Price Dispersion: $p^{IB} - p^{OC}$	(2) Price Discrimination: $p^{IB} - p^{IC}$	(3) Loyalty Premium: $p^{IC} - p^{OC}$
Search: μ	7.763***	22.604***	-14.680***
	(2.256)	(6.286)	(4.563)
Costs: C	-0.387***	-0.768***	0.376***
	(0.064)	(0.176)	(0.129)
#Competitors	-0.057	-4.640***	4.566***
	(0.199)	(0.551)	(0.402)
Average HH size	15.393	-19.351	34.073
	(11.631)	(31.769)	(23.447)
Income <25k €/a	43.174***	167.128***	-123.080***
	(15.088)	(38.807)	(27.924)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	7.59	7.59	7.59
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A6: IV estimates of the impact of consumer search on markups

	(1) Markup Incumbent Base: $\ln(m^{IB})$	(2) Markup Incumbent Cheapest: $\ln(m^{IC})$	(3) Markup Overall Cheapest: $\ln(m^{OC})$
Search: $\ln(\mu)$	0.407***	-4.555***	-1.591***
	(0.080)	(0.932)	(0.535)
Costs: $\ln(C)$	-3.398***	-6.171***	-28.350***
	(0.074)	(0.873)	(0.630)
#Competitors	0.002***	0.177***	0.014***
	(0.001)	(0.008)	(0.004)
Average HH size	0.012	4.553***	2.647***
	(0.069)	(0.843)	(0.519)
Income <25k €/a	0.158**	-3.825***	-1.002**
	(0.075)	(0.697)	(0.433)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	23.66	23.66	23.66
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A7: IV estimates of the impact of consumer search on markup dispersion

	(1) Markup Dispersion: $\ln(m^{IB} - m^{OC})$	(2) Markup Discrimination: $\ln(m^{IB} - m^{IC})$	(3) Markup Loyalty Premium: $\ln(m^{IC} - m^{OC})$
Search: $\ln(\mu)$	0.329***	2.186***	-1.715***
	(0.063)	(0.450)	(0.347)
Costs. $\ln(C)$	-1.287***	-3.978***	0.133
	(0.069)	(0.468)	(0.351)
#Competitors	0.001**	-0.044***	0.058***
	(0.000)	(0.003)	(0.003)
Average HH size	0.060	-0.284	1.408***
	(0.059)	(0.413)	(0.328)
Income <25k €/a	0.113*	0.874***	-1.020***
	(0.064)	(0.307)	(0.249)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	23.66	23.66	23.66
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A8: IV estimates of the impact of consumer search on Lerner Indices

	(1) Lerner Index Incumbent Base: $\ln(L^{IB} - L^{OC})$	(2) Lerner Index Incumbent Cheapest: $\ln(L^{IB} - L^{IC})$	(3) Lerner Index Overall Cheapest: $\ln(L^{IC} - L^{OC})$
Search: $\ln(\mu)$	0.031***	-0.124***	-0.034***
	(0.006)	(0.023)	(0.007)
Costs: $\ln(C)$	-0.527***	-0.341***	-0.355***
	(0.007)	(0.022)	(0.008)
#Competitors	0.000***	0.004***	0.000***
	(0.000)	(0.000)	(0.000)
Average HH size	0.011**	0.066***	0.008
	(0.006)	(0.021)	(0.006)
Income <25k €/a	0.009	-0.067***	-0.010*
	(0.006)	(0.016)	(0.005)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	23.66	23.66	23.66
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A9: IV estimates of the impact of consumer search on Lerner Index dispersion

	(1) Lerner Index Dispersion: $\ln(L^{IB} - L^{OC})$	(2) Lerner Index Discrimination: $\ln(L^{IB} - L^{IC})$	(3) Lerner Index Loyalty Premium: $\ln(L^{IC} - L^{OC})$
Search: $\ln(\mu)$	0.064***	0.154***	-0.090***
	(0.011)	(0.026)	(0.019)
Costs: $\ln(C)$	-0.172***	-0.185***	0.015
	(0.012)	(0.025)	(0.019)
#Competitors	0.000	-0.003***	0.003***
	(0.000)	(0.000)	(0.000)
Average HH size	0.004	-0.055**	0.059***
	(0.011)	(0.024)	(0.018)
Income <25k €/a	0.018*	0.075***	-0.057***
	(0.010)	(0.019)	(0.014)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	23.66	23.66	23.66
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A10: IV estimates of the impact of consumer search on prices – non-linear specifications

	(1) Incumbent Base: $\ln(p^{IB})$	(2) Incumbent Cheapest: $\ln(p^{IC})$	(3) Overall Cheapest: $\ln(p^{OC})$
Search: $\ln(\mu)$	0.044***	-0.337***	-0.082***
	(0.010)	(0.033)	(0.009)
Search^2: $\ln(\mu)^2$	-0.002***	0.070***	0.016***
	(0.001)	(0.003)	(0.001)
Costs: $\ln(C)$	0.243***	0.158***	0.424***
	(0.008)	(0.025)	(0.008)
#Competitors	0.000***	0.003***	-0.000**
	(0.000)	(0.000)	(0.000)
Average HH size	0.020***	0.079***	0.013*
	(0.008)	(0.026)	(0.007)
Income <25k €/a	0.007	-0.006	0.007
	(0.008)	(0.021)	(0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	17.76	17.76	17.76
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A11: IV estimates of the impact of consumer search on dispersion – non-linear specifications

	(1) Price Dispersion: $\ln(p^{IB} - p^{OC})$	(2) Price Discrimination: $\ln(p^{IB} - p^{IC})$	(3) Loyalty Premium: $\ln(p^{IC} - p^{OC})$
Search: $\ln(\mu)$	0.558***	3.806***	-4.147***
	(0.071)	(0.440)	(0.453)
Search^2: $\ln(\mu)^2$	-0.084***	-0.739***	0.849***
	(0.006)	(0.036)	(0.036)
Costs: $\ln(C)$	-0.496***	1.365***	-3.182***
	(0.060)	(0.335)	(0.341)
#Competitors	0.003***	-0.029***	0.051***
	(0.000)	(0.002)	(0.003)
Average HH size	0.042	-0.252	2.018***
	(0.056)	(0.332)	(0.354)
Income <25k €/a	0.016	0.108	-0.183
	(0.063)	(0.268)	(0.288)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	17.76	17.76	17.76
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A12: IV estimates of the impact of consumer search on prices – estimations without covariates

	(1) Incumbent Base: $\ln(p^{IB})$	(2) Incumbent Cheapest: $\ln(p^{IC})$	(3) Overall Cheapest: $\ln(p^{OC})$
Search: $\ln(\mu)$	0.040***	-0.119***	-0.030***
	(0.010)	(0.032)	(0.009)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	16.68	16.68	16.68
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A13: IV estimates of the impact of consumer search on dispersion – Estimations without covariates

	(1) Price Dispersion: $\ln(p^{IB} - p^{OC})$	(2) Price Discrimination: $\ln(p^{IB} - p^{IC})$	(3) Loyalty Premium: $\ln(p^{IC} - p^{OC})$
Search: $\ln(\mu)$	0.295***	1.429***	-1.852***
	(0.070)	(0.404)	(0.496)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	16.68	16.68	16.68
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A14: IV estimates of the impact of consumer search on prices – Search instrumented only by U40

	(1) Incumbent Base: $\ln(p^{IB})$	(2) Incumbent Cheapest: $\ln(p^{IC})$	(3) Overall Cheapest: $\ln(p^{OC})$
Search: $\ln(\mu)$	0.038***	-0.162***	-0.043***
	(0.008)	(0.030)	(0.008)
Costs: $\ln(C)$	0.233***	0.455***	0.491***
	(0.009)	(0.029)	(0.009)
#Competitors	0.000***	0.005***	0.000***
	(0.000)	(0.000)	(0.000)
Average HH size	0.020**	0.090***	0.015**
	(0.008)	(0.028)	(0.008)
Income <25k €/a	0.010	-0.089***	-0.012*
	(0.008)	(0.021)	(0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	46.86	46.86	46.86
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A15: IV estimates of the impact of consumer search on dispersion – Search instrumented only by U40

	(1) Price Dispersion: $\ln(p^{IB} - p^{OC})$	(2) Price Discrimination: $\ln(p^{IB} - p^{IC})$	(3) Loyalty Premium: $\ln(p^{IC} - p^{OC})$
Search: $\ln(\mu)$	0.360***	1.969***	-2.040***
	(0.065)	(0.394)	(0.418)
Costs: $\ln(C)$	-0.853***	-1.762***	0.381
	(0.067)	(0.362)	(0.392)
#Competitors	0.001**	-0.045***	0.069***
	(0.001)	(0.003)	(0.003)
Average HH size	0.024	-0.377	2.160***
	(0.061)	(0.353)	(0.388)
Income <25k €/a	0.117*	0.983***	-1.177***
	(0.066)	(0.274)	(0.300)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	46.86	46.86	46.86
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A16: IV estimates of the impact of consumer search on prices – non-linear specifications of level-level estimations

	(1) Incumbent Base: p^{IB}	(2) Incumbent Cheapest: p^{IC}	(3) Overall Cheapest: p^{OC}
Search: μ	12.582**	-85.103***	-19.414***
	(5.081)	(28.511)	(6.578)
Search^2: μ^2	-0.294**	2.376***	0.551***
	(0.123)	(0.692)	(0.160)
Costs: C	0.300***	0.602***	0.552***
	(0.026)	(0.142)	(0.033)
#Competitors	0.138	5.705***	0.527**
	(0.186)	(1.028)	(0.237)
Average HH size	3.717	155.686**	31.607*
	(12.779)	(71.820)	(16.616)
Income <25k €/a	9.944	-41.505	1.321
	(7.885)	(33.343)	(7.825)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-Stage F stat.	3.43	3.43	3.43
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29,090	29,090	29,090

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A17: IV estimates of the impact of consumer search on dispersion – non-linear specifications of level-level estimations

	(1) Price Dispersion: $p^{IB} - p^{OC}$	(2) Price Discrimination: $p^{IB} - p^{IC}$	(3) Loyalty Premium: $p^{IC} - p^{OC}$
Search: μ	31.703***	97.349***	-65.442***
	(10.353)	(31.638)	(22.599)
Search^2: μ^2	-0.838***	-2.662***	1.819***
	(0.251)	(0.768)	(0.548)
Costs: C	-0.251***	-0.301*	0.050
	(0.052)	(0.158)	(0.113)
#Competitors	-0.379	-5.556***	5.169***
	(0.375)	(1.142)	(0.815)
Average HH size	-27.800	-152.103*	123.616**
	(26.174)	(79.606)	(56.996)
Income <25k €/a	8.418	51.155	-42.736
	(13.719)	(37.618)	(26.832)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Kleibergen-Paap F stat.	3.43	3.43	3.43
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	29051	29051	29051

Note: Standard errors clustered at zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBC*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.