

Discussion Paper No. 18-042

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

Klaus Gugler, Sven Heim, Maarten Janssen,
and Mario Liebensteiner

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 18-042

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

Klaus Gugler, Sven Heim, Maarten Janssen,
and Mario Liebensteiner

Download this ZEW Discussion Paper from our ftp server:

<http://ftp.zew.de/pub/zew-docs/dp/dp18042.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von
neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung
der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other
economists in order to encourage discussion and suggestions for revisions. The authors are solely
responsible for the contents which do not necessarily represent the opinion of the ZEW.

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

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

September 18, 2018

Abstract

We study how consumer search affects pricing in markets with incumbents and entrants using panel data on German electricity retail markets. Consumers observe the baseline price of the incumbent and decide whether or not to search. Incumbent providers can price discriminate between searching and loyal consumers. Empirically we show that local incumbents increase their baseline rate while entrants decrease their tariffs if consumer search increases. Moreover, the incumbent price discriminates more strongly in markets with more consumer search. Using a theoretical model, we show that these pricing patterns are consistent with the strategic interaction of profit-maximizing firms.

Keywords: Search, Price Dispersion, Price Discrimination, Electricity

JEL Classification: D43, D83, L11, L13, Q40

*We thank *ene't* for giving us access to their data on consumer search at price comparison websites. We thank Klenio Barbosa, Anette Boom, Tomaso Duso, Daniel Garcia, Alessandro Gavazza, Ulrich Laitenberger, Dieter Pennersdorfer, Steven Puller, Karl Schlag, Philipp Schmidt-Dengler, Nils-Hendrik von der Fehr, Mike Waterson, Matthijs Wildenbeest, Biliana Yontcheva, and Christine Zulehner for valuable comments. We also thank *Axiom* for giving us data on household characteristics. The project was financially supported by the State Government of Baden-Württemberg, Germany, through the research program Strengthening Efficiency and Competitiveness in the European Knowledge Economies (SEEK). 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 Center 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

[¶]University of Kaiserslautern, mario.liebensteiner@wiwi.uni-kl.de

1 Introduction

Many markets where liberalization has taken place are characterized by a substantial 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. Knowing the alternative contracts provided, consumers also have to be willing to switch. Reacting to the presence of entrants, the incumbent may use this asymmetry and price discriminate between consumers with high search costs and consumers with lower search costs who may consider leaving. This paper studies the asymmetric interaction between the incumbent and the entrants, consumers' willingness to search for and switch to alternative providers, and the price discrimination strategy of the incumbent. In our empirical analysis, we investigate these relationships using data from the German retail electricity markets. 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 consumer search behavior is endogenously determined by the pricing behavior of firms. Our theoretical model shows the general mechanism that explains why many liberalized markets are still characterized by high retail prices and price dispersion.¹

The empirical part of our analysis employs a unique data set on retail electricity prices and consumer search intensity 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 search for cheaper offers and switch to any alternative tariff that is available in their zip code. Even though in recent years most consumers use online platforms to search for cheaper rates,² in 2015 76% of households stayed with the incumbent, with 33% remaining at the expensive baseline tariff, while 43% switched to a cheaper incumbent tariff, and only 24% switched to a new entrant.³ Hence, some two decades after liberalization the former incumbent still prices well above costs and the market is characterized by substantial price dispersion.

A key aspect of our dataset is the fact that we have data on consumer search

¹This is also true for electricity markets around the world. See, e.g., Cabral (2017) for evidence related to different European countries and Hortacsu et al. (2017) for evidence on the USA.

²According 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.

³See the 2015 report of the German regulatory authority (Bundesnetzagentur, 2015).

intensity at the zip code level. In particular, we have data on the actual number of search queries at online price comparison platforms, and given that most of the search for lower prices is via these platforms, we interpret these data as a direct measure of search intensity at the local level. Such data are generally not available and other empirical studies on consumer search markets often rely on indirect measures of consumer search activity.⁴ Moreover, we are not aware of any paper having a direct measure of consumer search intensity in a panel data context.⁵

In terms of prices, we have access to the baseline and online price of the incumbent and the lowest online price of an entrant at the zip code level. Using these data, we empirically show that more consumer search tempts incumbents to increase their baseline rates and entrants to decrease their tariffs. Moreover, the incumbent increases the extent of price discrimination and sets its cheapest online tariff much lower when consumers search more. Price discrimination allows the incumbent to segment markets according to search intensity, and to simultaneously appropriate surplus from loyal consumers and prevent the entrant from stealing the business of many searching consumers. Search intensity and pricing are likely to affect each other simultaneously as search intensity may be a function of prices (e.g. Tappata, 2009; Yang and Ye, 2008) and retailers' pricing strategies may take into account the consumers' level of information (i.e. their search efforts). We address the potential reverse causality of pricing on search intensity by using instrumental variables.

The theoretical model we build serves two purposes. First, it shows that the empirical patterns can be easily understood as resulting from the interaction of the different incentives the incumbent, entrants and consumers have. Second, by means of the model, we are able to evaluate whether consumers are better or worse off because of price discrimination. The model has the following elements. Retail firms produce a homogeneous good so that consumers only search for better prices. Consumers differ in search and switching costs. 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. In the context of the German electricity market, or other markets where price comparison websites are available, one can think of search as consulting a platform so that at a cost consumers have ac-

⁴Many papers proxy for lower search costs or consumers' level of information by access to the internet or online versus offline purchases (Brown and Goolsbee, 2002; Brynjolfsson and Smith, 2014). Similarly, Pennersdorfer et al. (2015) use commuters versus non-commuters to distinguish between informed and uninformed consumers in the retail gasoline market and Chandra and Tappata (2011) use the distance between gas stations as a proxy for consumer information.

⁵While some papers have direct measures for search intensity, these measures are time series rather than panels. For instance, Byrne and De Roos (2017) have local price data for gasoline stations but their measure of search intensity is a time series of web traffic at the government's transparency site (Fuelwatch). Similarly, Lewis and Marvel (2011) also use local gasoline prices but use a time series of web traffic at a website that reports gasoline prices (Alexa).

cess to all prices available. 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, including the ones we find in the empirical part of the paper, namely that if the fraction of consumers with small search costs increases, price dispersion and price discrimination increase simultaneously with the fraction of consumers who search online, while the incumbent raises its baseline price to consumers who do not search. By means of our model, we can also analyze the counterfactual situation where the incumbent is prohibited from price discriminating between online and offline marketing channels. We show that searching consumers are better off under price discrimination, while “loyal” consumers are worse off. For a large class of search cost distributions, the latter effect dominates so that on average, consumers are worse off because of price discrimination.

The German electricity market is just one example of a market where important market liberalizations have taken place over the last 20–30 years. Many of these markets (including electricity markets in many states of the USA⁶ 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 German electricity market. In all these markets, new firms have entered, but because consumers buy on 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 have 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.⁷ 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. In most of the other markets mentioned (e.g. railways, postal services or airlines), the natural incumbency effect is likely

⁶Some 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).

⁷There 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.

to be much smaller as consumers do not have an ongoing contract for continuous delivery.⁸

As the focus of the paper is on the impact of consumer search on price dispersion and price discrimination in liberalized markets, it contributes to different strands of literature. There is a large and varied theoretical literature on how consumer search affects price dispersion in homogeneous goods markets (Stahl (1989), and Janssen and Moraga-González (2004) for theoretical contributions). Most of the theoretical models are based on random search, whereas in liberalized markets, it is important to account for the incumbency effect as consumers are served by their local incumbent at the baseline tariff unless they proactively search for and switch to a cheaper alternative tariff.

Several empirical studies focus on price dispersion and search intensity (see, e.g., Sorensen (2000), Chandra and Tappata (2011), and De los Santos et al. (2012) for empirical studies). An early example is Brown and Goolsbee (2002) who use the variation in internet usage 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. 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. (2015) 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. However, the above literature does not explicitly deal with the incumbency effects and the resulting asymmetry between firms that is important to understand the functioning of liberalized markets. More precisely, the literature generally assumes that the shares of informed and uninformed consumers are randomly assigned between firms which clearly does not hold true for markets with an incumbency effect.

⁸Our findings are also of importance to other sectors which have not been liberalized recently, such as retail banking, where consumers who search (online) may get much better deals than loyal consumers. For example, in the recent article “American banks pay depositors less than online accounts”, in *The Economist* (see, *The Economist* (2018/02/15) at <https://www.economist.com/news/finance-and-economics/21737031-they-seem-be-relying-power-inertia-retain-their-customers-a-risky>) it is stated that established U.S. banks generally offer substantially lower interest rates on savings accounts compared to online rates offered to clients at internet portals and that the gap widens significantly. And more interestingly, the established banks even lowered the offered interest rates while online banks drastically increased them in recent years. While for established banks, the offered interest rate dropped on average from 0.11% in 2014 to only 0.09% in February 2018, online customers’ rates rose from 0.87% to 1.52% in the same period. These patterns are attributed to differences in search behaviors. As in our case, online offers are partly generated by firms that have only online operations, but also by firms (banks) that offer different prices on different retail channels. Our approach focuses, however, on single (local) incumbents competing with entrants, whereas in the banking sector there are many “incumbents” in the same market.

A growing literature deals explicitly with search in electricity markets, but most of these papers focus on how consumers search. Giulietti et al. (2014) analyze the retail electricity market in the United Kingdom and find that roughly half the households had relatively high search costs. In contrast to our paper they analyze the 2002–2005 period where consumers searched sequentially as information on suppliers’ prices was most commonly obtained from a salesperson (for a particular supplier) coming to the customer’s door. Hortacsu et al. (2017) analyze switching in the Texas retail electricity market and find that even though households generally switch to alternative retailers only rarely, they do switch more after experiencing a “bill shock”. Moreover, they also find that households attach a brand advantage to the incumbent. In contrast to our study, both papers do not observe the actual search behavior of consumers.

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 that give benefits to high-valuation consumers in particular. For the same sector, Miravete (2011) finds that incumbents – as a response to entry – temporarily increase the foginess of their tariffs relative to the monopoly phase, while entrants use foggy pricing far less frequently. In the paper, foggy pricing refers to the practice of offering tariff options that are always more expensive than other offered tariff options regardless of usage profile, aiming to obfuscate consumers with high search costs and selling them contracts with deceptive prices. For the US airline industry, Goolsbee and Syverson (2008) 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 non-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) sets his argument, 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 Gilbuhk (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 behavior between incumbents and entrants is affected by the presence of more searching consumers.

Our paper explains that an 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 appropriate surplus from consumers with high search cost and prevent entrants stealing the business of consumers with lower search cost. This pricing strategy prevents many consumers from switching supplier and at the same time allows the incumbent to appropriate market power rents.

The rest of the paper is structured as follows. The next section describes the German retail electricity market in more detail. We then investigate the search-pricing nexus in the German retail electricity market. Section 3 describes the empirical identification strategy, section 4 discusses the data, and section 5 presents the econometric results and shows the robustness of the results to alternative specifications. Section 6 provides a theoretical model explaining the empirical findings and evaluating the welfare effects of price discrimination. Section 7 concludes.

2 The Retail Electricity Market

In 1999 Germany's electricity liberalization brought about the end of local monopolies by allowing entry to local markets and giving consumers the choice to switch from their former local incumbent to alternative suppliers. It was believed that increased competition and freedom of consumer choice would eventually result in lower retail margins with large economic benefits for consumers. Prior to market liberalization, the local incumbent served all customers in its distribution grid area at a regulated tariff. After liberalization, the incumbents are legally obliged to supply electricity at a baseline tariff to all households which do not proactively choose another supplier.⁹ However, the incumbents' baseline tariffs are no longer regulated

⁹There is always only a single incumbent at each address which is defined by law. By law, the incumbent is defined as the local electricity retailer with the largest customer base. In theory, a different retailer may become the new incumbent because of that. However, due to the very low switching rates this is never the case in practice and the original incumbent almost never changed. The only exceptions where the incumbent changed were due to a few mergers of municipal utilities in the past two decades.

and households are free to switch to any alternative tariff offered for their local address, either by one of the many new entrants or by their local incumbent.

Retail entry in the German electricity market can be characterized as fairly easy since it follows a regulated, non-discriminatory procedure, and involves low entry costs and risks. This is also witnessed by the large number of active firms: there are on average 133 firms per zip code area with a range of 55 to 192. There are several reasons for local retail entry, such as (i) direct electricity contracts with local electricity producers;¹⁰ (ii) vertical integration, so that operations happen around the locations of the power plants; and (iii) local bureaucracy, such as the obligation to become a member of the regional balancing group coordinator in its service area.

Another important characteristic is that all retailers that compete in a certain area have almost identical costs. Some cost components like grid charges and concession fees differ not only over time but also at a very local level. However, they are equal for all retailers in the local market. Other cost components like subsidies for renewable energies only change over time but do not have local variation. Costs for purchasing wholesale electricity are also almost identical across retailers since wholesale electricity prices are determined centrally at the European Energy Exchange (EEX).¹¹ Some other costs, such as administrative or advertisement costs, may differ across retailers but only account for a very minor part of (variations in) the retail costs. Thus, while costs are very similar for all retailers in the local market, due to locally varying cost components, the costs vary substantially across local markets in Germany.

Many incumbents operate only at a very local level and 46% of the incumbents only have a single zip code in their incumbency area. These small incumbents are mostly municipal utilities. Larger incumbents often have several zip codes in their incumbency area. The incumbency areas of incumbents with more than one zip code cover 5 zip codes at the median and 32 at the mean. Hence, as the costs differ at a very local level, they also often vary within the same incumbency area, and so do the prices: incumbents with more than one zip code set on average 3.5 different prices in their incumbency areas. For incumbents with more than one price zone, the prices within the same incumbency area differ on average by 10.4 Euro/3.5 MWh.¹² To

¹⁰If a firm operates in the whole country, it has to meet its demand (and demand fluctuations) throughout Germany. For this reason, such a company may have to operate at the power exchange to fulfill its delivery commitments. Operating at the power exchange, however, is associated with operational, managerial and transaction costs (e.g. it may necessitate having a trading department, which may be costly). Thus, smaller providers may prefer direct contracts with local electricity producers, and thus only operate locally.

¹¹Even if firms buy electricity through direct contracts with electricity producers, the spot price still represents the opportunity costs of purchasing electricity.

¹²The largest observed difference of the base tariffs within the same incumbency area of 134 Euro/MWh was in 2012 by E.ON Avacon Vertrieb GmbH, which served 189 zip codes with 14 different price zones.

give an example, in 2011 the incumbency area of Envia Mitteldeutsche Energie AG, a relatively large incumbent operating in 312 zip codes, had 25 different price zones with prices ranging from 930 to 982 Euro/3.5 MWh. Appendix Figure A8 illustrates the base tariffs set by Envia Mitteldeutsche Energie GmbH within its incumbency area.

At the start of the liberalization process, all consumers were automatically supplied by their local incumbent and paid the baseline tariff. However, German law grants consumers the right to switch away from the *incumbent baseline* tariff at any time with two weeks' notice. Consumers who switch from the baseline tariff generally make a one-year contract with their new supplier, which is automatically renewed if the consumer does not cancel the contract in time.¹³ Thus, once a consumer has decided to switch, the costs related to being locked in are higher.

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), Energieverbraucherportal.de (energy consumption portal) and mut-zum-wechseln.de (courage to change). Verivox started to provide search services in electricity in 2000, Toptarif in 2007, and Check24 in 2008. Despite this fairly recent trend of searching through platforms, in 2011 80% of switchers had already searched online for alternative providers, according to a survey.¹⁴

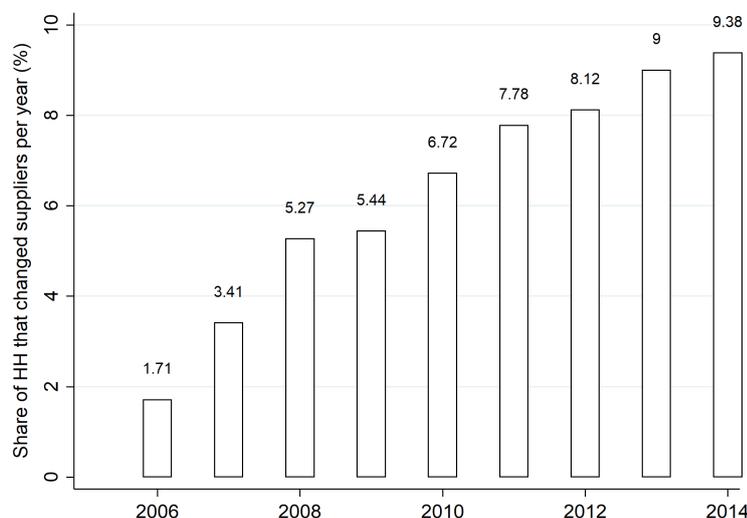
The switching rate has been growing in recent years (see Figure 1), as online price comparison platforms have significantly reduced the costs of searching for cheaper providers (something that is also acknowledged in other markets; see e.g. Bar-Isaac et al., 2012). A comparison portal requires a consumer to enter all relevant details (zip code, expected yearly electricity consumption, whether the contract is for private or commercial use). 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 year¹⁵ and whether or not the listed tariffs should include one-off bonuses. The platform then lists the "personalized" prices of all providers that are active in the indicated zip code ranked from lowest to highest. For each tariff, the

¹³Consumers may also opt for monthly contracts in some tariffs, but this is rarely done. According to a market report by the German regulatory authority Bundesnetzagentur (2013, p. 150), the average contract period is 10 months, suggesting that the majority choose yearly contracts. During the term of the contract, consumers can only switch if their supplier changes the price or they move to another zip code in which their current retailer is not active. Consumers then have an extraordinary termination right.

¹⁴http://www.atkearney.at/documents/3709812/3710656/BIP_Der_Strom_und_Gasvertrieb_im_Wandel.pdf, March 3, 2016.

¹⁵Consumers have an extraordinary termination right if their retailer adjusts the price. Retailers may also adjust prices if they are advertised as "guaranteed", e.g. retailers are allowed to adjust prices if the EEG surcharge or grid fees change even if they offered a price guarantee.

Figure 1: Average switching rates of households in German retail electricity markets



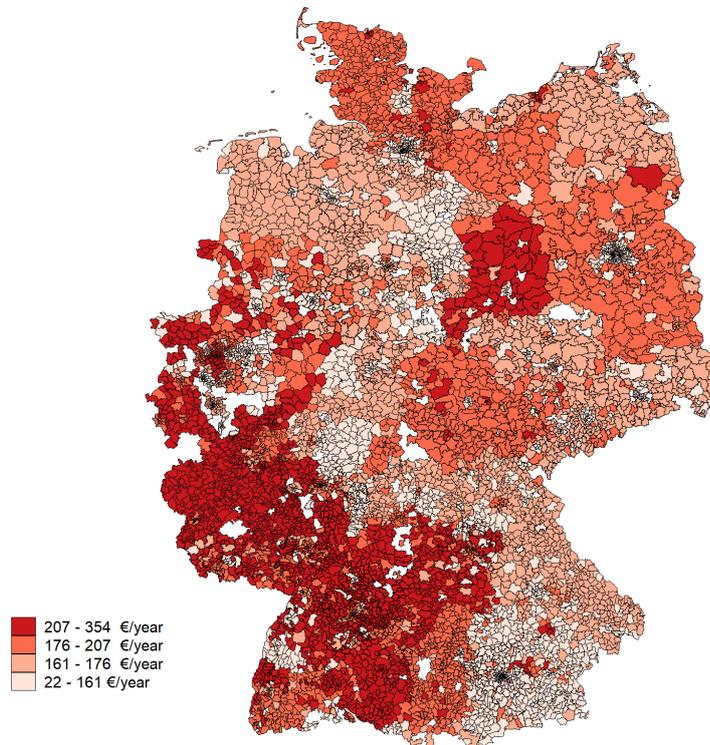
Note: Data on supplier changes are obtained from Germany’s regulatory authority (Bundesnetzagentur, 2015), data on the number of German households are from the German Federal Statistical Office. Switching rates include not only switching from the incumbent to an entrant but also to a large extent switching among entrant suppliers.

platform also provides information on how much consumers can save 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, but for all consumers who are familiar with online shopping, the search costs seem relatively small compared to the potential savings of switching from the incumbents’ baseline tariff to the overall cheapest tariff, which are on average almost 200 Euro per year for a standard two-person household with 3,500 kWh consumption. The potential gains from search range from 77 to 354 Euro per year, depending on the households’ location, as shown in Figure 2 for the year 2012.

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 and registration, at no additional cost.¹⁶

¹⁶In many other countries, the switching process for electricity providers is comparable to the one in Germany’s retail electricity markets. 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.” Similarly, Hortacsu et al. (2017) mention for the Texas retail electricity market that switching is a relatively easy process (without monetary costs) and that search costs are the fundamental

Figure 2: Potential gains from search (2012)



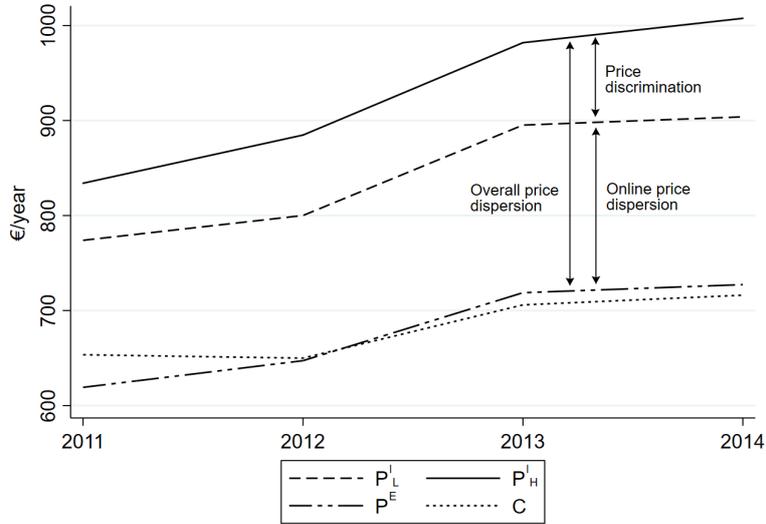
Finally, retail electricity can be considered a fairly homogeneous product¹⁷ and for the consumers there are no retailer specific differences regarding the quality of supply. If an entrant fails to deliver, the incumbent provider has the legal obligation to guarantee the continuing electricity supply at the baseline tariff for the affected consumers without any temporary interruption. However, not all consumers may be aware of this safety net.¹⁸ Hence, theoretically it should not matter for the end-consumer which retailer delivers the electricity, although to some extent it still may matter in practice, equipping the incumbent with some brand value.

driver of consumer inertia. In contrast, in other markets switching fees may be substantial. For example, in the US cable TV market, consumers have to pay for upfront installations and equipment (Shcherbakov, 2016).

¹⁷The main product differentiating feature is the differentiation between regular tariffs and “green” tariffs, which exclusively sell electricity produced by renewable sources such as wind, solar or run-of-river plants. In the empirical analysis, we omit consumers that only search for green certified tariffs. However, such consumers only present a very small share of all searching consumers (approximately 3%) and our results stay fully robust once we include them.

¹⁸The 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. To date, there have been only two insolvencies of German incumbents (Stadtwerke Gera and Stadtwerke Wanzleben, both in 2014). Consumers were not affected, at least insofar as electricity supply continued during the insolvency procedures. Hortacsu et al. (2017) mention the possibility that customers may believe that the incumbent supplier may exhibit a higher supply security although this is in fact not true.

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



Note: Prices (P_H^I , P_L^I , P^E) and costs (C) are presented net of value added taxes.

As other prices than the incumbent base tariff can only be observed by consumers who proactively search, an incumbent is able to set alternative tariffs online for consumers who engage in search, and in this way price discriminate between consumers with low and high search costs. The incumbent’s online tariff is typically set below the baseline tariff but still above the cheapest overall tariff set by an entrant. Figure 3 shows that there are considerable price differences between the incumbent’s baseline tariff (i.e. price incumbent high, P_H^I), the incumbent’s cheapest tariff (i.e. price incumbent low, P_L^I), and the overall cheapest entrant tariff (i.e. price entrant, P^E). We focus on the cheapest entrant price, assuming that consumers who switch away from the incumbent most likely choose the cheapest tariff available.¹⁹ As a result, we observe three forms of price dispersion: (i) Overall *price dispersion* ($P_H^I - P^E$), which is the difference between the incumbent’s baseline tariff and the overall cheapest tariff; (ii) *price discrimination* by the incumbent ($P_H^I - P_L^I$), measured by the difference between the incumbent’s baseline tariff and the incumbent’s cheaper online tariff; and (iii) *online price dispersion* measured by the difference between the incumbent’s cheaper tariff and the cheapest entrant tariff ($P_L^I - P^E$).²⁰

¹⁹Our assumption that the majority of consumers may switch to the cheapest price of an alternative supplier (given their brand loyalty is low enough to switch away from the incumbent) is also supported in the theoretical and empirical literature. Gulietti et al. (2014) state that “in the absence of search costs, someone who switches will choose the cheapest alternative, other things equal.” Hence, at a tariff comparison platform, a consumer faces a list of all available tariffs starting with the cheapest and thus may also choose the cheapest tariff. Corroborating empirical evidence stems, e.g., from Baye et al. (2006), who estimate that an online retailer selling handheld PCs – a product that is significantly more differentiated than retail electricity – receives 60% more clicks by being cheapest relative to being second-cheapest on a UK online comparison platform (i.e. Kelkoo.com) in 2003.

²⁰We employ the price range as our dispersion measure, which is a commonly used measure in

Figure 3 also depicts the (approximated) costs of retailers.²¹ We see that costs and prices have increased over time (mostly due to increased taxes and levies to finance the integration of renewables). Evidently, even nearly two decades after the retail liberalization in the industry, the incumbent baseline tariff remains well above costs. Moreover, the figure emphasizes that incumbents price discriminate by setting a high baseline price for their consumers with high search costs and a lower price for consumers with lower search costs who may consider leaving. However, we observe that the cheaper incumbent’s price is still well above costs. By contrast, the cheapest tariffs set by a entrants are very close to marginal costs. Despite having a liberalized market with free consumer choice for nearly 20 years, a large share of consumers (around 76%, as stated in the introduction) stays with an expensive incumbent tariff and tariffs remain highly dispersed.

3 Identification of the Effects of Consumer Search on Pricing

To examine the effect of consumer search intensity on pricing strategies, we first explain our identification strategy and then describe our data and results. The relationship we are interested in can be described by a model of the form:

$$\ln(y_{it}) = \beta \ln(\mu_{it}) + \gamma \mathbf{x}_{it} + \delta_i + \eta_t + \epsilon_{it}, \quad (1)$$

where the subscripts i and t indicate zip codes and years, respectively. The dependent variable y either denotes an electricity tariff (P_H^I , P_L^I , P^E) or a price difference measure ($P_H^I - P^E$, $P_H^I - P_L^I$, $P_L^I - P^E$) 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 such as 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 not 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

the literature (Baye et al., 2006). In our case, the price range best reflects the potential gains from search.

²¹In Section 4, we discuss our approach to approximate retailers’ costs.

search impacts the dependent variable in percentage terms. Assuming that search patterns at other major comparison websites are 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.²² Hence, our parameter of interest β reflects the percentage change in pricing behavior for a one percent change in search activity.²³

As consumers may search more if they expect higher gains from searching when prices are high and/or more dispersed (e.g. Tappata, 2009; Yang and Ye, 2008), there potentially is a reverse causal relationship between prices and search intensity. Therefore, we instrument for consumer search intensity μ using two instruments. Our first instrument is the share of young households in a zip code measured as the share of households with a household head below the age of 40 ($U40$). The underlying idea is that younger people are probably more familiar with the internet in general and with online shopping in particular. As a second instrument, we employ the regional availability of broadband internet (BBA), measured as the share of households in a zip code for which an internet speed of 16 Mbit/s is offered for their local address. The idea here is that fast internet makes online shopping more convenient.²⁴

Using these instruments, we estimate the following linear projection of μ in the first-stage:

$$\ln(\mu_{it}) = \alpha \mathbf{z}_{it} + \gamma^{FS} \mathbf{x}_{it} + \delta_i^{FS} + \eta_t^{FS} + u_{it}, \quad (2)$$

where the vector \mathbf{z} consists of our two instruments $U40$ and BBA . The superscript FS indicates that the parameters are from the first-stage estimation. Plugging the first-stage prediction of search (i.e. $\widehat{\ln(\mu)}$) into Equation (1) should yield unbiased estimates for β .

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. retailers may face a higher risk of shortfalls of payments by younger customers) we also control for income. As income is included in \mathbf{x} , we assume that the instrument vector \mathbf{z} is uncorrelated with the

²²The assumption that the searching patterns we observe are similar across platforms is reasonable: We also have data on consumer search at *Verivox* – another major price comparison platform in Germany – for the year 2014; however the data are only provided as percentages of search in a respective zip code relative to the overall search in Germany, which is why we cannot merge these data with our actual search data. We find a correlation coefficient of 85% between search intensity at *Verivox* and the platforms we use here, indicating that search at *Verivox* does not seem to differ much from search observed in our dataset.

²³However, the estimations are fully robust if estimated in levels, as we show later.

²⁴We will also experiment with alternative internet speeds in the robustness section.

error term ϵ in Equation 1 such that

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

Moreover, it may be that broadband internet availability affects electricity demand or that younger households consume more or less than older households. However, we believe that these demand effects do not translate into direct pricing effects in retail electricity markets since electricity demand is highly inelastic. To tackle remaining doubts regarding our identification strategy, we apply an alternative set of instruments in the spirit of Hausman and Taylor (1981). That is, we use average values of the instruments in the 50 surrounding zip codes (while disregarding the actual zip code i) as alternative instruments. Thus, even if our instruments were to directly affect prices in some way, their values from neighboring zip codes should not, although they should still be correlated with the values in the respective zip code itself.²⁵ The results are fully robust to these alternative instruments, as we show in the robustness section.

4 Data and Variables

We use panel data at the German zip code level for the period 2011–2014.²⁶ A key asset of our data is the information about local online search queries at major price comparison portals, which enables us to construct a direct measure of consumer search. Another particular advantage is that we do not have to make assumptions about market delineation, as consumers at retail electricity markets can only choose among electricity suppliers that sell to their local address.

The data stem from four sources. From *ene't*, a German software and data provider for the electricity industry, we received detailed data on individual consumer searches at several online price comparison sites as well as retail electricity tariffs and cost components. The database marketing company *Acxiom* provided data on structural household characteristics in Germany. We also use data from the *European Energy Exchange* (EEX) to obtain a proxy for the purchase costs of wholesale electricity (as a component of utilities' total costs). Moreover, we use data

²⁵The share of young households (U40) in neighboring zip codes should be correlated with the zip code of observation due to demographic effects (e.g. trends in migration, existence of a university, etc.). Also, broadband availability (BBA) in zip code i is correlated with BBA in neighboring zip codes, e.g. due to regional roll-out.

²⁶We have 8,226 zip codes in our data. However, there is an overlap of incumbency areas in some of the zip codes. For instance, A is the incumbent in one part of a zip code while B is the incumbent in the other part of the same zip code. We drop all zip codes which have more than one incumbent, which reduces the number of zip codes in our data to 7,249. However, the results we present later only change marginally if we also include these zip codes into our data and use averages of the different incumbent prices.

Table 1: Summary statistics

Variable	Unit, data source	Mean	SD	Min	Max	Obs
<i>Dependent variables</i>						
Incumbent's baseline price (P_H^I)	€/a, ene't	1,006	78.1	800	1,204	25,861
Incumbent's cheapest tariff (P_L^I)	€/a, ene't	929	86.7	716	1,117	25,861
Entrant's overall cheapest tariff (P^E)	€/a, ene't	808	59.0	667	903	25,861
Overall price dispersion ($P_H^I - P^E$)	€/a, ene't	199	38.9	77	354	25,861
Price discrimination ($P_H^I - P_L^I$)	€/a, ene't	77	41.6	0	282	25,861
Online price dispersion ($P_L^I - P^E$)	€/a, ene't	122	46.5	0	259	25,861
<i>Endogenous variable</i>						
Consumer search intensity (μ)	%, ene't	9.23	6.3	0	34.7	25,861
<i>Instruments</i>						
Head of HH below age of 40 (U_{40})	%, <i>Acxiom</i>	24.7	5.2	8.4	55.1	25,861
Broadband internet availability (BBA)	%, <i>breitbandatlas.de</i>	63.8	32.3	0	100	25,861
<i>Control variables</i>						
Net costs	€/a, ene't & <i>EEX</i>	683	42.3	560	823	25,861
Competitors	#, ene't	133	25	55	192	25,861
Average HH size	%, <i>Acxiom</i>	2.11	0.2	1.52	2.54	25,861
Income <25k €/a	%, <i>Acxiom</i>	0.39	0.1	0.02	0.83	25,861

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

on regional broadband internet availability from *breitbandatlas.de*.

As mentioned in the section above, consumers typically have yearly contracts and we therefore aggregate all data at the yearly level. Table 1 provides the summary statistics of the variables in our regressions. In Table A1 in the Appendix, we additionally report the between and within standard deviations of our key variables.

Prices

ene't provided all tariff data for P_H^I , P_L^I and P^E . For P_H^I and P^E we have monthly data. We use the yearly averages as our variable for these prices. For P_L^I the initial data are yearly average prices already. In the estimations we use gross prices (including 19% VAT) since they present the relevant price for end-consumers, and is displayed accordingly in the search platforms. 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.²⁷

Table 1 shows that, on average, a household with 3,500 kWh annual consumption of electricity pays around 1,006 EUR per year if it stays with the incumbent's default

²⁷3,500 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_H^I and P^E (not for P_L^I). Regression estimates using P_H^I and P^E as well as $P_H^I - P^E$ for alternative consumption levels yield robust results.

tariff. The incumbent’s cheaper tariff is around 8% lower at 929 EUR, while the overall cheapest entrant tariff is around 808 EUR (which is 20% cheaper than the incumbent default tariff). All prices exhibit substantial variation in our observation period regarding their spatial and time dimensions. Figure A2 in the Appendix shows the spatial distribution of the incumbents’ baseline tariffs for the year 2012. From the figure it can be seen that prices are generally higher in eastern and northern Germany. The main reason is that grid charges are higher there because of the many wind farms in these areas. Figure A4 in the Appendix shows the spatial cost distribution. We accommodate 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 code fixed effects. Notably, the price data also provide substantial temporal variation. This is shown for the incumbent base tariffs in Appendix Figure A3 but is also very similar for the other tariffs. The between and within standard deviations of the prices are also reported in A1 in the Appendix. Moreover, there is significant variation in the identity of the entrant offering the lowest tariff across zip codes and time. In our observation period, there are 182 different retailers that have the cheapest overall price in a zip code for at least one month. Moreover, in a zip code there are on average 14.4 different retailers that have the cheapest price in at least one month during our observation period. On average a retailer is the cheapest for 2.2 consecutive months. Overall this suggests that there is substantial heterogeneity and variation over time in the cheapest firm.²⁸

Consumer Search

ene’t provided data on individual online search queries, enabling us to construct data on regional consumer search intensity. The database covers detailed information on all search queries conducted at several well-known online price comparison platforms including Toptarif.de (top-tariff), Stromtipp.de (power-hint), Energieverbraucherportal.de (energy-consumer-portal), and mut-zum-wechseln.de (courage-to-change), of which Toptarif.de is by far the largest platform.²⁹ 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 interface, the type of search query (household or industrial customer), consumer preferences (e.g. whether a consumer only wants to get energy tariffs having a “green” certificate displayed), as well as a search session ID indicating the order of the queries from the same con-

²⁸This is in line with the following possible dynamic pricing strategy of entrants: they may try to attract consumers using the lowest price in the market and then raise their prices in subsequent periods, hoping consumers will not switch again.

²⁹Toptarif 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.

sumer.³⁰ 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.³¹ Since our focus is on household consumers, we disregard search by industrial consumers. Furthermore, we exclude 551,256 search sessions, which exclusively consider eco-label (i.e. “green”) certified tariffs.³² Those searches are most likely not predominantly price driven but related to product differentiation and, on average, €152 more expensive than the cheapest tariff.

We construct our measure of search intensity as the number of search sessions within a zip code per year divided by the number of households:

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

Since we observe some extreme outliers in some zip codes, apparently resulting from price comparing software “bots” or data crawling researchers, we truncate 2% of the upper bound of the sample distribution of our consumer information measure.³³

At the mean, 9.1% of households within a zip code search for retail tariffs at one of our sample comparison websites, whereas there is substantial variation ranging from 0% to 34.7%. An illustration of the spatial variation of search intensity is given in Appendix Figure A5. Moreover, Table A1 shows that there is also substantial within zip code variation.

Instruments

The data on the share of young households, i.e., households with a household head younger than 40 (*U40*), are obtained from *Acxiom* at a zip code-year resolution. Data on local broadband internet availability (*BBA*), which is the share of households for which internet speeds of 16 MBit/s or higher are offered, are obtained from

³⁰We 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.

³¹It 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 date. Therefore, we treat each search session as conducted by an individual household.

³²Nevertheless, our results are fully robust to the inclusion of eco-label searches.

³³Figure A6 in the Appendix provides a histogram on consumer information before and after trimming the data.

Breitbandatlas.de with the same resolution. The *BBA* data are at the municipal level and we aggregate them to the zip code level to match them with our remaining data. We choose the 16 MBit/s as it provides substantial variation and gives the best first-stage F -statistic of excluded instruments. A speed of, for example, 1 Mbit/s was available in more than 90% of all zip codes in 2011 and is actually too slow to use the internet conveniently. A speed faster than 16 Mbit/s has a very low coverage and for most applications does not make internet usage significantly more convenient than 16Mbit/s; at least not during the period 2011–2014. An alternative would have been to use 6 Mbit/s. In this case, the first-stage F -test is a little lower but the results do not change much.³⁴ *BBA* exhibits considerable within zip code variation while it is less pronounced for $U40$, as shown in Table A1.

Control Variables

We compute a variable reflecting retailers’ costs in order to control for regional cost differences. Detailed data on cost components are primarily obtained from *ene’t* and include, for example, grid charges, concession fees, renewable energy surcharges (“EEG Umlage”), CHP (combined heat and power) 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 time. 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 one-year ahead price presents the standard purchasing strategy for retailers.³⁵

To measure competition within the zip code, we use the number of electricity retail suppliers in the zip code, as provided by *ene’t*. The number of competitors in a zip code varies between 55 and 198 in our observation period.³⁶ Moreover, as can be inferred from Table A1 in the Appendix, the within zip code standard deviation is 19.8, suggesting that there is also a significant fluctuation of retailers within a zip code over time.

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

³⁴These results are available in the Appendix Tables A23 and A26.

³⁵Even if retailers purchase electricity through other channels than via the power exchange (e.g. bilateral contracts, OTC, etc.), the price from the power exchange still represents the opportunity cost.

³⁶These numbers may seem high but correspond well with Bundesnetzagentur (2015) and may be the consequence of low entry barriers in the electricity retail market in Germany.

with an annual income of less than €25,000 (Income <25k €/a) may be relevant, as low incomes may increase the risk of a consumer’s payment default.

5 Results

In Table 2, we present the results from our IV estimations for the three retail prices of interest, P_H^I , P_L^I and P^E . As we do not have information on all price comparison sites, we use a log-log specification to be able to interpret the coefficients as elasticities.³⁷ The instruments are sufficiently strongly correlated with the endogenous variable as shown by the high values of the first-stage F -test. Also, the Durbin-Wu-Hausman test suggests that the consumer search intensity μ should indeed be treated as endogenous as the null hypothesis of consumer search being an exogenous regressor is clearly rejected. 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 A2 in the Appendix.³⁸

Coming to the results, column 1 of Table 2 provides evidence that the incumbent reacts to a higher search intensity by slightly increasing its baseline tariff. For a change in consumer search intensity by 10%, the incumbent raises its tariff by 0.38%. Thus, while we theoretically show in Section 6 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. Column 2 shows in contrast that the incumbent reacts to more search activity in its zip code by reducing its cheaper online tariff considerably. For a 10% increase in search activity, the incumbent decreases its cheapest tariff by 1.16%. 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 with the incumbents’ cheapest tariffs. For every 10% increase in search intensity in a zip code the overall cheapest tariff in the market decreases by 0.32%. Thus, the incumbents’ cheapest tariffs react more strongly to consumer search than the overall cheapest tariff.

The empirical effects can be explained along the lines of Proposition 2, presented in the next Section discussing the theoretical model. If there are more low search cost consumers, then there will be more competition online and thus lower prices. To prevent too many consumers from switching to the entrant, the incumbent has to decrease its online price more aggressively than the entrants: the incumbent would

³⁷However, the results are fully robust to level-level specifications as shown in Table A7.

³⁸For information purposes, we also provide results from OLS panel estimations in Tables A3 and A4 in the Appendix. The direction of the sign and the significance are as in the IV estimations but the magnitudes are much lower, suggesting that neglecting endogeneity leads to an underestimation of the impact of consumer search on pricing.

Table 2: IV estimates of the impact of consumer search on prices (log-log)

	(1) Incumbent Base: P_H^I	(2) Incumbent Cheapest: P_L^I	(3) Overall Cheapest: P^E
Search (μ)	0.038*** (0.009)	-0.116*** (0.025)	-0.032*** (0.007)
Costs	0.238*** (0.011)	0.474*** (0.026)	0.531*** (0.009)
#Competitors	0.000*** (0.000)	0.005*** (0.000)	0.000*** (0.000)
Average HH size	0.024*** (0.008)	0.054** (0.023)	0.005 (0.006)
Income <25k €/a	0.009 (0.008)	-0.069*** (0.018)	-0.007 (0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F stat.	23.67	23.67	23.67
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

lose a larger markup when losing a customer as the incumbents' cheapest prices are higher than the overall cheapest prices in the market. At the same time, if there is still a considerable fraction of consumers with high enough search costs, the incumbent has an incentive to increase the margin on its baseline tariff as it will not lose too many consumers by doing so relative to the loyal consumers.

We now briefly explain the impact of the other control variables. We estimate the cost pass-through to the end-user retail tariffs, which is much higher in the competitive segments of the electricity retail market. For the incumbents' baseline tariffs, we estimate a pass-through of only around 24%, whereas 48% of cost increases are passed on to consumers for the incumbents' cheapest tariffs and 53% for the cheapest entrants' tariffs. The 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 costs to a lesser extent.

Average household size increases all tariffs. The reaction of the incumbent's baseline tariff to a larger share of low-income households ($Income < 25k \text{ €/year}$) is statistically insignificant, whereas the incumbent reacts with a significant decrease in its cheaper online tariff. To a lesser extent, the same is true for the cheapest entrant. This may imply that price sensitive consumers may not be willing to pay

a high online price 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 online price.

Possibly surprisingly, the number of competitors positively affects all prices. However, the estimated effect of an additional competitor on prices is close to zero and thus economically negligible. Given low entry barriers and the already high number of competitors, it seems that additional entry does not have an economically significant impact on prices.³⁹

Table 3 presents estimations of the impact of consumer search on the three price dispersion measures. Column 1 focuses on *overall price dispersion*, measured as the incumbent’s baseline tariff (P_H^I) minus the overall cheapest tariff (P^E). Evidently, price dispersion goes up if more consumers search, since the incumbent slightly increases its baseline tariff and at the same time the overall cheapest price declines with search: more search in a zip code region may indicate a larger relative mass of consumers with low in contrast to intermediate search costs, which prompts the incumbent to increase its baseline tariff, while tariffs in the competitive segment decline due to increased competitive pressure. For every 10% increase in search intensity, the extent of price dispersion goes up by 2.7%, 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 searching consumers, which is still above the overall cheapest tariff in the market, and a high incumbent baseline tariff for loyal consumers who do not search. Price discrimination becomes more pronounced with increasing search intensity. An increase in the share of searching consumers by 10% widens the gap between the incumbent’s baseline tariff and its cheaper tariff by 13.7%. 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. This can be explained in line with Proposition 1 of our theoretical model in the next section: more searching consumers imply more price discrimination if there are relatively sufficient many consumers left who are loyal and “always” buy at the baseline price of the incumbent.

We also see that *online price dispersion*, measured as the difference between the incumbent’s cheapest tariff and the overall cheapest tariff in the market, narrows

³⁹Regardless of this, these coefficients have to be interpreted with caution, as they may not reflect causal effects: prices and number of competitors may influence each other. In the robustness section, we also present results from estimations where we additionally instrument for the number of competitors using Hausman-type instruments (Hausman and Taylor, 1981) but the results only change marginally. Moreover, all results stay fully robust if we drop all covariates, as shown in Tables A5 and A6 in the Appendix.

Table 3: IV estimates of the impact of consumer search on dispersion (log-log)

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.272*** (0.057)	1.368*** (0.340)	-1.361*** (0.356)
Costs	-0.967*** (0.067)	-1.911*** (0.325)	-0.100 (0.360)
#Competitors	0.002*** (0.000)	-0.042*** (0.002)	0.066*** (0.003)
Average HH size	0.107** (0.055)	0.064 (0.294)	1.839*** (0.328)
Income <25k €/a	0.093 (0.065)	0.750*** (0.240)	-0.948*** (0.274)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F stat. F stat.	23.67	23.67	23.67
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

considerably with search intensity. The more consumers search in a market, the more the incumbent is forced to approach the overall cheapest price. For a 10% increase in search intensity, the “loyalty premium” narrows by 13.6%. 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. These results are in line with Proposition 3 below, where we explain that in zip codes where more consumers search the incumbent decreases its online price more aggressively than the entrants as he would lose a larger markup when losing a customer. This implies increased competition and less online price dispersion.

Overall, we find that the loyal consumers who stay with the incumbent’s baseline tariff get milked when there are more searching consumers in a local market. In contrast, those who are willing to search either get a cheaper incumbent tariff, which includes a brand 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 by slightly increasing its baseline tariff while at the same time drastically reducing its cheaper online tariff. Entrants react to more search with somewhat lower prices, giving rise to increasing overall price dispersion and price discrimination, as well as increasing the alignment of incumbent and entrant

prices in the competitive online segment.

Robustness Checks

The above results are robust to a large variety of alternative specifications as shown below.

Level-Level Estimation

In the main specification, we focus on log-log specifications to allow interpretation of the results as elasticities. However, the results remain fully robust if we run level-level specifications of the models. They are reported in Tables A7 and A8 in the Appendix.

Non-Linear Effect of Search on Prices

We now allow for a non-linear relationship between search and prices, by adding a μ^2 in Equation 1. We instrument for μ^2 by using the square of the first-stage estimate of μ from Equation 2 as the instrument for μ^2 (see Wooldridge (2010, p. 262) on this approach). The results remain robust and are reported in the Appendix Tables A9 and A10. A graphical illustration is provided in Graph A7 of the Appendix.

Alternative Outcome Variables

We also estimate models with different outcome variables. In the first set of these models (Appendix Tables A11 and A12), we estimate the impact of search on retailers' markups for different tariffs and on differences in these markups. We compute markups as the differences between (net) prices and costs. In the second set (Tables A13 and A14) we use Lerner Indices as the dependent variables. They are computed as the ratio of markups to prices.⁴⁰ The results for markups and Lerner Indices are as one would expect from the results of the price estimations.

Alternative Instruments

Though we do not assume that our instruments *BBA* and *U40* are correlated with the error term in Equation 1, we still apply a robustness check in which we use the average values of the instruments in the 50 surrounding zip codes (while disregarding the actual zip code of observation i) as alternative instruments. This is in the spirit of Hausman and Taylor (1981), who use average values of the endogenous variable in surrounding areas as instruments. In contrast, our approach uses average values of the initial instrument as a new instrument. The results are fully robust to these alternative instruments and are shown in the Appendix Tables A15 and A16.

Additionally Instrumenting for the Number of Competitors

⁴⁰As there are also negative values for the markups and Lerner Indices, we do the log-transformation on the absolute values and put the sign back after that.

The prices and the number of competitors may affect each other. In order to test if potential endogeneity of the number of competitors affects our results, we also instrument for the number of firms in a zip code by using the average number of firms in the surrounding 50 zip codes in the spirit of Hausman and Taylor (1981). There is sufficient variation regarding the number of firms in the zip code itself and the surrounding 50 zip codes: we observe an absolute difference between the former and the latter of 7 and also a standard deviation of 7. The results only change marginally, as shown in Tables A17 and A18 in the Appendix, suggesting robustness of our results to instrumenting for the number of competitors.

Alternative Clustering of Standard Errors

Many incumbents operate only locally and 46% of the incumbents only have a single zip code in their incumbency area. These small incumbents are mostly municipal utilities. However, larger incumbents often have several zip codes in their incumbency area and charge locally differing baseline tariffs. The different *price zones* are not necessarily at the zip code level.⁴¹ Hence, as a robustness check, we cluster standard errors in two alternative ways. In the first version, we allow the residuals to correlate within incumbent price zones and cluster standard errors at that level, instead of the zip code level (see Appendix Tables A19 and A20). In the second version, the clustering is at the incumbency area level (see Appendix Tables A21 and A22). In both cases the results remain fully robust.

6 A simple search theoretic model

In this Section, we consider a simple model to (i) show that the empirical findings are perfectly compatible with the essential incentives of the electricity providers and consumers and (ii) to provide a counterfactual analysis of a scenario where electricity providers are not allowed to price discriminate between online platforms and traditional marketing channels. The model applies to any liberalized market where an incumbent firm competes with entrants for a homogeneous product and the incumbent is able to price discriminate between searching and loyal consumers.

The model closely follows the institutional details described above. All consumers observe the regular (baseline) price P_H^I of the incumbent at no additional cost.⁴² There is an online price comparison website consumers can consult at a

⁴¹See Section 2 for more details.

⁴²Note that our model is static and treats incumbents and entrants as asymmetric. In real markets, the following dynamic aspects may be important: once some consumers have switched to entrants, they later gain some incumbency effect as these consumers will have to search if they want to switch away from their provider. Thus, over time entrants and incumbents may become more symmetric to each other. In the theoretical model, we have abstracted from these considerations as individual entrants in the German electricity market typically have a very small market share. In

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 of the paper, z are the instruments that are exogenous to the prices across regions, but that do affect differences in search behavior across different zip codes). 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^E of the overall cheapest firm (usually an entrant) and the cheapest (online) price P_L^I of the incumbent. We assume that once on the website, a consumer compares prices without additional search cost.

Apart from their search cost, consumers also have some brand loyalty to the incumbent. Brand loyalty 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 loyalty), we assume that brand loyalty is proportional to search cost and say that the brand loyalty of a consumer with search cost s is denoted by θs , with $\theta < 1$. 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 and are more loyal to the incumbent.⁴³ Once a consumer with search cost s is online and observes both prices P^E and P_L^I then he will continue to buy from the incumbent if $P_L^I - \theta s < P^E$.

The sequence of actions is as follows. In the first-stage, the incumbent and entrant choose P_H^I, P_L^I and P^E simultaneously.⁴⁴ At the beginning of the second-stage, consumers only observe P_H^I and decide whether or not to search (given their expectations of the online prices). If they do not search they buy at P_H^I from the incumbent. If they do 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_H^I (different from the equilibrium level), then he will continue to believe that P_L^I and P^E are at their equilibrium levels.

This model abstracts from possible dynamic considerations and from the interaction between many entrants. In the different local electricity markets in Germany, nearly twenty years after the first market liberalization, entrants are likely to have built up some small groups of more or less loyal consumers, giving them some market power as well. On the other hand, entrants also compete with each other, which

Table 1, we report that on average there are more than 130 firms active in every zip code, though the incumbent provider continues to have around 76% market share. This implies that on average entrants have less than 0.2% market share.

⁴³Even 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.

⁴⁴In the Appendix, we consider an alternative “Stackelberg” version of the model where the incumbent first chooses its baseline price P_H^I , and P_L^I and P^E are chosen at the moment P_H^I is given and observed by the entrant. In this case, online prices (and consumer expectations of them) react to the baseline price and the incumbent takes this reaction into account when setting P_H^I . The qualitative results do not change, however.

clearly lowers this incumbency effect. In our model, we have only one entrant (as, in our data, we observe only the lowest entrants' price) without loyal consumers. If the entrants' incumbency and competition effect in the real market are of the same order of magnitude, our static model is a reasonable approximation of the market we intend to describe.

We will be looking for an 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 loyalty (and thus also low search cost) buy from the entrant, while other online consumers buy from the incumbent at its cheapest (online) price. The *ex ante* utility of a consumer with search cost s to buy at the three different prices is given by $v - P^E - s$, $v - P_L^I + (\theta - 1)s$ and $v - P_H^I + \theta s$ 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 search costs are $\hat{s}_1 = (P_L^I - P^E) / \theta$ and $\hat{s}_2 = (P_H^I - p_L^{I^e})$ such that all consumers with $s < \hat{s}_1$ buy from the entrant, all consumers with $\hat{s}_1 < s < \hat{s}_2$ buy at the online price of the incumbent and all consumers with $s > \hat{s}_2$ buy at the regular price of the incumbent. Note that in the definition of \hat{s}_2 we have $P_L^{I^e}$, the online price of the incumbent that consumers expect to find if they search, but before engaging in search, and not the actual 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.⁴⁵ Note also that we do not have to deal with expected prices in the definition of \hat{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, have incurred the search cost, and observe both prices.

In this theory section we assume, without loss of generality, that the firms have no supply cost. 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:

$$\pi_E = F(\hat{s}_1; z)P^E = F\left(\frac{P_L^I - P^E}{\theta}; z\right)P^E$$

and

⁴⁵To determine the equilibrium value of P_L^I and P_H^I it is important that a consumer will not observe a deviation from P_L^I before the decision whether or not to search is made.

$$\begin{aligned}\pi_I &= [F(\widehat{s}_2; z) - F(\widehat{s}_1; z)] P_L^I + (1 - F(\widehat{s}_2; z)) P_H^I \\ &= \left[F(P_H^I - p_L^{I^e}; z) - F\left(\frac{P_L^I - P^E}{\theta}; z\right) \right] P_L^I + (1 - F(P_H^I - p_L^{I^e}; z)) P_H^I.\end{aligned}$$

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

$$F\left(\frac{P_L^I - P^E}{\theta}; z\right) - f\left(\frac{P_L^I - P^E}{\theta}; z\right) \frac{P^E}{\theta} = 0, \quad (3)$$

$$F(P_H^I - P_L^I; z) - F\left(\frac{P_L^I - P^E}{\theta}; z\right) - f\left(\frac{P_L^I - P^E}{\theta}; z\right) \frac{P_L^I}{\theta} = 0, \quad (4)$$

and

$$-f(P_H^I - P_L^I; z)(P_H^I - P_L^I) + (1 - F(P_H^I - P_L^I; z)) = 0, \quad (5)$$

where $f(\cdot)$ is the density function that is associated with $F(\cdot)$. Note that the fraction of actively searching consumers is given by $F(P_H^I - P_L^I; z)$.

For a given z , these three F.O.C.s determine the equilibrium values of $P_H^{I^*}$, $P_L^{I^*}$ and P^{E^*} and the associated levels of price discrimination and price dispersion and the fraction of active searchers.⁴⁶ To explain differences in the observed level of price discrimination and price dispersion and the fraction of active searchers between regions, we have to see how the equilibrium levels of prices change with changes in z .

The next result provides a general statement on the conditions affecting price discrimination and online price dispersion. The effects of the loyalty parameter θ are clear-cut: all prices and online price dispersion are increasing in θ , whereas overall price dispersion is decreasing and price discrimination is unaffected. The effects of changes in the search cost distribution, reflected by changes in the parameter z , are more rich and different patterns are possible.

Proposition 1 *The effects of an increase in brand loyalty θ is that all prices and online price dispersion increase, while overall price dispersion decreases and price discrimination is unaffected. The effects of changes in the parameters underlying the search cost distribution z are as follows. Price discrimination increases if, and only if, the inverse hazard rate evaluated at the equilibrium values $\frac{1 - F(P_H^{I^*} - P_L^{I^*}; z)}{f(P_H^{I^*} - P_L^{I^*}; z)}$ is increasing in z . Moreover, P^E and online price dispersion are positively related to P_L^I if the density functions are non-increasing, i.e., $\partial f\left(\frac{P_L^I - P^E}{\theta}; z\right) / \partial (P_L^I - P^E) \leq 0$. Finally, online price dispersion and price discrimination are linked by $1 + f(P_H^I -$*

⁴⁶Asterisks indicate equilibrium values.

$$P_L^I; z)(P_H^I - P_L^I) = f\left(\frac{P_L^I - P^E}{\theta}; z\right) \left(\frac{P_L^I + P^E}{\theta}\right).$$

The economic intuition behind the result on price discrimination is clear: for a given value of P_L^I the incumbent faces a trade-off in its decision whether or not to increase P_H^I . Raising P_H^I increases the profits for all consumers $1 - F(P_H^{I*} - P_L^{I*}; z)$ who will stay on the baseline tariff; on the other hand, a fraction proportional to the density $f(P_H^{I*} - P_L^{I*}; z)$ will decide to search. At the margin, those that decide to search will eventually buy at the incumbent's online price P_L^{I*} as the marginal consumer has a higher search and switching cost. The incumbent will lose $P_H^{I*} - P_L^{I*}$ per (marginal) consumer who searches. If, *evaluated at the equilibrium values*,⁴⁷ the inverse hazard rate is increasing in z , relatively more consumers will stay on the baseline tariff if z increases, making price discrimination more profitable.

Also, to understand online price dispersion, if P_L^I increases, then there is a larger potential demand for the entrant and, under “normal” demand conditions, it should increase its price, but not to the full extent (thereby also increasing sales). Online price dispersion and price discrimination are linked by how the incumbent sets 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. To get reasonably simple expressions for the F.O.C.s, we have adopted a piece-wise linear search cost distribution:

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

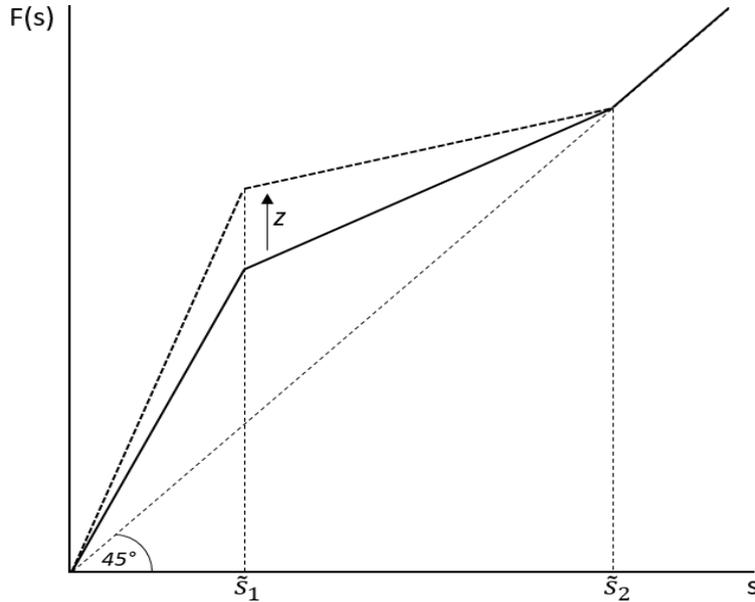
where to have a proper piece-wise linear distribution function, we should have $\alpha = \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1}$, $\beta = \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}$ and $\tilde{s}_2 > \tilde{s}_1$ and $z > 0$. If $z = 1$, we have the uniform distribution.⁴⁸ We focus on parameter values such that $\hat{s}_1 < \tilde{s}_1 < \hat{s}_2 < \tilde{s}_2$, 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.⁴⁹

⁴⁷Most distributions covered in standard statistics textbooks have an inverse hazard rate that is decreasing. Note, however, the difference: We ask the inverse hazard rate to be increasing in an exogenous parameter z and we do not require it to be increasing over the full domain of possible search cost values, whereas the common description is one where $(1 - F(x))/f(x)$ is decreasing in x .

⁴⁸An exponential search cost distribution according to which $F(s; z) = 1 - e^{-zs}$ on $s \in [0, \infty)$ would have been an alternative choice, but in that case $P_H^I - P_L^I = 1/z$ and $F(P_H^I - P_L^I) = 1 - e^{-1}$, which is independent of z . Thus, with an exponential distribution we cannot explain changes in the fraction of searchers between regions.

⁴⁹The qualitative results continue to hold if $P_H^I - P_L^I > \tilde{s}_2$ although the specific formulas will be different.

Figure 4: A piece-wise linear search cost distribution



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

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 \tilde{s}_1 and/or in \tilde{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 \tilde{s}_1 or in \tilde{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. As the instruments we used in the empirical part of the paper relate to households having very low search cost, 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 $F(P_H^{I*} - P_L^{Ie})$ is also endogenously 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 $P_H^{I*} - P_L^{Ie}$. Using (5) it is easy to see that for the case where the search cost distribution is piece-wise linear

$$P_H^I - P_L^I = \frac{\tilde{s}_2 - \tilde{s}_1 - (z-1)\tilde{s}_1\tilde{s}_2 - (\tilde{s}_2 - z\tilde{s}_1)(P_H^I - P_L^I)}{(\tilde{s}_2 - z\tilde{s}_1)},$$

so that the equilibrium level of price discrimination equals

$$P_H^{I*} - P_L^{I*} = \frac{\tilde{s}_2 - \tilde{s}_1 - (z-1)\tilde{s}_1\tilde{s}_2}{2(\tilde{s}_2 - z\tilde{s}_1)} \quad (6)$$

and thus that the equilibrium fraction of online searchers equals

$$F(P_H^{I*} - P_L^{I*}) = \frac{(z-1)\tilde{s}_1\tilde{s}_2 + \tilde{s}_2 - \tilde{s}_1}{2(\tilde{s}_2 - \tilde{s}_1)}.$$

Finally, applying the piece-wise linear search cost distribution to (3) and (4), it is easy to see that the relation between the equilibrium online prices is given by

$$P^{E*} = \frac{1}{2}P_L^{I*}$$

so that

$$P_L^{I*} = \frac{\tilde{s}_2 - \tilde{s}_1 + (z-1)\tilde{s}_1\tilde{s}_2\theta}{3z(\tilde{s}_2 - \tilde{s}_1)},$$

which implies that

$$P^{E*} = \frac{\tilde{s}_2 - \tilde{s}_1 + (z-1)\tilde{s}_1\tilde{s}_2\theta}{6z(\tilde{s}_2 - \tilde{s}_1)}$$

and

$$P_H^{I*} = \frac{\tilde{s}_2 - \tilde{s}_1 + (z-1)\tilde{s}_1\tilde{s}_2\theta}{3z(\tilde{s}_2 - \tilde{s}_1)} + \frac{\tilde{s}_2 - \tilde{s}_1 - (z-1)\tilde{s}_1\tilde{s}_2}{2(\tilde{s}_2 - z\tilde{s}_1)}.$$

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 $P_L^{I*} - P^{E*}$ positively correlates with both prices.

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

$$\frac{\tilde{s}_2 - \tilde{s}_1 - \tilde{s}_1\tilde{s}_2}{\tilde{s}_1(5\tilde{s}_2 - 6\tilde{s}_1)} < z < \frac{\tilde{s}_2^2 - (\tilde{s}_2 - \tilde{s}_1)(1 - \tilde{s}_2)}{\tilde{s}_1\tilde{s}_2}. \quad (7)$$

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).

⁵⁰Note that these conditions are independent of θ . As an example, if $z = 1$ these inequalities reduce to $\frac{1}{6} < \tilde{s}_1 < \frac{1}{2}$ and $\tilde{s}_2 > \frac{1}{2}$. Or, when $\tilde{s}_1 = \frac{1}{5}$ and $\tilde{s}_2 = \frac{3}{5}$, $\frac{7}{9} < z < \frac{5}{3}$ has to hold.

Proposition 2 (*price levels*) *If (7) holds, then an increase in the fraction of online searchers $F(P_H^{I*} - P_L^{I*})$, initiated by an increase in z , coincides with a decrease in online prices P^{E*} and P_L^{I*} if and only if $\tilde{s}_2 - \tilde{s}_1 > \tilde{s}_2\tilde{s}_1$, while it coincides with an increase in the baseline price P_H^{I*} if θ is small enough, z is large enough, and/or $\tilde{s}_2 - \tilde{s}_1$ is small enough.*

The conditions mentioned in the proposition should be interpreted as follows. First, if $\tilde{s}_2 - \tilde{s}_1 < \tilde{s}_2\tilde{s}_1$, then \tilde{s}_1 is relatively close to \tilde{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. The reverse is true if $\tilde{s}_2 - \tilde{s}_1 > \tilde{s}_2\tilde{s}_1$. Second, the result for the incumbent's baseline price can be understood 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 $\tilde{s}_2 - \tilde{s}_1$ 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.

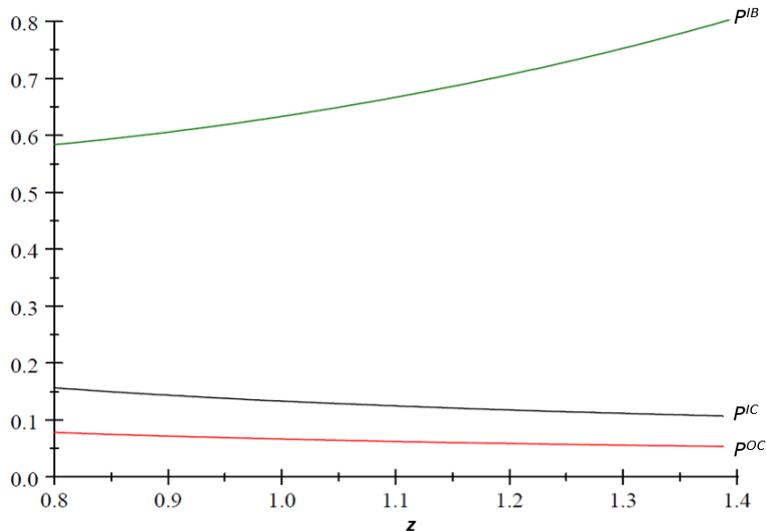
In line with our presentation of the empirical results, we present the results on price differences separately.

Proposition 3 (*Price discrimination and dispersion*) *If (7) holds, then an increase in the fraction of online searchers $F(P_H^{I*} - P_L^{I*})$, initiated by an increase in z , coincides with (i) an increase in price discrimination $P_H^{I*} - P_L^{I*}$ and (ii) a decrease in online price dispersion $P_L^{I*} - P^{E*}$, if and only if $\tilde{s}_2 - \tilde{s}_1 > \tilde{s}_1\tilde{s}_2$ and (iii) an increase in overall price dispersion $P_H^{I*} - P^{E*}$ if θ is small enough, z is large enough, or $\tilde{s}_2 - \tilde{s}_1$ is small enough.*

These results can be understood along similar lines as above. Proposition 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 5 depicts how the different prices change as a function of z when $\tilde{s}_2 = 3/5$ and $\tilde{s}_1 = 1/5$ and $\theta = 2/5$. As $\frac{\partial F(P_H^{I*} - P_L^{I*})}{\partial z}$ is a constant positive number, this figure can also be interpreted as how prices are linked to the fraction of searchers. One

Figure 5: Model prediction



Note: The figure predicts price changes as a function of z with $\tilde{s}_2 = 3/5$ and $\tilde{s}_1 = 1/5$ and $\theta = 2/5$. P_H^I , P_L^I , and P^E denote the incumbents' baseline tariffs, the incumbents' cheapest (online) tariffs, and the overall cheapest entrants' tariffs, respectively.

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 (overall) price dispersion, while online price dispersion is decreasing.

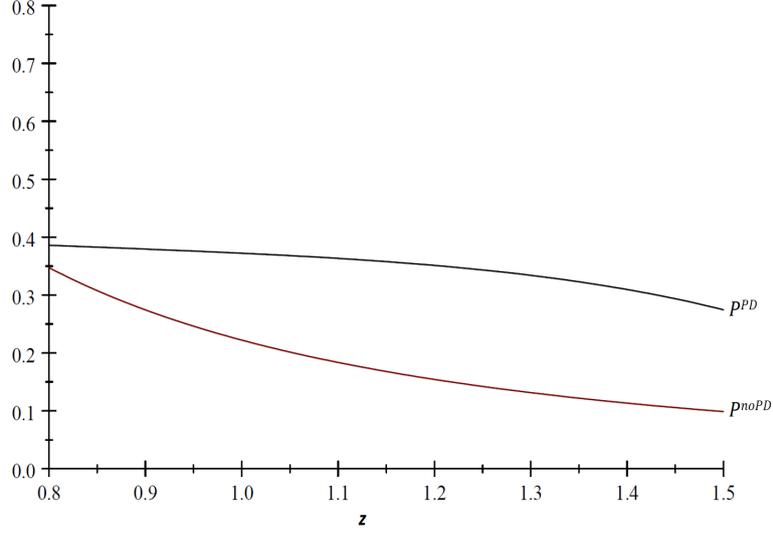
Thus, if markets (regions) mainly differ in the fraction of low search cost consumers (measured here by a difference in z), then an increase in online search is accompanied by more severe price discrimination by the incumbent, and if the search costs in the population are sufficiently spread ($\tilde{s}_2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

Welfare Analysis of Price Discrimination

We will now analyze whether consumers are better off without price discrimination. To this end, we simply force $P_H^I = P_L^I$ (and denote this value by P^I) and solve for the equilibrium values, denoting the price choice of the entrant under "no discrimination" by P_{ND}^E (to distinguish it from the price it chooses when the incumbent can price discriminate). As now we have that

$$\pi_E = F(\hat{s}_1; z)P^E = F\left(\frac{P^I - P_{ND}^E}{\theta}; z\right)P_{ND}^E$$

Figure 6: Average prices with and without price discrimination



Note: The figure predicts price changes as a function of z with $\tilde{s}_2 = 3/5$ and $\tilde{s}_1 = 1/5$ and $\theta = 2/5$. \bar{P} represents the average tariff with price discrimination; \bar{P}_{ND} represents the average tariff without price discrimination.

and

$$\pi_I = [1 - F(\hat{s}_1; z)] P^I = \left[1 - F\left(\frac{P^I - P_{ND}^E}{\theta}; z\right) \right] P^I,$$

it is easy to see that the two F.O.C.s are given by

$$F\left(\frac{P^I - P_{ND}^E}{\theta}; z\right) - f\left(\frac{P^I - P_{ND}^E}{\theta}; z\right) \frac{P_{ND}^E}{\theta} = 0,$$

and

$$1 - F\left(\frac{P^I - P_{ND}^E}{\theta}; z\right) - f\left(\frac{P^I - P_{ND}^E}{\theta}; z\right) \frac{P^I}{\theta} = 0.$$

Note that these conditions are very close to (3) and (4). In particular, it is clear that as $F(P_H^I - P_L^I; z) < 1$ in (4) in equilibrium $P_L^I < P^I$ and that because of the strategic complementarity of the price strategies, $P^E < P_{ND}^E$. Thus, searching consumers are better off with price discrimination. Intuitively, without price discrimination the incumbent has a larger share of “loyal” consumers he serves with the price P^I , compared to when it can price discriminate where P_L^I is meant to compete with the entrant’s price and the large share of loyal consumers is “addressed” by P_H^I . Thus, with price discrimination, there is simply more online competition to gain or attract searching consumers.

To compare P_H^I and P^I for the general case (and thus to make an overall comparison of the average price consumers pay⁵¹) is more difficult. Intuitively, though,

⁵¹One can also inquire into how the average price depends on the search intensity. The weighted

it would be natural to have that $P_H^I > P^I$ as under price discrimination the incumbent does not need to directly compete with the entrant's price when setting P_H^I . This is easily confirmed for the uniform distribution of search costs, where $z = 1$. In that case $P^{E*} = \frac{\theta}{6}$, $P_L^{I*} = \frac{\theta}{3}$, and $P_H^{I*} = \frac{1}{2} + \frac{\theta}{3}$, while $P_{ND}^{E*} = \frac{\theta}{3}$, $P^{I*} = \frac{2\theta}{3}$. As $\theta < 1$ it is easy to see that $P^{I*} < P_H^{I*}$.

For the case of the uniform distribution, it is also easy to calculate the average price consumers pay as $\frac{1}{2} \left(\frac{1}{2} + \frac{\theta}{3} \right) + \frac{1}{3} \frac{\theta}{3} + \frac{1}{6} \frac{\theta}{6} = \frac{1}{4} + \frac{11\theta}{36}$, for the case of price discrimination, while without price discrimination, the average price equals $\frac{2}{3} \frac{2\theta}{3} + \frac{1}{3} \frac{\theta}{3} = \frac{5\theta}{9}$. It follows that as $\theta < 1$, on average, the effect of the higher baseline price P_H^{I*} dominates and that consumers are worse off under price discrimination. Figure 6 confirms that, for the piece-wise linear distribution that we have considered above and for the same parameter values ($\tilde{s}_2 = 3/5$ and $\tilde{s}_1 = 1/5$ and $\theta = 2/5$), on average consumers are worse off under price discrimination. This average hides, however, that searching consumers are better off, and that loyal consumers are considerably worse off under price discrimination. In addition, with price discrimination, there will be fewer consumers switching to entrants than if price discrimination were banned. Note that the fraction of switchers is given by $F\left(\frac{P_L^I - P^E}{\theta}; z\right)$ for the case of price discrimination and by $F\left(\frac{P_L^I - P_{ND}^E}{\theta}; z\right)$ when price discrimination is banned. We have argued that $P_L^I < P^I$ and $P^E < P_{ND}^E$. As for the piece-wise linear distribution $P_{ND}^E = P^I/2$ and $P^E = P_L^I/2$ it follows that $\frac{P_L^I - P^E}{\theta} < \frac{P_L^I - P_{ND}^E}{\theta}$ so that fewer consumers switch under price discrimination.

7 Conclusion

After some 20 years of market liberalization and the entry of many companies, incumbents in Germany's local electricity retail markets still serve many consumers at higher prices than their competitors and prices remain highly dispersed. We provide an empirical analysis and a theoretical model detailing and explaining the pricing patterns. Using data on three tariffs (the incumbent's baseline price, the incumbent's cheapest price and the overall cheapest price) for each German zip code in the years 2011–2014, we empirically find that, on average, the incumbent's baseline tariff is increased 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, however the effect is much smaller. This pattern of price effects gives rise to increased overall price dispersion and price discrimination as well as reduced online price dispersion in local markets where consumers search more.

average price is given by $(1 - (F(\hat{s}_2))P_H^{I*} + (F(\hat{s}_2) - F(\hat{s}_1))P_L^{I*} + F(\hat{s}_1)P^{E*} = P_H^{I*} - F(\hat{s}_2)(P_H^{I*} - P_L^{I*}) - F(\hat{s}_1)(P_L^{I*} - P^{E*})$.

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 a strong 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 – presumably – attach a brand or loyalty premium to the incumbent. By engaging in price discrimination, the incumbent segments the markets into a segment of consumers with higher search costs and one with lower search costs.

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 appropriate surplus from consumers with high search and switching cost and help to prevent consumers who indicated they are willing to engage in search to switch to an entrant. In this way, few consumers may actually switch suppliers, allowing the incumbent to appropriate an important share of market revenue. This strategy may explain why even after some 20 years of market liberalization, incumbent's prices are still relatively high, and why a large fraction of consumers still buy from the incumbent.

In terms of welfare, we find that price discrimination is always beneficial for searching consumers, but that – for piece-wise linear search cost distributions – loyal consumers are worse off, and that on average the latter effect dominates, so that on average consumers are worse off. More generally, policies to foster search, while reducing prices in the competitive segments of the market, may also have a negative externality on non-searching consumers who must pay higher baseline prices if more consumers search.

We believe that this paper goes a long way towards describing and explaining the pricing behavior of incumbents and entrants after the liberalization of markets. However, once they have gained some market share, entrants may also try to increase their own prices, exploiting their own (small) incumbency effect. Moreover, the entrant may also engage in a price discrimination strategy where it sells at different prices to new customers online and to customers it attracted in previous years. These effects, although potentially small, are worthy of study. This requires, however, a dynamic model that is beyond the scope of the current paper.

References

- Bar-Isaac, H., Caruana, G., Cuat, V., 2012. Search, design, and market structure. *American Economic Review* 102 (2), 1140–1160.
- Baye, M. R., Morgan, J., Scholten, P., 2006. Information, search, and price dispersion. In: Hendershott, T. (Ed.), *Handbook of Economics and Information Systems*. Elsevier.
- 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.
- Byrne, D. P., De Roos, N., 2017. Consumer search in retail gasoline markets. *The Journal of Industrial Economics* 65 (1), 183–193.
- 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.
- 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, 2955–2980.
- 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 do incumbents respond to the threat of entry? evidence from the major airlines. *Quarterly Journal of Economics* 123 (4), 1611–1633.
- Hausman, J. A., Taylor, W. E., 1981. Panel data and unobservable individual effects. *Econometrica* 49 (6), 1377–1398.
- Hortacsu, A., Madanizadeh, S., Puller, S., 2017. Power to choose? An analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy* 9, 192226.
- 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.
- Lewis, M. S., Marvel, H. P., 2011. When do consumers search?. *The Journal of Industrial Economics* 59 (3), 457–483.
- 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., 2015. Information and price dispersion: Evidence from retail gasoline. CEPR Discussion Paper No. DP10771.
- 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.
- Shcherbakov, O., 2016. Measuring consumer switching costs in the television industry. *RAND Journal of Economics* 47 (2), 366–393.
- 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.
- Yang, H., Ye, L., 2008. Search with learning: understanding asymmetric price adjustments. *RAND Journal of Economics* 39 (2), 547–564.

Appendix

A Additional Figures and Tables

Figure A1: Screenshot of a typical online comparison platform

The screenshot shows the TOPTARIF website interface. At the top, there's a navigation bar with the TOPTARIF logo and a menu icon. Below the navigation bar, there's a main header with the text "BEQUEM UND RISIKOFREI ONLINE WECHSELN" and a circular graphic with an ice cream cone. To the right of the graphic is a form titled "DEINE VORTEILE:" and "DEINE DATEN:". The "DEINE VORTEILE:" section lists three benefits: "Dein persönliches Angebot aus über 17.000 Tarifen", "Unverbindlich und kostenfrei vergleichen", and "Ohne Risiko den Anbieter wechseln". The "DEINE DATEN:" section has input fields for "Postleitzahl" (68169), "Personen" (1, 2, 3, 4), and "Verbrauch" (3500 kWh/Jahr). A red button "JETZT VERGLEICHEN!" is at the bottom of the form. A circular badge on the right says "JETZT BIS ZU 84€ MEHR FÜR DICH!".

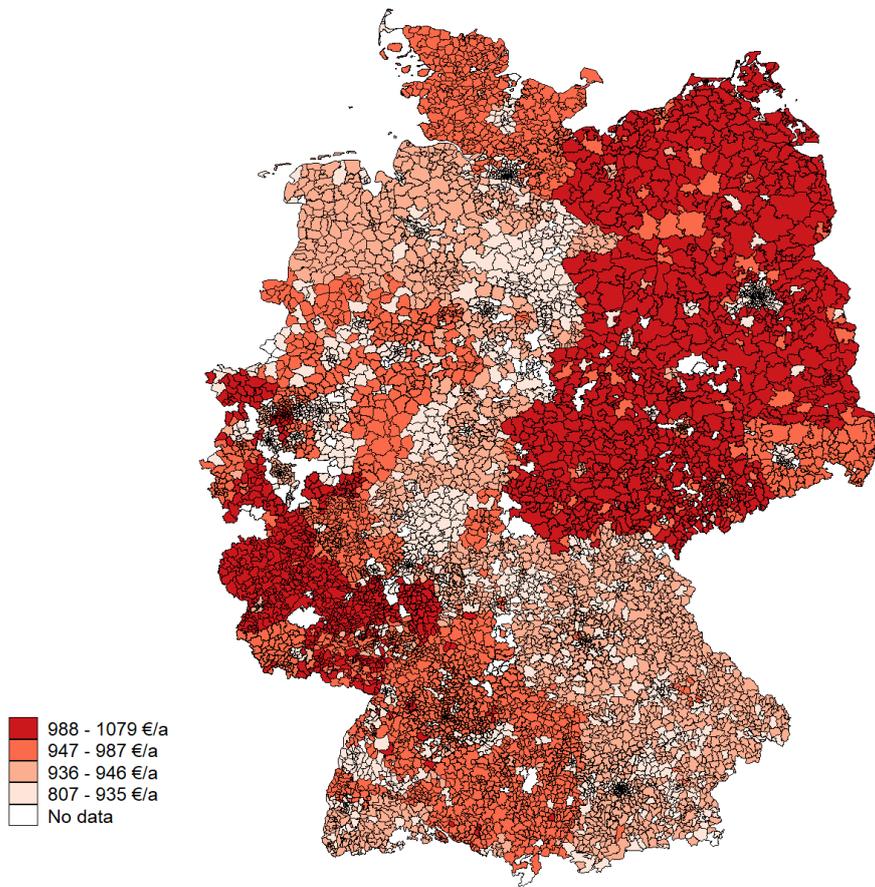
Below the form is a list of four energy tariffs. Each tariff card shows the provider logo, name, and details. The first tariff is from eprimo, with a price of 233,68 € and a cost of 842,56 € im 1. Jahr. The second is from BEV Energie, with a price of 233,60 € and a cost of 842,64 € im 1. Jahr. The third is from Sw Hamm, with a price of 222,05 € and a cost of 854,19 € im 1. Jahr. The fourth is from enQu, with a price of 221,28 € and a cost of 854,96 € im 1. Jahr. A red button "MEHR ZUM TARIF >" is at the bottom of each card.

Annotations on the right side of the screenshot point to specific elements:

- Zip code (Mannheim) - points to the "Postleitzahl" field.
- Household size - points to the "Personen" field.
- Suggested yearly consumption - points to the "Verbrauch" field.
- Price of cheapest tariff - points to the "ERSPARNIS" value of 233,68 €.
- Savings per year compared to incumbent baseline tariff - points to the "KOSTEN" value of 842,56 €.

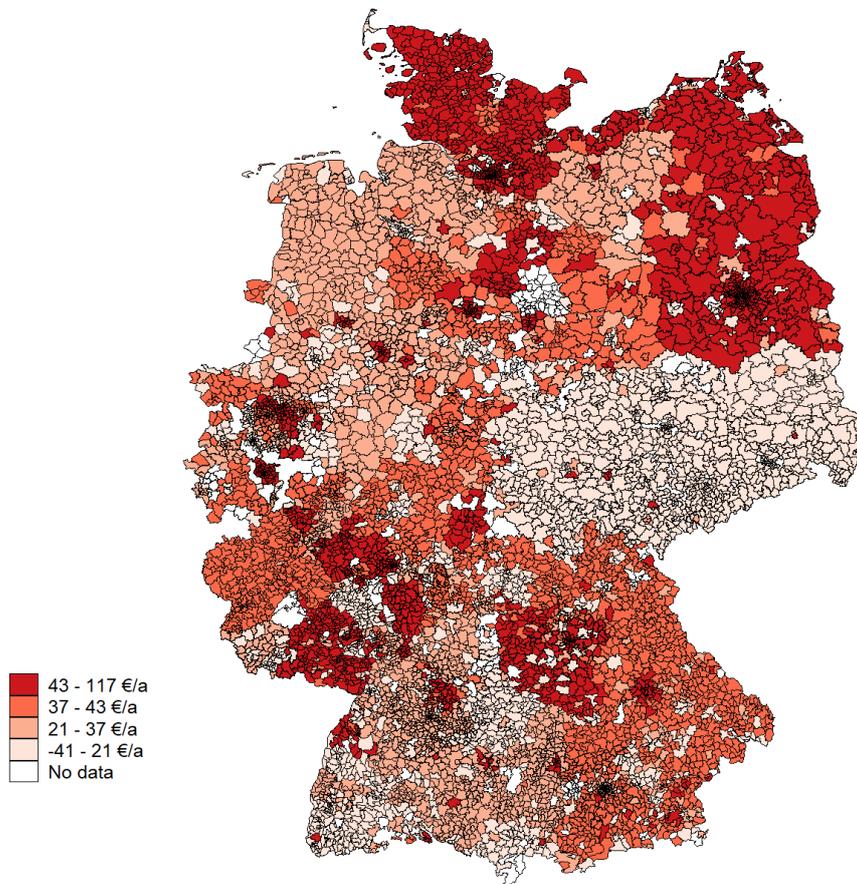
Note: Comparison platforms (here www.toptarif.de) list all available tariffs for a consumer given its expected annual consumption level for its local zip code, starting with the cheapest available tariff (including annual savings compared to the default incumbent baseline tariff). Site accessed on September 18, 2018.

Figure A2: Incumbents' baseline tariffs (2012)



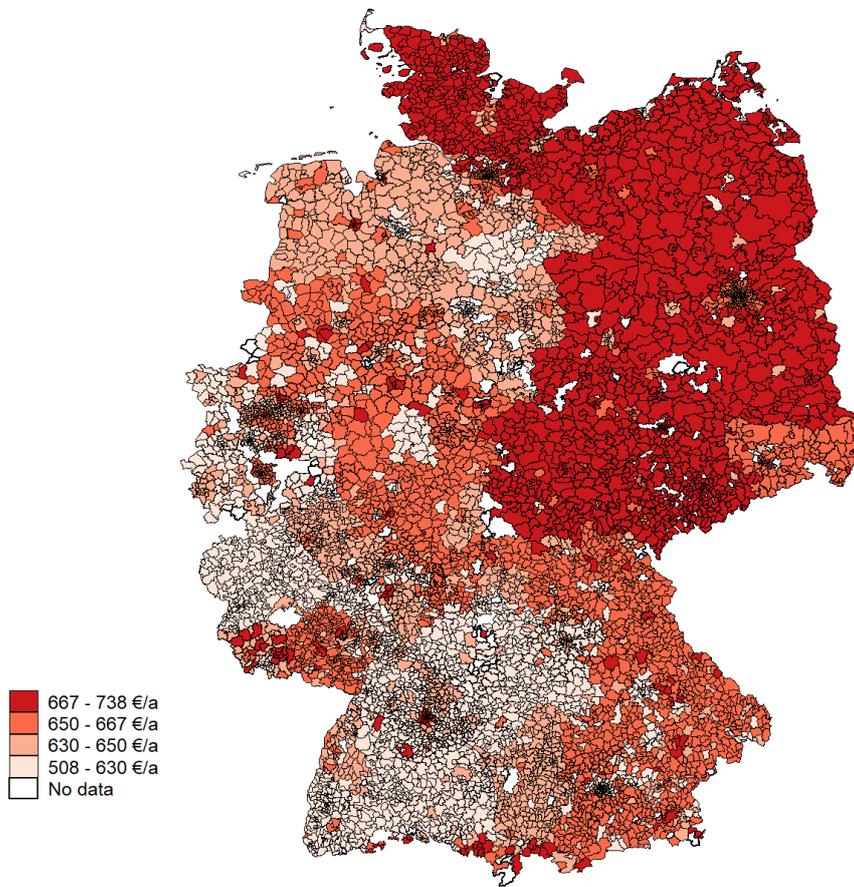
Note: The figure presents the spatial distribution of the incumbents' baseline tariffs for the year 2012.

Figure A3: Change of incumbents' baseline tariffs from 2011 to 2012



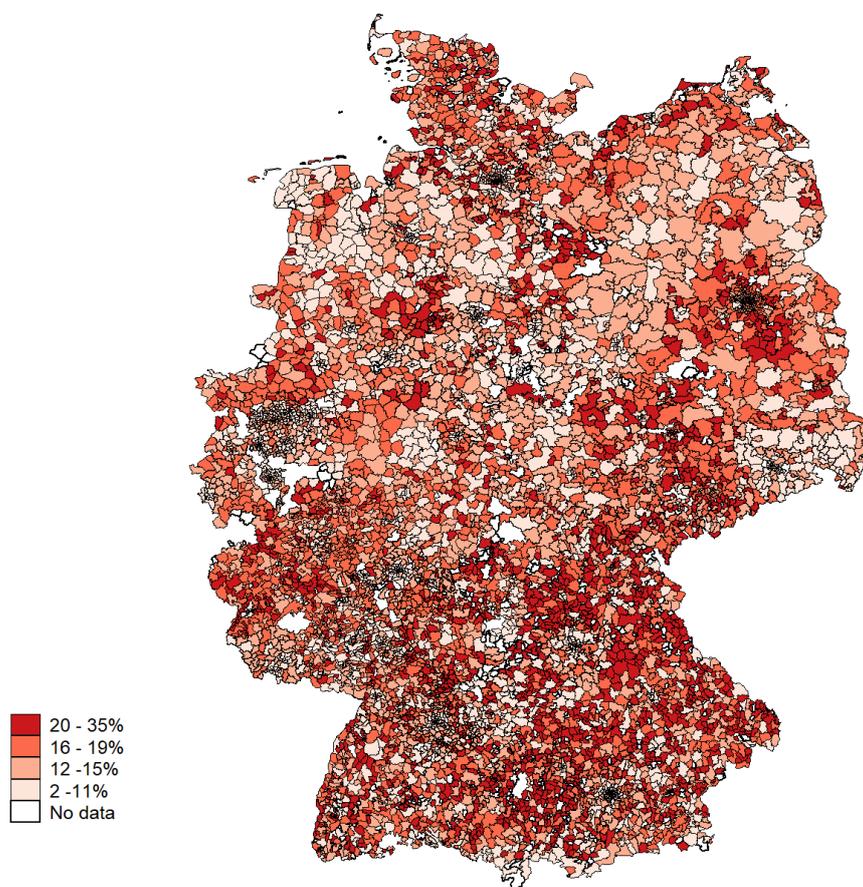
Note: The figure presents the spatial distribution of changes in the incumbents' baseline tariffs from 2011 to 2012.

Figure A4: (Net) costs (2012)



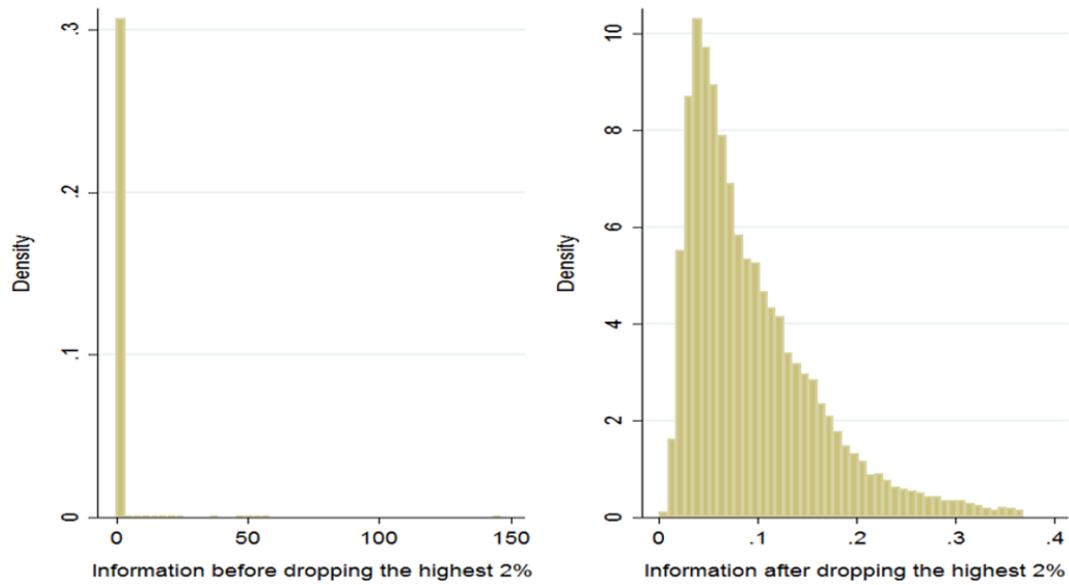
Note: The figure presents the spatial distribution of the (net) costs for electricity retailers in 2012.

Figure A5: Consumer search intensity (2012)



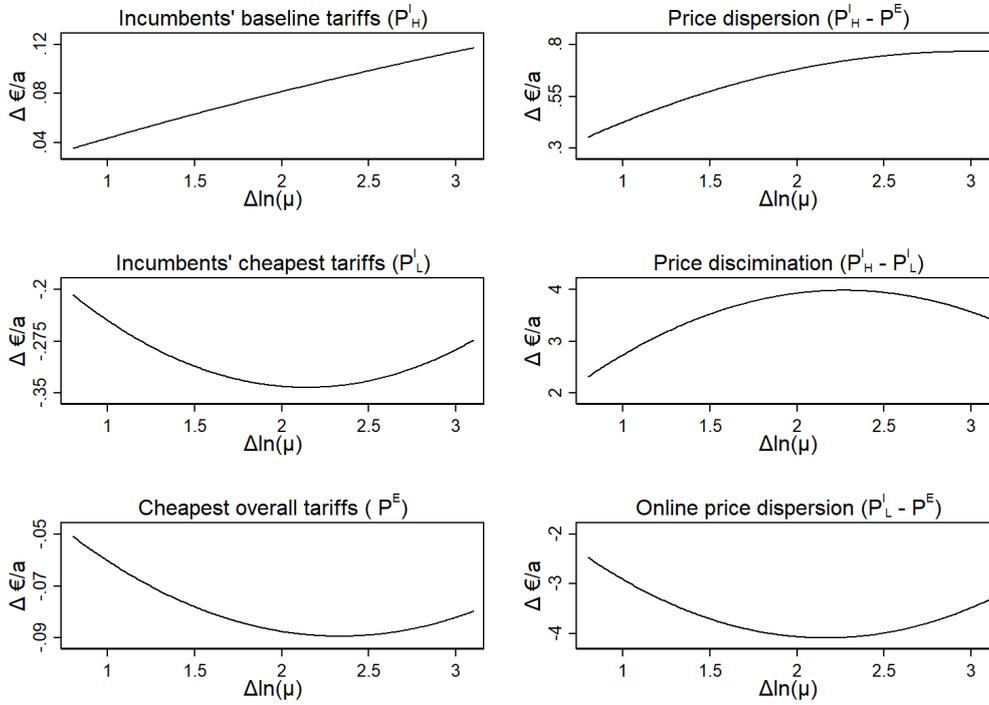
Note: The figure presents the spatial distribution of consumer search intensity in 2012.

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, not from actual consumers searching for electricity tariffs.

Figure A7: Quadratic impact of consumer search on prices and dispersion



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

Figure A8: Price zones of Envia Mitteldeutsche energie GmbH

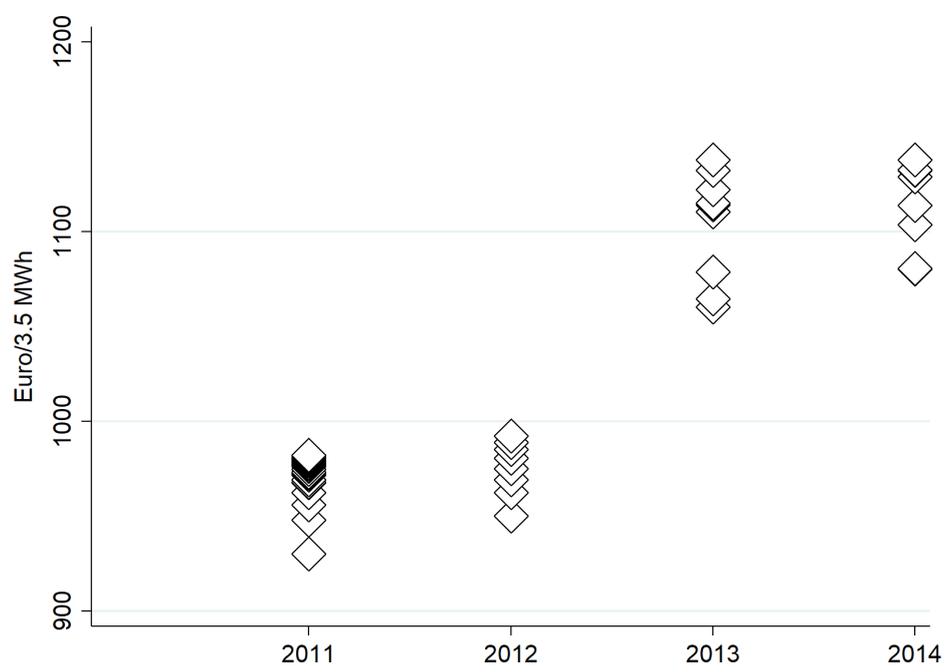


Table A1: Decomposition of standard deviations between and within zip codes

Variable	Mean	SD between	SD within
Incumbents' baseline tariff (P_H^I)	1,006	37.5	68.9
Incumbents' cheaper online tariffs (P_L^I)	929	37.4	78.7
Entrants' overall cheapest tariffs (P^E)	808	20.8	55.6
Competitors (#)	133	15.2	19.8
Net costs	683	29.0	31.2
Consumer search intensity (μ)	9.23	3.30	5.4
Head of HH below age of 40 (U40)	24.7	5.1	0.8
Broadband internet availability (BBA)	63.8	29.7	13.1

Table A2: First-stage regressions of consumer search (μ)

	(1) $\ln(\mu)$	(2) $\ln(\mu)$	(3) μ	(4) μ
U40	0.0164*** (0.0027)	0.0195*** (0.0028)	0.1016*** (0.0280)	0.1015*** (0.0295)
BBA	0.0005*** (0.0002)	0.0005** (0.0002)	0.0065*** (0.0018)	0.0048** (0.0017)
$\ln(\text{Costs})$		0.4702*** (0.1278)		
Costs				0.0296*** (0.0020)
#Competitors		0.0056*** (0.0006)		0.0843*** (0.0054)
Average HH size		0.7712*** (0.0999)		3.981*** (0.1.041)
Income <25k €/a		-0.0302 (0.0993)		3.911*** (1.0308)
Zip code fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	25,861	25,861	25,861	25,861

Notes: 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 prices (log-log)

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.003*** (0.000)	-0.004*** (0.001)	-0.001*** (0.000)
Costs	0.256*** (0.008)	0.422*** (0.017)	0.515*** (0.007)
#Competitors	0.001*** (0.000)	0.004*** (0.000)	0.000 (0.000)
Average HH size	0.046*** (0.005)	-0.015 (0.014)	-0.015*** (0.004)
Income <25k €/a	0.008 (0.007)	-0.067*** (0.014)	-0.006 (0.005)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	25,861	25,861	25,861

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

Table A4: OLS estimates of the impact of consumer search on dispersion (log-log)

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.013*** (0.002)	0.052*** (0.016)	-0.046** (0.020)
Costs	-0.842*** (0.047)	-1.304*** (0.218)	-0.732*** (0.251)
#Competitors	0.003*** (0.000)	-0.034*** (0.001)	0.058*** (0.001)
Average HH size	0.277*** (0.032)	0.891*** (0.189)	1.044*** (0.213)
Income <25k €/a	0.092* (0.054)	0.721*** (0.191)	-0.886*** (0.233)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	25,861	25,861	25,861

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

Table A5: IV estimates of the impact of consumer search on prices (log-log) – estimations without covariates

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.040*** (0.010)	-0.060** (0.024)	-0.017** (0.007)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	18.84	18.84	18.84
Durbin-Wu-Hausman test	0.00	0.01	0.02
Observations	25,861	25,861	25,861

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

Table A6: IV estimates of the impact of consumer search on dispersion (log-log) – Estimations without covariates

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.209*** (0.057)	0.706** (0.318)	-0.887** (0.379)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	18.84	18.84	18.84
Durbin-Wu-Hausman test	0.00	0.03	0.01
Observations	25861	25861	25861

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

Table A7: IV estimates of the impact of consumer search on prices (level-level)

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	4.427*** (1.530)	-13.022*** (3.930)	-2.851*** (0.897)
Costs	0.243*** (0.048)	0.958*** (0.121)	0.674*** (0.029)
#Competitors	0.237* (0.140)	4.617*** (0.349)	0.278*** (0.080)
Average HH size	19.159** (7.938)	15.217 (20.061)	-1.149 (4.922)
Income <25k €/a	22.991** (10.065)	-108.285*** (24.056)	-15.904** (6.200)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	7.91	7.91	7.91
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

Table A8: IV estimates of the impact of consumer search on dispersion (level-level)

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	7.387*** (2.086)	17.725*** (4.750)	-9.500*** (3.388)
Costs	-0.434*** (0.065)	-0.726*** (0.146)	0.267** (0.105)
#Competitors	-0.054 (0.188)	-4.413*** (0.419)	4.270*** (0.304)
Average HH size	20.812* (11.085)	3.390 (23.884)	15.755 (17.494)
Income <25k €/a	39.428*** (13.980)	132.376*** (29.352)	-89.455*** (20.785)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	7.91	7.91	7.91
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

Table A9: IV estimates of a non-linear impact of consumer search on prices (log-log)

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.046*** (0.010)	-0.320*** (0.031)	-0.077*** (0.008)
Search ²	-0.003*** (0.001)	0.075*** (0.003)	0.016*** (0.001)
Costs	0.251*** (0.010)	0.128*** (0.028)	0.454*** (0.009)
#Competitors	0.000*** (0.000)	0.003*** (0.000)	-0.000*** (0.000)
Average HH size	0.023*** (0.008)	0.058** (0.025)	0.005 (0.007)
Income <25k €/a	0.006 (0.008)	0.009 (0.022)	0.010* (0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Kleibergen-Paap F -stat.	17.11	17.11	17.11
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA and for μ^2 by the square of the first-stage prediction of μ . *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A10: IV estimates of a non-linear impact of consumer search on dispersion (log-log)

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.507*** (0.065)	3.524*** (0.407)	-3.794*** (0.417)
Search ² (μ^2)	-0.084*** (0.006)	-0.775*** (0.037)	0.876*** (0.037)
Costs	-0.573*** (0.066)	1.684*** (0.372)	-4.142*** (0.369)
#Competitors	0.004*** (0.000)	-0.025*** (0.002)	0.047*** (0.002)
Average HH size	0.101* (0.054)	0.018 (0.312)	1.930*** (0.340)
Income <25k €/a	0.003 (0.065)	-0.071 (0.277)	-0.018 (0.294)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Kleibergen-Paap F -stat.	17.11	17.11	17.11
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA and for μ^2 by the square of the first-stage prediction of μ . *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A11: IV estimates of the impact of consumer search on markups (log-log)

	(1) Markup Incumbent Base m_H^I	(2) Markup Incumbent Cheapest m_L^I	(3) Markup Overall Cheapest m^E
Search (μ)	0.373*** (0.071)	-2.905*** (0.766)	-2.397*** (0.591)
Costs	-3.300*** (0.082)	-8.468*** (0.763)	-30.573*** (0.704)
#Competitors	0.003*** (0.001)	0.171*** (0.006)	0.014*** (0.004)
Average HH size	0.079 (0.067)	3.746*** (0.682)	3.247*** (0.574)
Income <25k €/a	0.125 (0.080)	-3.315*** (0.612)	-0.557 (0.463)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	23.67	23.67	23.67
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

Table A12: IV estimates of the impact of consumer search on markup dispersion (log-log)

	(1) Markup Dispersion $m_H^I - m^E$	(2) Markup Discrimination $m_H^I - m_L^I$	(3) Markup Online Dispersion $m_L^I - m^E$
Search (μ)	0.274*** (0.057)	1.706*** (0.392)	-1.239*** (0.298)
Costs	-1.041*** (0.069)	-2.190*** (0.395)	-0.359 (0.333)
#Competitors	0.002*** (0.000)	-0.038*** (0.003)	0.056*** (0.002)
Average HH size	0.113** (0.055)	0.110 (0.338)	1.204*** (0.283)
Income <25k €/a	0.089 (0.065)	0.670** (0.279)	-0.825*** (0.230)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	23.67	23.67	23.67
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

Table A13: IV estimates of the impact of consumer search on Lerner Indices (log-log)

	(1) Lerner Index Incumbent Base L_H^I	(2) Lerner Index Incumbent Cheapest L_L^I	(3) Lerner Index Overall Cheapest L^E
Search (μ)	0.029*** (0.006)	-0.092*** (0.020)	-0.027*** (0.006)
Costs	-0.522*** (0.008)	-0.389*** (0.021)	-0.390*** (0.008)
#Competitors	0.000*** (0.000)	0.003*** (0.000)	0.000*** (0.000)
Average HH size	0.015*** (0.006)	0.043** (0.018)	0.001 (0.006)
Income <25k €/a	0.008 (0.006)	-0.052*** (0.014)	-0.005 (0.005)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	23.67	23.67	23.67
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

Table A14: IV estimates of the impact of consumer search on Lerner Index dispersion (log-log)

	(1) Lerner Index Dispersion $L_H^I - L^E$	(2) Lerner Index Discrimination $L_H^I - L_L^I$	(3) Lerner Index Online Dispersion $L_L^I - L^E$
Search (μ)	0.056*** (0.010)	0.121*** (0.023)	-0.064*** (0.017)
Costs	-0.131*** (0.012)	-0.133*** (0.023)	0.003 (0.019)
#Competitors	0.000* (0.000)	-0.003*** (0.000)	0.003*** (0.000)
Average HH size	0.014 (0.010)	-0.028 (0.020)	0.042*** (0.015)
Income <25k €/a	0.013 (0.010)	0.059*** (0.017)	-0.047*** (0.013)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	23.67	23.67	23.67
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,861	25,861	25,861

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

Table A15: IV estimates of the impact of consumer search on prices (log-log) – Hausman type instruments for search

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.045*** (0.009)	-0.157*** (0.026)	-0.030*** (0.006)
Costs	0.234*** (0.011)	0.491*** (0.031)	0.527*** (0.009)
#Competitors	0.000*** (0.000)	0.005*** (0.000)	0.000*** (0.000)
Average HH size	0.022** (0.009)	0.084*** (0.026)	0.006 (0.006)
Income <25k €/a	0.012 (0.009)	-0.061*** (0.022)	-0.005 (0.005)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	32.90	32.90	32.90
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,661	25,661	25,661

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA in the 50 surrounding zip codes. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A16: IV estimates of the impact of consumer search on dispersion (log-log) – Hausman type instruments for search

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.220*** (0.053)	1.810*** (0.352)	-2.277*** (0.385)
Costs	-0.935*** (0.062)	-2.126*** (0.392)	0.403 (0.431)
#Competitors	0.002*** (0.000)	-0.045*** (0.003)	0.072*** (0.003)
Average HH size	0.131*** (0.049)	-0.200 (0.340)	2.397*** (0.369)
Income <25k €/a	0.082 (0.060)	0.635** (0.291)	-0.859*** (0.313)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	32.90	32.90	32.90
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,661	25,661	25,661

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA in the 50 surrounding zip codes. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A17: IV estimates of the impact of consumer search on prices (log-log) – additionally instrumented for #Competitors

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.031*** (0.008)	-0.125*** (0.027)	-0.029*** (0.006)
#Competitors	0.001*** (0.000)	0.006*** (0.000)	0.000 (0.000)
Costs	0.254*** (0.011)	0.501*** (0.028)	0.524*** (0.009)
Average HH size	0.023*** (0.008)	0.051** (0.024)	0.006 (0.006)
Income <25k €/a	0.014* (0.008)	-0.061*** (0.019)	-0.007 (0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Kleibergen-Paap F -stat.	15.30	15.30	15.30
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,661	25,661	25,661

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA . Instrumented for #Competitors by #Competitors in the 50 surrounding zip codes. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A18: IV estimates of the impact of consumer search on dispersion (log-log) – additionally instrumented for #Competitors

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.220*** (0.053)	1.453*** (0.356)	-1.539*** (0.382)
#Competitors	0.005*** (0.001)	-0.052*** (0.003)	0.082*** (0.003)
Costs	-0.856*** (0.066)	-2.199*** (0.347)	0.391 (0.395)
Average HH size	0.100* (0.052)	0.137 (0.304)	1.764*** (0.345)
Income <25k €/a	0.122** (0.062)	0.698*** (0.253)	-0.824*** (0.293)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Kleibergen-Paap F -stat.	15.30	15.30	15.30
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,661	25,661	25,661

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA . Instrumented for #Competitors by #Competitors in the 50 surrounding zip codes. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A19: IV estimates of the impact of consumer search on prices (log-log) – alternative clustering of standard errors I

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.039* (0.022)	-0.115** (0.058)	-0.032** (0.016)
Costs: C	0.238*** (0.045)	0.476*** (0.107)	0.530*** (0.038)
#Competitors	0.000** (0.000)	0.005*** (0.001)	0.000 (0.000)
Income <25k €/a	0.009 (0.016)	-0.070*** (0.023)	-0.007 (0.010)
Average HH size	0.023 (0.019)	0.053 (0.055)	0.004 (0.015)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	10.19	10.19	10.19
Observations	25,861	25,861	25,861

Note: Standard errors clustered at the incumbents' price zone level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA . *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A20: IV estimates of the impact of consumer search on dispersion (log-log) – alternative clustering of standard errors I

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.292** (0.136)	1.398* (0.787)	-1.361* (0.754)
Costs: C	-0.979*** (0.287)	-1.963* (1.147)	-0.088 (1.351)
#Competitors	0.002 (0.001)	-0.042*** (0.010)	0.066*** (0.010)
Income <25k €/a	0.100 (0.131)	0.762** (0.308)	-0.926*** (0.310)
Average HH size	0.104 (0.131)	0.058 (0.749)	1.858** (0.924)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	10.19	10.19	10.19
Observations	25,861	25,861	25,861

Note: Standard errors clustered at the incumbents' price zone level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA . *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A21: IV estimates of the impact of consumer search on prices (log-log) – alternative clustering of standard errors II

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.039* (0.023)	-0.115* (0.063)	-0.032** (0.015)
Costs	0.238*** (0.063)	0.476*** (0.175)	0.530*** (0.041)
#Competitors	0.000** (0.000)	0.005*** (0.001)	0.000 (0.000)
Income <25k €/a	0.009 (0.016)	-0.070** (0.032)	-0.007 (0.010)
Average HH size	0.023 (0.017)	0.053 (0.068)	0.004 (0.016)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	9.95	9.95	9.95
Observations	25,861	25,861	25,861

Note: Standard errors clustered at the incumbency area level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA . *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A22: IV estimates of the impact of consumer search on dispersion (log-log) – alternative clustering of standard errors II

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.292** (0.131)	1.398* (0.829)	-1.361 (0.926)
Costs	-0.979*** (0.337)	-1.963 (1.593)	-0.088 (2.187)
#Competitors	0.002 (0.002)	-0.042*** (0.015)	0.066*** (0.017)
Income <25k €/a	0.100 (0.133)	0.762** (0.386)	-0.926** (0.387)
Average HH size	0.104 (0.108)	0.058 (0.870)	1.858 (1.529)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	9.95	9.95	9.95
Observations	25,861	25,861	25,861

Note: Standard errors clustered at the incumbency area level in parentheses. Estimation is by GMM. Instrumented for μ by $U40$ and BBA . *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A23: IV estimates of the impact of consumer search on prices (log-log) – alternative instrument for *BBA* also including mobile internet availability of at least 16 Mbit/s

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.040*** (0.009)	-0.078*** (0.022)	-0.024*** (0.006)
Costs	0.252*** (0.011)	0.496*** (0.022)	0.548*** (0.008)
#Competitors	0.000*** (0.000)	0.005*** (0.000)	0.000*** (0.000)
Average HH size	0.023*** (0.008)	0.031 (0.020)	-0.000 (0.006)
Income <25k €/a	0.009 (0.009)	-0.068*** (0.016)	-0.006 (0.005)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	22.45	22.45	22.45
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,578	25,578	25,578

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and mobile internet availability for smart phones in addition to fixed network internet of at least 16 Mbit/s. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A24: IV estimates of the impact of consumer search on dispersion (log-log) – alternative instrument for *BBA* also including mobile internet availability of at least 16 Mbit/s

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.232*** (0.055)	0.930*** (0.312)	-0.960*** (0.326)
Costs	-0.955*** (0.064)	-1.803*** (0.284)	0.287 (0.319)
#Competitors	0.002*** (0.000)	-0.040*** (0.002)	0.066*** (0.002)
Average HH size	0.129** (0.053)	0.284 (0.265)	1.603*** (0.303)
Income <25k €/a	0.088 (0.063)	0.746*** (0.212)	-0.985*** (0.249)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage <i>F</i> -stat.	22.45	22.45	22.45
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,578	25,578	25,578

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and mobile internet availability for smart phones in addition to fixed network internet of at least 16 Mbit/s. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A25: IV estimates of the impact of consumer search on prices (log-log)– instrument for *BBA* at 6 Mbit/s internet speed instead of 16 Mbit/s

	(1) Incumbent Base P_H^I	(2) Incumbent Cheapest P_L^I	(3) Overall Cheapest P^E
Search (μ)	0.031*** (0.009)	-0.139*** (0.029)	-0.042*** (0.008)
Costs	0.243*** (0.010)	0.486*** (0.029)	0.535*** (0.010)
#Competitors	0.000*** (0.000)	0.005*** (0.000)	0.000*** (0.000)
Average HH size	0.028*** (0.008)	0.067*** (0.026)	0.010 (0.008)
Income <25k €/a	0.008 (0.008)	-0.069*** (0.020)	-0.008 (0.006)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	21.41	21.41	21.41
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,578	25,578	25,578

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBA* but at 6 Mbit/s instead of 16 Mbit/s. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Table A26: IV estimates of the impact of consumer search on dispersion (log-log) – instrument for *BBA* with 6 Mbit/s internet speed instead of 16 Mbit/s

	(1) Price Dispersion $P_H^I - P^E$	(2) Price Discrimination $P_H^I - P_L^I$	(3) Online Price Dispersion $P_L^I - P^E$
Search (μ)	0.304*** (0.063)	1.682*** (0.379)	-1.480*** (0.371)
Costs	-0.980*** (0.071)	-2.077*** (0.352)	-0.080 (0.362)
#Competitors	0.002*** (0.000)	-0.044*** (0.003)	0.066*** (0.003)
Average HH size	0.089 (0.059)	-0.078 (0.318)	1.812*** (0.332)
Income <25k €/a	0.087 (0.067)	0.746*** (0.253)	-0.964*** (0.273)
Zip code fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
First-stage F -stat.	21.41	21.41	21.41
Durbin-Wu-Hausman test	0.00	0.00	0.00
Observations	25,578	25,578	25,578

Note: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBA* but at 6 Mbit/s instead of 16 Mbit/s. *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

B Proofs of Propositions

Proposition 1. The effect of an increase in brand loyalty θ is that all prices and online price dispersion increase, while overall price dispersion decreases and price discrimination is unaffected. Price discrimination increases if, and only if, the inverse hazard rate evaluated at the equilibrium values $\frac{1-F(P_H^I - P_L^I; z)}{f(P_H^I - P_L^I; z)}$ is increasing in z . Moreover, online price dispersion is positively related to P_L^I if the density functions are non-increasing, i.e., $\partial f\left(\frac{P_L^I - P^E}{\theta}; z\right) / \partial(P_L^I - P^E) \leq 0$. Finally, online price dispersion and price discrimination are linked by $1 + f(P_H^I - P_L^I; z)(P_H^I - P_L^I) = f\left(\frac{P_L^I - P^E}{\theta}; z\right)\left(\frac{P_L^I + P^E}{\theta}\right)$.

Proof. The effects of changes in the loyalty parameter θ can be understood as follows. First, it clearly follows from (5) that $P_H^I - P_L^I$, which is the measure of price discrimination, is independent of θ . Recognizing this, it follows from deducting (3) from (4) that $\frac{P_L^I - P^E}{\theta}$ is also independent of θ , or in other words, that $P_L^I - P^E$, which is the measure of online price dispersion, is proportional to θ . From each of the individual equations (3) and (4) it then follows that both P_L^I and P^E have to be proportional to θ . Thus, all prices are increasing in θ . Overall price dispersion is decreasing in θ as P_H^I increases less than proportionally.

To understand the effects of changes in z , we first consider the result on price discrimination. Taking the total differential of (5) with respect to $P_H^I - P_L^I$ and z yields

$$\begin{aligned} & \left[-2f(P_H^I - P_L^I; z) - \frac{\partial f(P_H^I - P_L^I; z)}{\partial(P_H^I - P_L^I)}(P_H^I - P_L^I) \right] d(P_H^I - P_L^I) \\ & = \left[\frac{\partial f(P_H^I - P_L^I; z)}{\partial z}(P_H^I - P_L^I) + \frac{F(P_H^I - P_L^I; z)}{\partial z} \right] dz. \end{aligned} \quad (8)$$

As profit maximization implies that the second-order condition of (5) with respect to $P_H^I - P_L^I$ is negative, it should be that in an equilibrium,

$$-2f(P_H^I - P_L^I; z) - \frac{\partial f(P_H^I - P_L^I; z)}{\partial(P_H^I - P_L^I)}(P_H^I - P_L^I) < 0.$$

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

$$-\frac{\partial f(P_H^I - P_L^I; z)}{\partial z}(1 - F(P_H^I - P_L^I; z)) - \frac{F(P_H^I - P_L^I; z)}{\partial z}f(P_H^I - P_L^I) > 0,$$

which using (5) can be rewritten as

$$-f(P_H^I - P_L^I; z) \left[\frac{\partial f(P_H^I - P_L^I; z)}{\partial z} (P_H^I - P_L^I) + \frac{F(P_H^I - P_L^I; z)}{\partial z} \right] > 0.$$

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

To investigate online price dispersion, we take the total differential of (3) with respect to P_L^I and P^E to obtain

$$0 = \frac{1}{\theta} \left[f \left(\frac{P_L^I - P^E}{\theta}; z \right) - f' \left(\frac{P_L^I - P^E}{\theta}; z \right) \frac{P^E}{\theta} \right] dP_L^I + \frac{1}{\theta} \left[-2f \left(\frac{P_L^I - P^E}{\theta}; z \right) + f' \left(\frac{P_L^I - P^E}{\theta}; z \right) \frac{P^E}{\theta} \right] dP^E,$$

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 the second term in square brackets must be negative. If $f' \left(\frac{P_L^I - P^E}{\theta}; z \right) \leq 0$, then the first term in square brackets is positive, and its absolute value is smaller than the first term in square brackets. Thus, $0 < dP^E/dP_L^I < 1$. Therefore, $0 < d(P_L^I - P^E)/dP_L^I < 1$.

Finally, to understand how price discrimination and online price dispersion are related, we substitute (3) and (5) into (4) to get the condition stated in the Proposition. *Q.E.D.*

Proposition 2 (price levels). If (7) holds, then an increase in the fraction of online searchers $F(P_H^{I*} - P_L^{I*})$, initiated by an increase in z , coincides with a decrease in online prices P^{E*} and P_L^{I*} if and only if $\tilde{s}_2 - \tilde{s}_1 > \tilde{s}_2 \tilde{s}_1$, while it coincides with an increase in the baseline price P_H^{I*} if θ is small enough, z is large enough, or $\tilde{s}_2 - \tilde{s}_1$ is small enough.

Proof. It is clear that

$$\frac{\partial(P_H^{I*} - P_L^{I*})}{\partial z} = \frac{(\tilde{s}_2 - \tilde{s}_1)\tilde{s}_1(1 - \tilde{s}_2)}{2(\tilde{s}_2 - z\tilde{s}_1)^2} > 0.$$

From the expressions determining equilibrium prices, it follows that

$$2 \frac{\partial P^{E*}}{\partial z} = \frac{\partial P_L^{I*}}{\partial z} = -\frac{\theta}{3z^2} \left(\frac{\tilde{s}_2 - \tilde{s}_1 - \tilde{s}_1 \tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \right),$$

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

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

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 ($\tilde{s}_2 - \tilde{s}_1$ is small enough). *Q.E.D.*

Proposition 3 (price discrimination and dispersion). If (7) holds, then an increase in the fraction of online searchers $F(P_H^{I*} - P_L^{I*})$, initiated by an increase in z , coincides with (i) an increase in price discrimination $P_H^{I*} - P_L^{I*}$ and (ii) a decrease in online price dispersion $P_L^{I*} - P^{E*}$, if and only if $\tilde{s}_2 - \tilde{s}_1 > \tilde{s}_1 \tilde{s}_2$ and (iii) an increase in overall price dispersion $P_H^{I*} - P^{E*}$ if θ is small enough, z is large enough, or $\tilde{s}_2 - \tilde{s}_1$ is small enough.

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

$$\frac{\partial(P_H^{I*} - P_L^{I*})}{\partial z} = \frac{(\tilde{s}_2 - \tilde{s}_1)\tilde{s}_1(1 - \tilde{s}_2)}{2(\tilde{s}_2 - z\tilde{s}_1)^2} > 0$$

and

$$\frac{\partial F(P_H^{I*} - P_L^{I*})}{\partial z} = \frac{\tilde{s}_1 \tilde{s}_2}{2(\tilde{s}_2 - \tilde{s}_1)} > 0,$$

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

$$\frac{\partial(P_L^{I*} - P^{E*})}{\partial z} = -\frac{\theta}{6z^2} \left(\frac{\tilde{s}_2 - \tilde{s}_1 - \tilde{s}_1 \tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \right)$$

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

$$\frac{\partial(P_H^{I*} - P^{E*})}{\partial z} = \frac{\theta}{6z^2} \left(-1 + \frac{\tilde{s}_1 \tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \right) + \frac{(\tilde{s}_2 - \tilde{s}_1)\tilde{s}_1(1 - \tilde{s}_2)}{2(\tilde{s}_2 - z\tilde{s}_1)^2},$$

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 ($\tilde{s}_2 - \tilde{s}_1$ is small enough). *Q.E.D.*

C 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 modeled. Interestingly, even if the search cost distribution shifts the 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 z 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 \tilde{s}_2 , coincides with less online search. The reason is that a different part of the search cost distribution is shifted. An increase in z represents more consumers with the lowest search cost levels, whereas an increase in \tilde{s}_2 (for $z > 1$)

– or a decrease in \tilde{s}_2 (for $z < 1$) – implies that the 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.

Claim 1 (a) An increase in \tilde{s}_2 leads to an increase in the fraction of online searchers $F(P_H^{I*} - P_L^I)$, an increase in price discrimination $P_H^{I*} - P_L^I$ and an increase in online prices, if and only if $z < 1$. (b) If $z < 1$ an increase in \tilde{s}_1 leads to a decrease in the fraction of online searchers $F(P_H^{I*} - P_L^I)$, an increase in price discrimination $P_H^{I*} - P_L^I$ and a decrease in online prices. If $z > 1$, the impact on the fraction of online searchers is ambiguous, whereas the other effects are reversed.

D Sequential price setting game

When the firms compete in a sequential price setting game, in which the incumbent sets its baseline rate first, the respective profits of the entrant and incumbent do not change and are as given in the main text:

$$\pi_E = F(\hat{s}_1; z)P^E = F\left(\frac{P_L^I - P^E}{\theta}; z\right)P^E$$

and

$$\begin{aligned}\pi_I &= [F(\hat{s}_2; z) - F(\hat{s}_1; z)]P_L^I + (1 - F(\hat{s}_2; z))P_H^I \\ &= \left[F(P_H^I - P_L^{I^e}; z) - F\left(\frac{P_H^I - P^E}{\theta}; z\right) \right] P_L^I + (1 - F(P_H^I - P_L^{I^e}); z)P_H^I.\end{aligned}$$

In taking the first-order conditions, one has to be careful in this “Stackelberg” environment where, in the second stage, the incumbent sets the online price P_L^I simultaneously with the entrant choosing P^E , and the incumbent chooses the baseline price P_H^I in the first stage. In this case, when setting online prices, both players have to take the number of consumers who search online, i.e., $F(P_H^I - P_L^{I^e})$, as given. Thus, if (as explained in the main text) both online prices react to the incumbent baseline price, the F.O.C.s (evaluated at the equilibrium where $P_L^{I^e} = P_L^{I*}$) for the online prices (for given P_H^I), do not change either, so that they are given by:

$$F\left(\frac{P_L^I - P^E}{\theta}; z\right) - f\left(\frac{P_L^I - P^E}{\theta}; z\right) \frac{P^E}{\theta} = 0$$

and

$$F(P_H^I - P_L^I; z) - F\left(\frac{P_L^I - P^E}{\theta}; z\right) - f\left(\frac{P_L^I - P^E}{\theta}; z\right) \frac{P_H^I}{\theta} = 0,$$

respectively. This determines the online prices for given P_H^I : $P^E(P_H^I)$ and $P_L^I(P_H^I)$.

However, in determining the baseline price under “Stackelberg”, the incumbent and the consumers take these reactions into account. Thus, when observing P_H^I , consumers realize that the second stage prices will be affected by a change in P_H^I . Thus, the incumbent sets P_H^I such that

$$0 = -f(P_H^I - P_L^I; z)(P_H^I - P_L^I)\left(1 - \frac{\partial P_L^I}{\partial P_H^I}\right) - \frac{P_H^I}{\theta} f\left(\frac{P_L^I - P^E}{\theta}; z\right) \frac{\partial(P_L^I - P^E)}{\partial P_H^I} \\ + \left[F(P_H^I - P_L^I; z) - F\left(\frac{P_L^I - P^E}{\theta}; z\right) \right] \frac{\partial P_L^I}{\partial P_H^I} + (1 - F(P_H^I - P_L^I; z))$$

This expression has several new terms compared to the F.O.C. for P_H^I in the simultaneous choice model analyzed in the main text as, when setting P_H^I the incumbent (and the consumers) now consider how both online prices and the market shares change in response to changes in P_H^I .

For general distribution functions, it is not possible to solve these three equations in a meaningful way. Thus, in the rest of this appendix we consider the piece-wise linear distribution, where (as in the main text) we consider $\frac{P_L^I - P^E}{\theta} < \tilde{s}_1 < P_H^I - P_L^I < \tilde{s}_2$. As in the main text, the solution to (3) yields $P^E = P_L^I/2$, while in combination with (4) we have

$$\left(\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}\right) P_L^I = \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} P_H^I + \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1}.$$

Note from this equation it is clear that online prices are increasing in P_H^I but not to the full extent. In particular, $0 < \frac{\partial(P_L^I - P^E)}{\partial P_H^I} < \frac{\partial P_L^I}{\partial P_H^I} < 1$. Thus, if $\frac{P_L^I - P^E}{\theta} < \tilde{s}_1 < P_H^I - P_L^I < \tilde{s}_2$ the incumbent base line price solves

$$0 = -\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} (P_H^I - P_L^I) \left(1 - \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}\right) - \frac{z}{2\theta} \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} P_H^I \\ + \left[\frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} - \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} (P_H^I - P_L^I) - z \frac{P_L^I - P^E}{\theta} \right] \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} \\ + 1 - \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} - \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} (P_H^I - P_L^I),$$

or as

$$\left(\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}\right) P_L^I = \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1} P_H^I + \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1}$$

we have that

$$0 = -2 \left(\frac{3z}{2\theta} P_L^I - \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \right) - \frac{z}{\theta} \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} P_L^I \\ + 1 - \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \left(1 - \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right),$$

which can be simplified to

$$\frac{z}{\theta} \left(3 + \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right) P_L^I = 1 + \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \left(1 + \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right).$$

Thus, we have that the different equilibrium prices for the incumbent are given by

$$P_L^I = \frac{1 + \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \left(1 + \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right)}{\frac{z}{\theta} \left(3 + \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right)},$$

so that

$$P_H^I = \frac{1 + \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - \tilde{s}_1} \left(1 + \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right)}{\frac{z}{\theta} \left(3 + \frac{\frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}}{\frac{3z}{2\theta} + \frac{\tilde{s}_2 - z\tilde{s}_1}{\tilde{s}_2 - \tilde{s}_1}} \right)} \left(1 + \frac{3z}{2\theta} \frac{\tilde{s}_2 - \tilde{s}_1}{\tilde{s}_2 - z\tilde{s}_1} \right) - \frac{(z-1)\tilde{s}_1\tilde{s}_2}{\tilde{s}_2 - z\tilde{s}_1}.$$

For the parameter values we considered before, where $\theta = 2/5$, $\tilde{s}_1 = 3/10$ and $\tilde{s}_2 = 3/5$, this results in

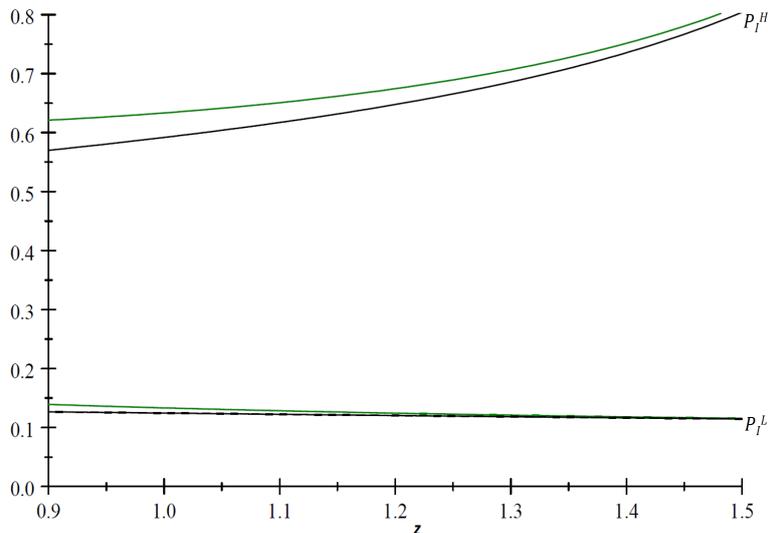
$$P_L^I = \frac{1 + \frac{.18(z-1)}{.3} \left(1 + \frac{6-3z}{\frac{9z}{0.8} + 6 - 3z} \right)}{\frac{5z}{2} \left(3 + \frac{6-3z}{\frac{9z}{0.8} + 6 - 3z} \right)}$$

and

$$P_H^I = \frac{1 + \frac{.18(z-1)}{.3} \left(1 + \frac{6-3z}{\frac{9z}{0.8} + 6 - 3z} \right)}{\frac{5z}{2} \left(3 + \frac{6-3z}{\frac{9z}{0.8} + 6 - 3z} \right)} \left(1 + \frac{.9z}{0.8(.6 - .3z)} \right) - \frac{.18(z-1)}{.6 - .3z}.$$

Figure D1 plots these prices under sequential price setting as a function of z together with the corresponding prices for the simultaneous move game analyzed in the main text. The figure shows that the two different analyses (simultaneous versus

Figure D1: Price patterns: simultaneous versus sequential game



Notes: The figure predicts price changes under sequential price setting (green) and simultaneous price setting (black) as a function of z with $\tilde{s}_2 = 3/5$ and $\tilde{s}_1 = 1/5$ and $\theta = 2/5$. Since the entrants' online tariffs (P^E) are half of the incumbents' cheaper online tariffs (P_L^I), P^E is not shown for better clarity.

sequential choice of offline and online prices) show that equilibrium outcomes are very close to each other. The reason is twofold. First, as indicated above, for given and identical P_H^I , the online market is governed by the same incentives and F.O.C.s. Second, if in the sequential setting the incumbent wants to increase its baseline tariff compared to the simultaneous choice setting, the incumbent not only gains because all prices will increase, but also loses as more consumers will switch to the entrant instead of buying from the online incumbent's price. These opposing forces are such that the net effect is that the baseline price is almost identical in the two cases.

The figure shows that if online prices react to the baseline price, the same pattern with respect to changes in z emerges, namely that if z increases (and therefore, more consumers search online), online prices decrease, while the incumbent's baseline price increases. Thus, price discrimination between loyal and searching consumers increases and online price dispersion decreases.