Discussion Paper No. 16-029

Non-Sequential Search, Competition and Price Dispersion in Retail Electricity

Klaus Gugler, Sven Heim, and Mario Liebensteiner

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 16-029

Non-Sequential Search, Competition and Price Dispersion in Retail Electricity

Klaus Gugler, Sven Heim, and Mario Liebensteiner

Non-Sequential Search, Competition and Price Dispersion in Retail Electricity*

Klaus Gugler†, Sven Heim‡, and Mario Liebensteiner§

April 1, 2016

Abstract

We investigate the impact of consumer search and competition on pricing strategies in Germany's electricity retail. We utilize a unique panel dataset on spatially varying search requests at major online price comparison websites to construct a *direct measure of search intensity* and combine this information with zip code level data on electricity tariffs between 2011 and 2014. The paper stands out by explaining price dispersion by differing pricing strategies of *former incumbents* and *entrant firms*, which are distinct in their attributable shares in informed versus uninformed consumers. Our empirical results suggest causal evidence for an inverted U-shape effect of consumer search intensity on price dispersion in a clearinghouse environment as in Stahl (1989). The dispersion is caused by opposite pricing strategies of incumbents and entrants, with incumbents initially *increasing* and entrants initially *decreasing* tariffs as a reaction to more consumer search. We also find an inverted U-shape effect of competition on price dispersion, consistent with theoretical findings by Janssen and Moraga-González (2004). Again, the effect can be explained by opposing pricing strategies of incumbents and entrants.

Keywords: Search, Information, Competition, Price Dispersion, Electricity Retail

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

*We gratefully acknowledge *ene't* for providing us with extensive data on search requests and prices as well as cost information. We thank *axciom* for providing detailed data on structural characteristics. We thank Ulrich Laitenberger, Maarten Jansen, Dieter Pennersdorfer, Philipp Schmidt-Dengler and Biliana Yontcheva for helpful comments. 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).

†Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria. E-mail: klaus.gugler@wu.ac.at

‡ ZEW Centre for European Economic Research and MaCCI Mannheim Centre for Competition and Innovation, E-mail: heim@zew.de

§ Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria. E-mail: mario.liebensteiner@wu.ac.at

1. Introduction

For the *law of one price* – as known from the standard homogeneous goods Bertrand model – to be valid, consumers must be aware of all offered prices for the product (Stigler, 1961). Electricity may be considered as such a kind of product, since it does not matter for the final consumer where and how it was produced. Therefore, one may assume that the transition from local monopoly supply to a more competitive environment after market liberalization should lead to price convergence with additional firm entry, until the law of one price holds. We focus on the German retail market, which was liberalized in 1999 (when the EU Directive 96/92/EC came into force). Since then, customers have had the freedom to choose their electricity supplier. For example, in 2014 a household customer had the choice between 73 and 198 providers, depending on its location. Nevertheless, even though 17 years have passed since, our data reveal that price dispersion has been pronounced and persisting despite considerable firm entry.

One possible explanation would be that electricity only appears to be a homogenous product, yet there might be some form of product differentiation, such as differences in service quality, certification with a "green" label[1], or marketing. Giulietti et al. (2014, p.559), however, argue that "product differentiation (…) is unlikely to be important in explaining differences between large established firms selling such a fundamentally homogenous product." Moreover, Stigler (1961) argues that not all of the observed price dispersion can be attributed to (marginal) product differentiation. In his view, the existence of significant *search frictions* directly induces price dispersion even in homogeneous goods markets. Thus, "price dispersion is a manifestation – and, indeed is a measure of ignorance in the market." (Stigler 1961, p.214). While theory proofs that there will be no price dispersion when either all or no consumers are informed, there will be price dispersion in between (Stahl, 1989).

The nexus between search frictions, competition, and price dispersion for homogenous goods has received much attention in the literature. Theoretical studies (e.g. Stahl, 1989; Baye and Morgan, 2001; Janssen and Moraga-González, 2004) show that for homogenous goods, such as electricity, price dispersion may represent an equilibrium outcome when customers differ in their search intensities. Thus, it may be

[1] In our sample, less than 4% of all search requests relate to tariffs with eco-labels. Therefore, in the present application we exclude search requests that only consider eco label tariffs in order to rule out pricing effects from such a form of product differentiation. However, the results remain fully robust when these searches are also included.

beneficial for firms to either charge a high price for customers with low (or no) search intensity, or to charge a low price for customers who do engage in costly search for price quotations.[2] Moreover, the number of firms may affect price dispersion and there may be reverse causality running from price dispersion to search intensity.

We add on this topic by analyzing how both *consumer information* and *competition* impact on price dispersion and prices in the German electricity retail market for the period 2011–2014. This paper is the first empirical attempt to analyze these relations altogether. Importantly, we are the first to employ a unique dataset on actual consumer search requests as a *direct measure of search intensity* in a panel data context. Furthermore, we are able to consider the endogenous nature of search with respect to price and price dispersion. The most important decisive feature is our ability to distinguish the price setting behavior of the local *incumbent* suppliers – whose consumers are uninformed –from the cheapest *entrant firms* – whose consumers are informed. Hence, we are not only able to investigate the effects of both consumer search and competition on price dispersion, but also to deduct how the incumbents and entrants set their prices. It is this combination of (1) applying a direct measure of search intensity, (2) recognizing the endogeneity of search intensity and (3) taking advantage of heterogeneous shares of informed versus uninformed consumers between incumbents and entrants in our empirical analysis that makes our study painting a fuller picture of the intricate relations.

In our sample, each of the 8,824 German zip codes is served by one of the 777 local incumbent electricity retailer, which represents the former local monopoly supplier. Prior to market liberalization, the local incumbent would have served all local customers in its distribution grid area. After the liberalization, those customers with low search costs were more likely to change to a cheaper entrant supplier, while customers with high search costs were more likely to stay with the incumbent. From this, we claim that in particular the local *incumbent retailer is associated with customers with high search costs and thus low search intensity*, while the *entrant retailers attract customers with no or low search costs and thus high search intensity.*[3]

[2] In this sense, Giulietti et al. (2014) highlight that *search costs* represent the main reason for low switching behavior and for observing large price differences between incumbents and entrants in electricity retail in the UK.

[3] In this regard, there is evidence of significant search costs. Giulietti et al. (2014) point to high search costs in electricity retail in the UK, and of Hortacsu et al. (2015) find that customers of the incumbent only search in about 2% of months or approximately once every 4–5 years.

Thus, we may draw inference about price setting of incumbents and entrants from heterogeneity in consumer information, which has not been examined thus far.

Our data provide suitable and detailed spatially varying measures for the dependent variables (price dispersion, price of cheapest entrant, and price of incumbent) and for our main variables of interest, namely the number of firms competing for customers and consumer search intensity. The latter variable is hardly observable and, hence, generally unknown to the researcher (Baye et al., 2006).[4] We utilize information on all search requests (i.e. clicks) from the main German online price comparison platforms for electricity retail tariffs for the period 2011–2014 to develop a *direct measure of consumer information*. Price comparison websites display *all* retail tariffs available at the address of the respective customer. Hence, we interpret consumer search in the sense of a *non-sequential clearinghouse model* as in Stahl (1989).[5]

We stress that sequential search models may have been more accurate in picturing consumers' price quotations in electricity retail in the past when online comparison platforms were less prevalent and, thus, consumers had to search for information about individual firms' prices. For example, Giulietti et al. (2014) investigate the period 2002–2005 for the U.K. electricity retail market, where consumer search was mainly driven by door-to-door selling of sales representatives of the respective electricity suppliers, and thus sequential search may be an adequate model.[6] For the recent period, simultaneous search models (and in particular clearinghouse models; see Baye et al., 2006) may have become more precise due to the emergence of comprehensive online search platforms as the standard searching devices (Montgomery et al. 2004).[7] Hence, at present, most of the costly search refers to finding out about the yearly electricity consumption level of the household and finding and opening the comparison website. Once this step is overcome, all available local prices in the zip code are shown to the consumer simultaneously. Giulietti et al. (2014, p. 576) corroborate that by stating that "(…) non-sequential search might be a good approximation of search behavior (…) if consumers use price comparison sites." It is therefore reasonable that

[4] We come back to this issue in section 2.1 where we discuss the relevant empirical literature.

[5] See, for example, Brown and Goolsbee (2002), Pennersdorfer et al. (2015), and Tang et al. (2010) for similar approaches.

[6] Another example of sequential search would be consumers who call their local suppliers or visit their websites in order to make price quotations. After each quotation, the consumer may decide whether he/she continues to make another price quotation.

[7] Similarly, De los Santos et al. (2012) reach the conclusion that fixed sample size search models have become more accurate than sequential search models.

the consumer who engages in non-sequential search at an online comparison platform choses the cheapest price among all available local prices.[8]

In our econometric approach we face one caveat that we cannot rule out potential simultaneity of both variables of interest, consumer search and competition, with the dependent variables, price dispersion and prices of incumbents and entrants. For this reason, we make use of an instrumental variables approach in order to circumvent the endogeneity problem and to determine causal effects of both variables. Our empirical results correspond well with findings from the respective theoretical literature. In line with Stahl (1989), we find that price dispersion follows an inverse U-shape relation with consumer information. That is, while more information leads to an increase in price dispersion to begin with, eventually – after a large part of consumers have informed themselves – price dispersion declines again. Somewhat in contrast to Stahl (1989), we find that the incumbent's price increases albeit at a decreasing rate (inverse U-shape), and in line with Stahl (1989) the entrant's price decreases at a decreasing rate (U-shape) with the share of informed consumers. With respect to competition, price dispersion follows an inverse U-shape as indicated by Janssen and Moraga-González (2004). Additional firms in the market cause the incumbent's price to rise at a declining rate (inverse U-shape), whereas the entrant's price falls at a declining rate (U-shape).

In addition, we extend our findings on the price setting behavior of incumbent and entrant firms by alternative regressions on markups and Lerner indices. Besides, the robustness of our empirical findings is confirmed by running all regressions at different household sizes (at different yearly consumption levels). What is more, we show that semi-parametric estimation allowing for flexible functional relationships particularly support the parametric estimates.

Importantly, our results show that the assumption of *equal shares* of informed versus uninformed consumers across firms – as generally assumed in the theoretical (e.g. Baye and Morgan, 2001; Stahl, 1989) and empirical (e.g. Chandra and Tappata, 2011; Pennersdorfer et al., 2014; Tang et al., 2010) literature – is not necessarily tenable. Price setting incentives differ across firms, e.g. incumbents versus entrants, according to their "endowment" of informed and uninformed consumers. We contend

[8] This stands in marked contrast to a sequential search model, a consumer obtains one price quotation after another until he/she finds a price that lies below his/her reservation price. See Baye et al. (2005) for an overview and discussion of theoretical search models on both sequential and non-sequential consumer search.

that incumbent firms, which are associated with uninformed consumers, follow opposing and mutually exclusive pricing strategies with respect to both consumer search and competition, compared to entrant firms, which face substantially larger shares of informed consumers. This determines whether firms follow a pricing strategy to either attract market shares from new customers who are willing to engage in search ("business-stealing effect") or to extract rents from existing customers who do not engage in search ("surplus-appropriation effect").[9]

The paper is structured as follows. The next section reviews the relevant theoretical and empirical literature. Section 3 provides a background on electricity retail in Germany. Section 4 describes the data and section 5 the econometric model. Section 6 discusses the results, and section 7 concludes.

2. Relevant literature on price dispersion

2.1. Relevant empirical literature on price dispersion

The empirical literature on search frictions can be broadly categorized into studies (1) that look at the determinants and effects of price dispersion and (2) studies that structurally estimate search (and switching) costs or directly analyze the search behavior of consumers. Studies to (1) may be differentiated according to whether they incorporate direct information on search intensity and studies that do not. Moreover, some studies recognize the endogeneity of the intensity of search with price dispersion. Baye et al. (2006) provide an excellent survey.

Studies that do not employ direct measures of search intensity include e.g. Sorensen (2000), Barron et al. (2004), Baye et al. (2004), Lewis (2008), and Chandra and Tappata (2011). Sorensen (2000), a seminal study in the field, finds that price-cost margins and price dispersion are negatively correlated with the purchase frequency of drugs, which can be interpreted as indirect evidence that search costs are important determinants of pricing strategies. Most studies find a negative effect of the number of sellers on price dispersion, e.g. for retail gasoline Barron et al. (2004) and Lewis (2008), and for electronics products Chandra and Tappata (2011) and Baye et al. (2004). Two results are worth mentioning in our context.

---

[9] In a similar manner, Ericson (2014, p. 44) refers to these effects as "investment motive" – firms acquire market shares in order to extract consumers' rents in future periods – and "harvesting motive" – firms maximize profits from their existing customers.

Chandra and Tappata (2011) using detailed information on competitors' locations in the U.S. retail gasoline market infer from their results that there is a negative relation between price dispersion and search intensity. However, they do not employ a direct measure of search intensity stating that it would be "ideal" to have a control group of no search. We take incumbents in electricity retail as such a "control" group, and show below that the relations between price dispersion/average prices and search intensity are intricate, namely non-linear and moreover dependent on the type of firms (low/high search intensity of customer base, i.e. incumbent or entrant).

The novel feature of Lewis (2008) is that the study looks at differentiated products (high-brand versus other gasoline stations) and, somewhat similarly to us, assumes that customers of high-brand stations are less likely to search (we assume that incumbent customers are less likely to search). Accordingly, Lewis (2008) finds that while price dispersion declines with station density for low-brand stations, price dispersion does not decline with station density if the station is high-brand. This implies that consumer heterogeneity interacts with seller heterogeneity. While Lewis (2008)'s results are novel in that respect, the study does not employ a direct measure of search intensity, and thus cannot analyze its endogenous nature. In addition to analyzing different types of firms (incumbent versus entrants), we employ such a measure for search intensity (number of clicks per zip code looking for electricity retail tariffs) and also instrument for search intensity and competition by truly exogenous shifters (i.e. new households, moved households, age of household heads for search intensity; number of households for competition).

The main drawback of the above studies is that they do not employ direct measures of search intensity and therefore they can only indirectly draw inference. Brown and Golsbee (2002), in contrast, use the variation in the share of consumers searching on the internet as their measure of consumer information.[10] They find that increased internet usage has resulted in an increase of price dispersion (in the market for term-life insurance) at low levels and a decrease of price dispersion at high levels of internet usage. While we get similar results for electricity retail, there are two shortcomings in the Brown and Golsbee (2002) study we seek to avoid. First, it does not incorporate the competitive environment in life insurance, let alone type differences among firms, and second it does not tackle the potential endogeneity of internet usage, i.e. internet usage could be higher if there is more price dispersion, since expected gains

[10] Ellison and Ellison (2005) and Ellison and Ellison (2009) question the extent to which the internet has actually reduced consumer search costs due to "bait and wait" and "obfuscation" strategies used by firms.

from obtaining information from shopbots (i.e. online comparison platforms) will increase with the dispersion of prices (see also Baye and Morgan, 2001).

More recently, Tang et al. (2010) examined the impact of changes in shopbot use on prices and price dispersion in online book retailing. An increase in shopbot use leads to decreases in average prices and price dispersion. Against the theory of an inverted U-shaped relation between price dispersion and search intensity, they find a U-shaped relation. The authors explain this conundrum by the fact that they only observe parts of the theoretical distribution of search but not the theoretical endpoints where no one searches and where everyone searches. While Pennersdorfer et al. (2014) do find an inverted U-shaped relation between price dispersion in the Austrian gasoline retail market and the share of informed consumers (as proxied by the share of commuters), Sengupta and Wiggins (2014) do not find a significant relation between price dispersion and internet usage for airline fares.

Studies that structurally estimate search (and switching) costs include Giulietti et al. (2014, electricity retail), Hong and Shum (2006, economics and statistics textbooks), Honka (2014, US auto insurance) or Koulayev (2014, hotel search platforms), and studies that directly analyze the search behavior of consumers include Baye et al. (2009, broad range of products) and Hortacsu et al. (2015, residential electricity market).[11] Two findings may be generalizable. First, search costs are substantial. Giulietti et al. (2014), e.g., analyze switching behavior in electricity retail, and estimate that roughly half the households had search costs exceeding 52 Pounds in the year 2005. Second, search is endogenous. Baye et al. (2009) find that a firm enjoys a 60% jump in its clicks when it offers the lowest price at a comparison site. Hortacsu et al. (2015) find that – while households rarely search for alternative retailers in electricity – they search more after a "bill shock" in the previous months. Moreover, households attach a brand advantage to the incumbent.

Summarizing the empirical literature, there are a lot of studies looking at the determinants of prices or price dispersion using a variety of competition measures such as the number of firms and/or the types of firms. Most studies find a negative effect of the number of firms on price dispersion. Studies employing direct measures of search intensity are, however, rare, and they most often find an inverted U-shaped relation between price dispersion and share of informed consumers. However, these studies either do not account for the endogeneity of search or do not account for the type

---

[11] See Baye et al. (2006) for a more comprehensive overview as well as Kim et al. (2011), and De los Santos et al. (2012).

differences across firms. To the best of our knowledge, we are the first study employing (1) a direct measure of search intensity, (2) recognizing its endogenous nature, and (3) distinguishing between types of firms (incumbents versus entrants). As we will show below, this is essential, because firms behave differently depending on the importance they attribute to informed versus uninformed customers.

2.2.Relevant theoretical literature on price dispersion and hypotheses

Our empirical approach builds on the existing theoretical literature on the nexus between pricing, competition, and search costs. There are numerous theoretical models trying to explain the impact of asymmetric consumer information, which arises from differing search and/or switching costs, on price dispersion. Most relevant for our setting is a clearinghouse model developed by Stahl (1989), which is based on Varian's (1980) model. We also refer to other theoretical works that generally build on these models. Besides, Janssen and Moraga-González (2004) develop a non-sequential fixed sample size search model that bears important implications for our empirical analysis.

Stahl's (1989) model assumes a fraction of consumers μ who have no search costs ("shoppers") and thus are fully informed about the full price distribution in the market. These consumers have information about all available prices and buy from the cheapest supplier. The remaining share of consumers $(1-\mu)$ pays a search cost for each price quotation and engage in sequential search until they obtain a price quotation that is below their endogenously determined reservation price.

As opposed to Stahl (1989)'s model, where informed consumers are assumed to have zero search costs, in the electricity retail market informed consumers still have positive search costs. Nonetheless, the informed consumers obtain *all* available price quotes for their respective zip code from an online comparison website (i.e. a clearinghouse), which may severely lower their search costs compared to the uninformed consumers. Hence, the fraction of informed consumers who gather information from a clearinghouse will only buy at the lowest price.[12] The other fraction of uninformed or brand loyal (see Baye et al., 2004) consumers will stay with their incumbent suppliers and purchase at higher prices. Consequently, we interpret consumers who collect price quotes from online search platforms as informed consumers (i.e. "shoppers") as presented in Stahl's model (μ).

[12] Similarly, Chandra and Tappata (2011) assume "shoppers" to access a clearinghouse and buy from the cheapest supplier, whereas the uninformed "nonshoppers" buy from random stores.

According to Stahl (1989), when consumers have asymmetric search costs ($0 < \mu < 1$), price dispersion represents a Nash equilibrium because firms do not follow pure strategies but randomly draw from an equilibrium price distribution (i.e. mixed strategies).[13] Barron et al. (2004, p. 1049) stress that all firms get the same expected profit form mixed strategies: "a low price reduces the returns from sales to the uninformed, but this is exactly offset by the increase in expected returns arising from the increased likelihood that the seller will be the lowest-priced seller, and thus sell to all of the informed consumers." Moreover, there are two extreme cases, for which price dispersion vanishes. When all consumers are uninformed ($\mu = 0$), all firms charge monopoly prices (Diamond's outcome). On the contrary, when all consumers are fully informed ($\mu = 1$), all firms charge Walrasian prices (Bertrand outcome).

Stahl (1989) shows that as the proportion of informed consumers increases gradually from $\mu = 0$ to $\mu = 1$, the price distribution shifts downwards monotonically and the average price falls. Nevertheless, it is not clear-cut how the support of the price distribution behaves as μ shifts from zero to one. Stahl (1989) argues that as the number of "shoppers" increases, both the upper and the lower bounds of the support shift downward (see also Tang et al., 2010). This argument seems somewhat ad-hoc for the upper bound of the distribution, as expected prices may fall even though the upper bound of the distribution (i.e. the incumbent supplier's price in our study) may stay unchanged or may even increase.

In Stahl's (1989) model, price dispersion follows an inverted U-shape with an increasing fraction of informed consumers between the two extreme cases of monopoly pricing and the Bertrand outcome for the following reasons (see also Baye et al., 2006). From the starting point of $\mu = 0$, where all firms charge the same monopoly price, an increase in the fraction of informed consumers (μ) provides firms with an incentive to undercut prices in order to attract consumers who search for prices. As a result, price dispersion increases. Once μ reaches a certain threshold, more mass of the price support shifts downward and hence price dispersion starts to decline until it eventually converges to the Bertrand outcome. Pennersdorfer et al. (2014) underpin this statement by providing a mathematical proof. Chandra and Tappata (2011) provide an intuition for the behavior of price dispersion, as low gains from search arise for consumers when search intensity is either low or high, however, gains are higher when search intensity takes up an intermediate level.

[13] Varian (1980) provides the same finding.

Janssen and Moraga-González (2004) provide a non-sequential search model where search is endogenous to *firm entry* and where the fraction of uninformed consumers searches with differing intensity. Two counteracting forces are prevalent in the market: The *business-stealing effect* causes firms to charge low prices in order to attract informed consumers who engage in search. The *surplus-appropriation* effect lets firms charge high prices for the uninformed consumers who do not compare prices. Price dispersion may be an equilibrium result in this model as it may be beneficial for firms to either charge a high price for customers with medium or low search intensity or a low price for customers with high search intensity. Similarly, Chandra and Tappata (2011) postulate that, holding consumer information constant, an increase in the number of firms in the market brings about two effects. First, each firm loses parts of its uninformed consumers. Second, the probability of having the lowest price decreases exponentially. Both effects cause firms to set more extreme prices increasing price dispersion. Moreover, they show that *price dispersion increases with higher competition at a decreasing rate*.

In our setting, the consumers willing to stay with their local incumbent suppliers at higher tariffs, as they do not engage in search, correspond well with Janssen and Moraga-González's (2004) consumers with medium or low search intensity. On the other hand, the informed consumers of the cheapest entrant firms may represent the group of consumers with high search intensity.

Janssen and Moraga-González (2004) show that when consumers *search with high intensity*, an increase in the number of firms causes prices to drop at a declining rate. The reason is that it may be beneficial for a firm to undercut prices since with a low number of firms in the market the business-stealing effect dominates the surplus-appropriation effect. With a large number of firms, additional entry weakens the business-stealing effect and causes prices to fall less rapidly (and eventually increase). In other words, when the number of firms in the market is low, a firm may profit from charging the lowest price in order to attract the informed customers. Once there is a sufficiently large number of firms in the market, the probability of charging the lowest prices approaches zero and it may be worthwhile to charge higher prices and target the uninformed consumers. Consequently, our cheapest entrant's price may descend due to competition at a decreasing rate (U-shape).

On the contrary, when consumer *search intensity is low*, an increase in the number of firms in the market causes prices to first rise and then stagnate. "(…) When the number of firms in the market grows without limit, the probability that a firm is undercut by some other firm converges to 1. It is precisely for this reason that in the limit economy firms set their prices ignoring the fully-informed consumers altogether and concentrating on attracting the less-informed consumers." (Janssen and Moraga-González, 2004, p. 1107). This implies that our local incumbent suppliers may *increase* their prices or stay put when competition in the market goes up. What is more, both price effects for the entrants and incumbents cause price dispersion to follow an inverted U-shape relation with the number of firms in the market.

Morgan, Orzen, and Sefton (2006) support the results by Janssen and Moraga-González: An increase in the number of firms leads to lower average prices for the informed consumers. On the contrary, more firms in the market lead to higher average prices for the uninformed consumers because firms' incentives to undercut prices in order to attract "shoppers" vanish.

In Table 1 we summarize the main predictions of the relevant theoretical literature on our main variables of interest, which we will test empirically. Due to the reasons given above, we expect price dispersion to be related to both consumer information and number of competitors in an inverted U-shape manner.

For incumbent prices, the predictions on the initial effects of consumer information are not clear-cut. Incumbents may cater to the less informed consumers (surplus appropriation) or alternatively they may decrease prices if the upper bound of the price distribution also declines in response to more consumer search (business stealing). Which effect dominates is ultimately an empirical question. Eventually, i.e. with a lot of consumers searching in the market, we expect also incumbent prices to decline in response to additional consumer search. The predictions for the effects of consumer search on entrant's prices are more clear-cut. Entrants have to lure customers away from incumbents, thus they decrease prices in response to customer search, albeit at a decreasing rate.

Applying the theory of Janssen and Moraga-Gonzales (2004) to our problem, the predictions on the effects of the number of competitors on (incumbent and entrant) prices are clear-cut, albeit in opposite directions. Incumbents operate in a low to middle search intensity equilibrium, and therefore are expected to increase prices in response to firm entry (at a decreasing rate). Entrants operate in a high-search intensity situation and have to decrease prices to attract informed consumers (at a decreasing rate).

Table 1: Predicted effects of main variables of interest on price dispersion and prices

	Price dispersion		Price incumbent		Price entrant	
	Effect	Reason	Effect	Reason	Effect	Reason
Consumer information (μ)	+	Mixed strategy; incentive for firms to undercut prices to attract searchers increases (Stahl, 1989)	−/0/+	Incumbents may cater to less informed consumers with lower elasticity of demand (surplus appropriation). In contrast, Stahl (1989) and Tang et al. (2010) predict the incumbent's price (upper bound of the price distribution) to decline.	−	Intensity of price competition goes up; demand shifts to lower-priced firms (Stahl, 1989)
Consumer information squared (μ^2)	−	More mass of price support shifts downward; Extreme case, u=1, Bertrand outcome (Stahl, 1989)	−/0	Eventually incumbent prices may stagnate or decrease to retain some demand (Stahl, 1989).	+	Price approaches lower bound at marginal cost (Bertrand outcome, Stahl, 1989)
Competitors (N)	+	Decreased (increased) frequency to charge intermediate (extreme) prices; business stealing effect strengthened for entrants, surplus appropriation effect strengthened for incumbents (Janssen and Moraga-Gonzales, 2004)	+	Surplus appropriation dominates business stealing effect; probability to be undercut converges to 1; incumbent concentrates on less-informed consumers (i.e. low to middle search intensity equilibrium; Janssen and Moraga-Gonzales, 2004)	−	Business stealing dominates surplus appropriation effect (i.e. high-search intensity equilibrium; Janssen and Moraga-Gonzales, 2004)
Competitors squared (N^2)	−	Strengthening of the two effects (business stealing for entrants and surplus appropriation for incumbents) weakens with more firms (Janssen and Moraga-Gonzales, 2004)	−	As N approaches infinity, there is a price cealing, the monopoly price (Diamond outcome, Stahl, 1989) or firms randomize between marginal cost pricing and monopoly price (Janssen and Moraga-Gonzales, 2004)	+	As N approaches infinity, the business stealing effect is weakened since the probability of having the lowest price approaches zero (Janssen and Moraga-Gonzales, 2004)

3. Background on Electricity Retail

Almost 17 years have passed since European electricity markets were liberalized ending the former local monopoly regencies in electricity retail. In 2014, a German household could choose on average between 155 electricity retailers each providing 3.4 tariffs.

Table 2 shows potential yearly savings (i.e. price dispersions) between the incumbent's base tariff and the cheapest supplier for standard households.

Table 2: Potential savings (€/year) from switching from the incumbent to the cheapest entrant at different household sizes

	Obs	Mean	SD	Min	Max	% of all HH
1 P HH (2000 kw/year)	30,982	121.18	25.52	15.18	226.91	36.51
2 P HH (3500 kw/year)	30,968	196.04	39.38	8.07	353.52	35.29
4 P HH (5000 kw/year)	30,963	277.05	52.71	0.56	491.93	10.68

Notes: "HH" stands for household. "Obs." are zip code-year observations.

Although switching the provider substantially helps the budget (see Table 2) only a small fraction of German households has taken advantage of it until today. However, the share of prospective switching customers used to grow in recent years (see Figure 1) as online price comparison sites have significantly disburdened costs of searching cheaper providers, and customers can now easily switch the provider within the time period of a TV break.[14] Nevertheless, 79% of German households were still supplied by their former local *incumbent utility* in 2014 even though switching to one of the alternative retailers – referred to as *entrants* hereinafter – would clearly generate substantial savings.[15] In general there are three candidate explanations for these frictions in consumer switching, namely search costs, switching costs and brand effects. Since electricity is a homogenous product brand effects should not play a major role.[16] There is no reason to assume that the incumbent supplier is more secure in providing electricity since the incumbent has the legal obligation to guarantee a continuous

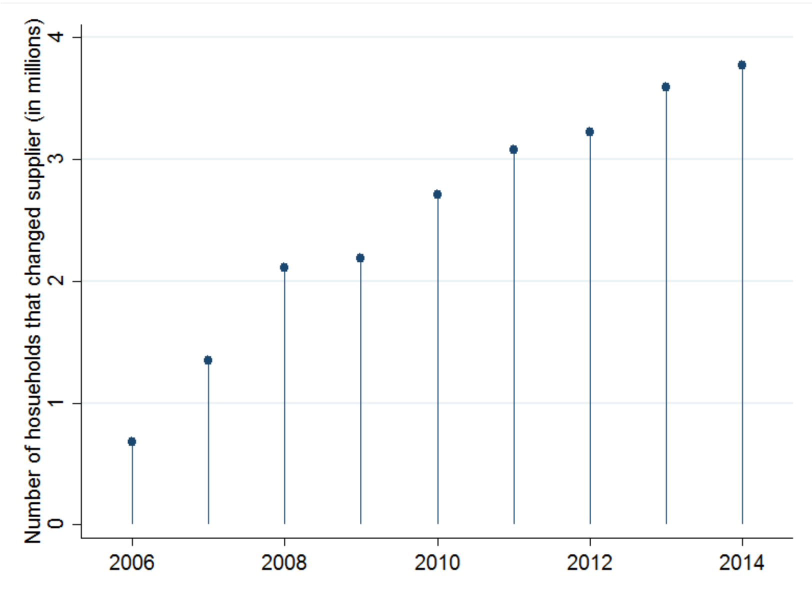
[14] According to a survey 80% of the switchers searched online for alternative providers (A.T. Kearney, 2012)

[15] See Monitoringreport (2014, p. 114) from the German Federal Cartel Office and the Federal Network Agency.

[16] Hortacsu et al. (2015) analyze the Texas residential electricity market between 2002 and 2006 and find a brand effect for the incumbent. However, they also point out that it diminishes over time due to consumers' learning.

provision of electricity to the customer.[17] Even when an entrant goes bankrupt, the former incumbent has to back its customers with electricity supply at no additional fees.[18] Nevertheless, not all consumers might be aware of this safety net. Switching costs are also not likely a candidate explanation since a) the switching process is an automated process and conducted by the provider a consumer opts for switching to[19] and b) the cancellation period for the incumbents' base tariffs is only two weeks by law. Therefore, we consider search costs as the main source for observed differences in end-user prices.

Figure 1: Development of supplier changes in electricity retail in Germany



Note: Data are obtained from Bundesnetzagentur (2015).

A comparison portal requires entering information on a customer's expected consumption over one year. Because electricity tariffs are multi-part tariffs and electricity consumption depends on exogenous factors (for instance weather) this opens

[17] In this regard, Hortacsu et al. (2015) mention the possibility that customers may believe that the incumbent supplier may exhibit a higher supply security although this is in fact not true.

[18] Indeed, two of the bigger alternative providers went bankrupt in 2011 (Teldafax) and 2013 (Flexstrom), respectively.

[19] The new supplier will automatically overtake all switching activities for its new customer, such as unsubscribing from the old supplier, registration, etc., at no switching fees. Giulietti et al. (2014) investigate the U.K. electricity retail market, where the switching process can be easily compared with Germany's. They highlight that "Search is perceived by consumers as being significantly more difficult than switching." (p. 561)

the possibility to over- or underestimate consumption and as a result the initial choice might not be the best choice (Waddams and Wilson, 2010). Furthermore, tariffs might include a switching bonus thereby making it only the cheapest tariff for the first year and without engaging in search again the tariff can be disadvantageous for the consumer after the first year. This tariff strategy relies on profits from automatic contract extension due to expected consumer inertia. Montgomery et al. (2004) argue that the use of price-comparison "shopbots" will persist on a rather low level, given the various cognitive costs of evaluating many alternatives. Nevertheless, online price comparison websites have been gaining substantial attention from potential switchers over time.

4. Data and Empirical Model

The novel and unique data provide a key feature of our paper, since we are able to empirically test the effects of both consumer search and competition on price dispersion (and prices) at the German zip code level for the period 2011–2014. It is of particular relevance that we are able to directly measure consumer search from our information about online search requests (i.e. clicks) for alternative tariffs at major price comparison portals. Besides having suitable measures for our dependent variables and variables of interest, we include a rich set of control variables. In section 4.1., we present descriptive statistics and information on the spatial distribution of some key variables. In section 4.2., we describe our estimation strategy for the econometric approach.

4.1.Data and Variables

Our data stems mainly from two sources. From *ene't*, a German software and data provider for the electricity industry, we received detailed data on consumer search activities, retail electricity prices and cost components. The marketing company *Acxiom* provided data on structural household characteristics in Germany. Since the *ene't* data are structured at the monthly frequency, we aggregate them to match with the *Acxiom* data, which are at the yearly level. This corresponds well with the length of a typical electricity contract. Our data span the period 2011–2014. The spatial data resolution is at the German zip code level. In total, the database contains 8,224 zip codes. This bears an advantage over other empirical papers, which have to make assumptions about the delineation of local.[20] Consumers can only choose among electricity tariffs from retailers that supply in their respective zip code. We first describe how we construct our main variables from the data and subsequently provide information on control variables.

Price Dispersion

Our measure for price dispersion (PD) represents the relevant price range, i.e. the difference between the incumbent's price and the cheapest entrant tariff.[21] We define the incumbent's price as the upper bound as even if there were more expensive entrants consumers would not consider them as alternatives. The cheapest price describes the lower bound as in general a provider mostly attracts consumers when it has the lowest price in the market.[22] Prices are measured as end-user tariffs including the fee for the electricity itself, grid charges and other charges, and taxes including the value added tax. The price data are at the zip code level. They are observed on a due date each month and subsequently transformed into yearly averages.

Since we have knowledge on the cost components of electricity supply (described later) we also compute *markups* and *Lerner indices*. Markups are calculated considering all known spatially varying costs such as fixed and variable grid charge components and concession fees as well as spatially constant but time varying costs, such as the renewable energy surcharge.[23] These data are described in greater detail below. We also utilize the day-ahead spot price (Phelix base) at the European Power Exchange (EPEX)[24] in order to approximate the costs of purchasing wholesale electricity. Even if electricity is purchased otherwise (e.g. through direct contracts, OTC markets, forward contracts, etc.) or generated by vertically integrated power plants, the spot price represents the opportunity cost of purchasing electricity. The cost components we observe (abstracting from heterogeneous advertising and hedging strategies etc.) do not differ between incumbents and entrants and thus the difference in markups between incumbents and entrants is the price dispersion plus the difference

[20] For example, in Gasoline retail Pennerstorfer et al. (2014) make assumptions about the search radius of consumers.

[21] The price range is a common measure for dispersion (see, e.g., Chandra and Tappata, 2011; Pennersdorfer et al. 2015). An alternative is the variation in prices, which is not available to us since we only observe the tariffs of the incumbent and cheapest entrant each month.

[22] An exemption may be the fraction of consumers who are willing to switch to a green power supplier at a higher tariff.

[23] In German "EEG Umlage".

[24] EPEX represents the relevant power exchange for Germany.

in value added tax due to the higher prices charged by the incumbent.[25] Accordingly, the Lerner Index represents the ratio of the markup to the price.

Consumer Information

We are interested in the impact of consumer information on actual prices and on price dispersion. The direct measure for consumer information is constructed from individual search queries provided by *ene't*. The data contain detailed information on individual search patterns enabling us to construct a dataset on regional consumer search intensity. The sample period ranges from March 2011 to December 2014. The database covers all search activities conducted on several well-known online price comparison platforms including Toptarif.de (top tariff), Stromtipp.de (power tip), Energieverbraucherportal.de (energy consumption portal) and mut-zum-wechseln.de (courage-to-change), of which Toptarif.de is the biggest platform by far.[26] For each query we observe the timestamp of the search request, the zip code for which the offered electricity tariffs are requested, the (expected) yearly consumption entered into the search mask, the type of search request (household or industrial customer) as well as a search session ID indicating the order of the queries of the respective searchers. In sum, we have information on 35,855,071 search requests 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.

In our analysis we will focus on household consumers. The main application is conducted on a typical two person household with a consumption of 3,500 kWh per year, which represents 35% of all households.[27] We are not able to observe actual switching, because clicking on a certain supplier's tariff at the online comparison website redirects the searcher to a website where the switch can 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.[28]

---

[25] There are in some cases also additional but rather small deviations between the incumbents' and the entrants' costs resulting from slightly different costs for metering services.

[26] Toptarif is one of the three big 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.

[27] Our findings are fully robust to the estimations for different household sizes (i.e. different consumption levels). We provide these results upon request.

[28] Brynjolsson and Smith (2001) confirm this and find that factors that drive clicks are reasonable and relatively unbiased indicators of sales in their study on online book purchases

On this basis we construct a measure of consumer information as follows. Because many searchers conduct several search requests within a search session (e.g. comparing prices for different levels of consumptions) we only count the number of search sessions and refer to a consumer conducting a search session as being fully informed regardless of the depth of the search activity.[29] Furthermore, we exclude 551,256 search sessions which exclusively consider eco-label certified tariffs. Those searches are most likely not predominantly price but rather ideology related and on average 152 Euro more expensive than the cheapest tariff.[30] We then compute our measure of consumer search μ by aggregating the search sessions within a zip code on a yearly basis and subsequently divide this value by the number of households within the zip code in the respective year. 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 measure of consumer information.[31]

Competition

To measure spatial competition we use the number of electricity retail suppliers within a zip code. The number of competitors in a zip code varies between 55 and 198 in our observation period. These data were also provided by *ene't*.

Control Variables

Data on costs are obtained from *ene't*. We distinguish between three cost factors with spatial variation, namely (i) fixed grid charges, (ii) variable grid charges, and (iii) concession fees.[32] Fixed and variable 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). The concession fee has to be paid by the system operator to the respective zip code (i.e. zip code) for the right to install and operate electric cables on public roads.

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

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

[31] Figure 4 in the Appendix provides a histogram on consumer information before and after the data trimming. Figure 5 and Figure 6 provide histograms on the distribution of the number of competitors and price dispersion.

[32] In the regressions, we cannot include other costs, which only vary with time but not across firms (e.g. wholesale spot price as a proxy for the purchase price of electricity), because these are captured by year fixed effects.

Hence, the concession fees vary at the zip code level. All of these cost components represent parts of the annual electricity bill.

Other control variables refer to structural household characteristics which we received from *Acxiom*. These data include the (i) average household size, (ii) the share of job seekers, and (iii) income brackets (which may control for consumers' sensitivity to prices and thus impact a household's likelihood to search).

Identification

We are aware of a likely reverse causality (i.e. the problem of endogeneity) between prices and both a) consumer information and b) the number of competitors.[33] There is also likely an effect of prices (and thus also price dispersion) on consumer search because consumers increase search efforts when their electricity bill is high[34] and price dispersion enables consumers to reduce their expenditures (thus increasing their gains from search). As regards b), the levels of prices and price dispersion may also affect entry decisions. For example, high prices may attract new entrants.

We circumvent the endogeneity issue by applying *instrumental variable techniques*. We employ the following instruments for *consumer information*: (1) the share of households with a household head below the age of 40 (U40), (2) the share of households that moved into the zip code (New HH), and (3) the share of households that moved away from the zip code (Moved HH). We instrument for *competition* with (4) the number of households in the zip code (# HH). The reason why we believe these measures to be adequate instruments are as follows:

With respect to U40, the assumption is that younger people are more familiar with the internet and thus have lower burdens to gather information online at price comparison websites. Under this assumption the share of U40 households will exogenously shift average search intensity in the respective zip code (but not prices or price dispersion). With respect to New HH, we emphasize that new households receive information on their electricity contract from their base supplier (the former

---

[33] In this study, we do not look at a potential relation between search intensity and the number of firms, but employ instrumental variable techniques. According to Janssen and Moraga-González (2004), search intensity depends on the number of firms in an intricate way. However, as the number of firms is always above 50 and also consumers have no information about the exact number of firms in the market in the absence of searching we argue that consumers' search intensity is unaffected by the number of firms in their local market.

[34] See also Hortacsu et al. (2015) who point to higher switching rates in the summer month as a reaction of consumers to high electricity bills from air conditioning.

incumbent), which confronts them with their contracts' monetary conditions and thus increases the likelihood to search for alternative tariffs.[35] Regarding Moved HH, we postulate that since moves away from a zip code are not predominantly spontaneous, a household is not likely to search for a new energy contract in the year it moves away. The number of households within the respective zip code represents a measure of market size, which we expect to exogenously impact competition. A larger market promises more potential customers and, as a result, has an influence on competition.[36]

Table 3 provides descriptive statistics of our variables employed in the regressions. Graphical illustrations of the spatial distribution of consumer search intensity, the number of competitors and price dispersion are made available in Figure 7 to Figure 9 in the Appendix.

---

[35] Indeed, a disproportionally high share of switchers are new households (Monitoring Report, 2014).

[36] We see no reason to assume that the number of households has an influence on prices through the presence of economies of scale in retail electricity. This is for the reason that electricity trading represents the main task of electricity retailers, which does not exhibit economies of scale – as opposed to other electricity sectors, such as power generation or transmission.

Table 3: **Summary statistics**

	Mean	S.D.	Min	Max	Obs.
Price and Price Dispersion					
Price dispersion (€/a)	196.04	39.38	8.07	353.52	30,968
Price incumbent (€/a)	1003.79	77.90	761.01	1204.15	30,968
Price entrant (€/a)	807.75	58.74	657.19	903.03	30,968
Information and Number of Competitors					
% Informed households (μ)	9.28	6.46	0.00	36.06	30,968
# Competitors (N)	132.33	24.63	54.58	198.00	30,968
Markups					
Markup dispersion (€/a)	164.89	33.28	3.38	297.07	30,968
Markup incumbent (€/a)	162.84	40.94	-72.37	299.24	30,968
Markup entrant (€/a)	-2.04	22.60	-118.60	67.22	30,968
Lerner Index					
Lerner index dispersion (€/a)	0.164	0.028	0.006	0.345	30,968
Lerner index incumbent (€/a)	0.161	0.034	-0.095	0.271	30,968
Lerner index entrant (€/a)	-0.003	0.029	-0.179	0.086	30,968
Costs					
Variable grid charge (€ cent/kw)	5.32	0.88	1.57	8.67	30,968
Fixed grid charge (€/a)	16.14	11.43	0.00	75.00	30,968
Concession fee (€ cent/kw)	1.49	0.35	0.00	2.39	30,968
Phelix Spot Price (€/MW)	44.40	6.69	35.89	54.03	30,968
Instruments					
% Moved households	24.60	5.00	7.70	55.00	30,968
% New households	5.60	2.10	0.70	79.00	30,968
% Head of household under the age of 40	5.50	2.00	1.20	79.50	30,968
# Households	5,025	4,666	16	29,891	30,968
Household characteristics					
% Job seekers	5.50	3.60	0.00	24.50	30,968
% Household head with income < 25 th. €/a	39.20	7.50	2.40	83.00	30,968
% Household head with income 25–50 th. €/a	32.20	2.50	9.90	59.10	30,968
Average household size (persons)	2.11	0.18	1.52	2.55	30,968

Notes: "Obs" are zip code-year observations. Negative markups and Lerner indices may occur since (cheapest) entrants pay a switching bonus and therefore may condone losses in the first year.

5. Econometric Model

We aim at estimating the impact of consumer search intensity as well as the degree of competition on price and price dispersion. Due to the above discussed endogenous relation between dispersion and consumer search we apply instrumental variable techniques. The two first-stage equations we estimate are as follows:

$$\mu_{it} = \delta_1 Num\_hh_{it} + \vartheta_{11} U40_{it} + \vartheta_{21} New_{it} + \vartheta_{31} Move_{it} + X_{it}\theta_1 + \gamma_{t1} + \rho_{i1} + u_{it1} \quad (1)$$

and

$$N_{it} = \delta_2 Num\_hh_{it} + \vartheta_{12} U40_{it} + \vartheta_{22} New_{it} + \vartheta_{32} Move_{it} + X_{it}\theta_2 + \gamma_{t2} + \rho_{j2} + u_{it2} \quad (2)$$

The subscripts i, j and t indicate zip codes, supply areas of the incumbents, and years, respectively. Our dependent variables, the measure for consumer search intensity (μ) and the number of competitors (N), are linear projections of the instruments *Num_hh*, which is the number of households within a zip code, $U40$, which is the share of households with the head of the household being under the age of 40, *New*, which represents the share of households that moved into the respective zip code, and *Move*, which represents the share of households that moved away. X is a vector of all control variables as described above. In our preferred regression we also add regional fixed effects for the 777 incumbents (ρ_j). The reason for including incumbent fixed effects instead of zip code fixed effects is that an incumbent generally operates in multiple zip codes and also the number of competitors rather varies on distribution grid areas (generally the incumbents supply area).[37] Finally, year dummies (γ_t) capture unobserved time varying components.

Besides the terms in levels, we create instruments for the squared terms of consumer information (μ_{it}^2) and the number of suppliers (N_{it}^2) in order to allow for non-linear relations in the structural model. The instruments are the squared predictions of the first stages for μ_{it} and N_{it}, respectively, from equations 1 and 2: $\hat{\mu}_{it}^2$ instruments for μ_{it}^2 and $\hat{N}_{it}^2$ for N_{it}^2.[38]

[37] As there are 8,224 zip codes the average incumbent operates in 10.6 zip codes.
[38] See Wooldridge (2010, p. 262) on this approach.

Thus, the structural equation we estimate takes the following form:

$$y_{it} = \beta_1 \hat{\mu}_{it} + \beta_2 \hat{\mu}_{it}^2 + \beta_3 \hat{N}_{it} + \beta_4 \hat{N}_{it}^2 + X_{it}\theta_3 + \gamma_{t3} + \rho_{j3} + u_{it3}. \quad (3)$$

where y denotes price dispersion, price incumbent, and price cheapest entrant, respectively. In alternative specifications, y may alternatively indicate markups or Lerner indices.

6. Results on Price Dispersion and Prices of Incumbents and Entrants

In this section we discuss the empirical results. Table 4 shows the estimates for Price Dispersion (PD). Furthermore, Table 5 and Table 6 contain regression estimates of incumbents' and the entrants' prices, respectively. In all tables we report reduced form and IV estimates with and without fixed effects (FE). In the case of IV regressions the respective test statistics suggest our instruments to work properly (in the fixed effects models). The instruments are sufficiently strong and the model is identified as a whole as shown by the high Kleibergen-Paap statistics.[39] Also, the Hansen J test indicates that the instruments are valid as they are orthogonal to the errors.[40] As the Durbin-Wu-Hausman test supports the endogeneity assumption by rejecting the null hypothesis of consumer search and the number of competitors being exogenous variables, the IV estimations with fixed effects (IV FE) are, thus, our preferred specifications over non-IV OLS and FE results.

The regression results strongly support the hypothesis of an inverted U-shaped impact of *consumer search* on price dispersion, as $\hat{\mu}$ is positively significant and $\hat{\mu}^2$ is

[39] The Kleibergen-Paap rk Wald F-statistic (2006) is the multivariate analogue of the first stage F-test. In case of multiple endogenous variables the Kleibergen-Paap statistic indicates whether the employed instruments do not only identify the endogenous variables individually in the corresponding first stage regressions (as the Angrist-Pischke first stage F-statistic does), but also whether the endogenous variables are simultaneously identified. This is not necessarily the case, for example, when the instruments are highly correlated. In this circumstance, the same instruments would separately identify each endogenous regressor but the equation as a whole would not be identified. Indeed, in the case of only a single endogenous variable the Kleibergen-Paap statistic is identical to the standard first stage F-statistic. The null hypothesis of the Kleibergen-Paap test is that the structural equation is under-identified (that is, the rank condition fails). Even though critical values do not exist for the Kleibergen-Paap statistic, the critical values calculated by Stock and Yogo (2005) are generally applied. As a rule of thumb a value of the test statistic above ten indicates identification of the model.

[40] First stage regressions are reported in Table 7 in the Appendix.

negatively significant.[41] This result underpins Stahl's (1989) theoretical finding that as μ varies between zero (Diamond Outcome) and one (Bertrand outcome), price dispersion increases first and then declines.

Moreover, we also find an inverted U-shape relation between price dispersion and *competition* ($\widehat{N}$ is positive and significant, $\widehat{N^2}$ is negative and significant). Interestingly, Tables 4 and 5 indicate that price dispersion is caused by *contrasting* pricing strategies of the incumbent and the entrant rather than by different intensities of the same strategies.

The incumbents' pricing reaction to increased *consumer search* is best described as inverse U-shaped – an initial *increase* in price is followed by a decrease after a certain degree of consumer information is achieved – presumably because the *surplus appropriation effect* initially outweighs the *business stealing effect* due to the much larger share of uninformed consumers of the incumbent.[42] This finding contradicts Stahl's (1989) theoretical argument that an increase in μ shifts also initially the upper bound of the price support downwards.

The contrary is true for the group of entrants. Their pricing strategy with respect to consumer information follows a U-shape since the *business stealing effect* clearly outplays *the surplus appropriation effect* as all consumers are with the incumbent to begin with when there are no informed consumers. The more consumers get informed, the higher the market shares of the entrants (due to switching). However, it may be that as not all consumers continue to stay informed (i.e. after having switched to the cheapest provider at a certain time they have a high probability to stay there regardless of whether there are cheaper providers after some time), there is also a surplus appropriation effect for the entrants when more consumers get informed.

With regard to the impact of the degree of *competition* on pricing we observe similar patterns. Price dispersion follows an inverse U-shape with additional firms in the market. This empirical results corresponds well with theoretical findings by Janssen and Moraga-Gonzales (2004). The underlying price setting behavior of entrants and incumbents is as follows.

---

[41] We also apply the method recently developed by Lind and Mehlum (2010) in order to test whether the estimated relationships are actually non-monotonic within the observed data spectrum.

[42] Recall that the incumbent has by far the largest market share since all consumers are initially assigned to the incumbent. Thus, switching requires search efforts and all consumers that are still with the incumbent may be viewed as uninformed. See Janssen and Moraga-Gonzales (2004) on *business stealing* and *surplus appropriation* effects.

As the number of firms increases, the entrants react with price decreases at a decreasing rate. The intuition behind this finding is that entrants try to undercut their competitors in order to expand their market shares. In this case, *the business stealing effect outweighs the surplus appropriation effect.* Moreover, with more and more firms in the market, the likelihood of charging the lowest tariff approaches zero. Hence, firm entry weakens the business stealing effect causing prices of entrants to fall less rapidly and eventually to increase.

On the other hand, the incumbent suppliers react with price increases at a decreasing rate due to intensified competition in the market. In this case, the *surplus appropriation effect offsets the business stealing effect* when N increases. The incumbent concentrates on its less-informed customers and charges a higher price as its incentives to undercut entrant prices to attract informed consumers disappears, since its probability of being undercut approaches one.

For our results we reach the conclusion that the general assumption in the theoretical and empirical literature of an equal endowment of informed (and uninformed) consumers across firms is inappropriate. We show that price setting differs across firms, as incumbents and entrants react differently to consumer information and competition, given their heterogeneous proportions of "shoppers" and "non-shoppers". As a consequence, opposed pricing strategies emerge for firms that are associated with uninformed consumers who are not willing to search for cheaper tariffs (i.e. incumbents) compared to firms associated with informed consumers who engage in search (i.e. entrants).

Table 4: Price Dispersion (PD) estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	435.383***	62.841***	1078.341***	736.278***
	(13.812)	(7.045)	(109.785)	(80.865)
Information^2	-1088.574***	-167.267***	-2977.277***	-2046.275***
	(38.815)	(18.942)	(290.151)	(213.607)
#Competitors (N)	2.026***	1.160***	0.865***	3.287***
	(0.088)	(0.056)	(0.284)	(0.358)
#Competitors^2	-0.007***	-0.003***	-0.005***	-0.009***
	(0.000)	(0.000)	(0.001)	(0.001)
Fixed Effects	NO	YES	NO	YES
Year Dummies	YES	YES	YES	YES
Extreme point (Inf.)	0.200	0.188	0.181	0.180
Inverse U-shape test (Inf.)	0.000	0.000	0.00	0.000
Extreme point (Comp.)	138.633	191.417	80.307	178.102
Inverse U-shape test (Comp.)	0.000	0.029	0.071	0.001
Kleibergen Paap F Stat.	-	-	85.879	56.582
Hansen J stat. (p-val.)	-	-	0.000	0.935
Wu-Hausman Test (p-val.)	-	-	0.000	0.000
Adj. R^2	0.262	0.431	-	-
#Obs.	30,968	30,968	30,968	30,968

Notes: Robust standard errors in parentheses. Estimation of IV models is by IV GMM. Instruments for μ and N are the share of new households, the share of moved households, the share of households where the head is below the age of 40 and the number of households in a zip code. The additional instruments for μ^2 and N^2 are the squares of their respective first stage predictions from equations (1) and (2). All estimations include control variables for regionally varying costs, household characteristics such as income and average household size (persons) as well as year dummies. Significant for * $p < 0.1$,** $p < 0.05$,*** $p < 0.01$.

Table 5: Price estimates for the incumbent

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information	447.602***	53.800***	957.849***	322.813***
	(12.836)	(5.614)	(106.646)	(60.993)
Information^2	-1088.929***	-112.360***	-2375.961***	-734.000***
	(37.290)	(14.755)	(280.656)	(161.437)
#Competitors	1.830***	0.923***	0.184	2.039***
	(0.080)	(0.046)	(0.267)	(0.259)
#Competitors^2	-0.007***	-0.002***	-0.004***	-0.006***
	(0.000)	(0.000)	(0.001)	(0.001)
Fixed Effects	NO	YES	NO	YES
Year Dummies	YES	YES	YES	YES
Extreme point (Inf.)	0.206	0.239	0.202	0.220
Inverse U-shape test (Inf.)	0.000	0.000	0.000	0.000
Extreme point (Comp.)	136.236	184.870	25.567	174.491
Inverse U-shape test (Comp.)	0.000	0.000	1.000	0.001
Kleibergen Paap F Stat.	-	-	85.879	56.582
Hansen J stat. (p-val.)	-	-	0.025	0.309
Wu-Hausman test (p-val.)	-	-	0.000	0.000
Adj. R^2	0.831	0.962	-	-
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

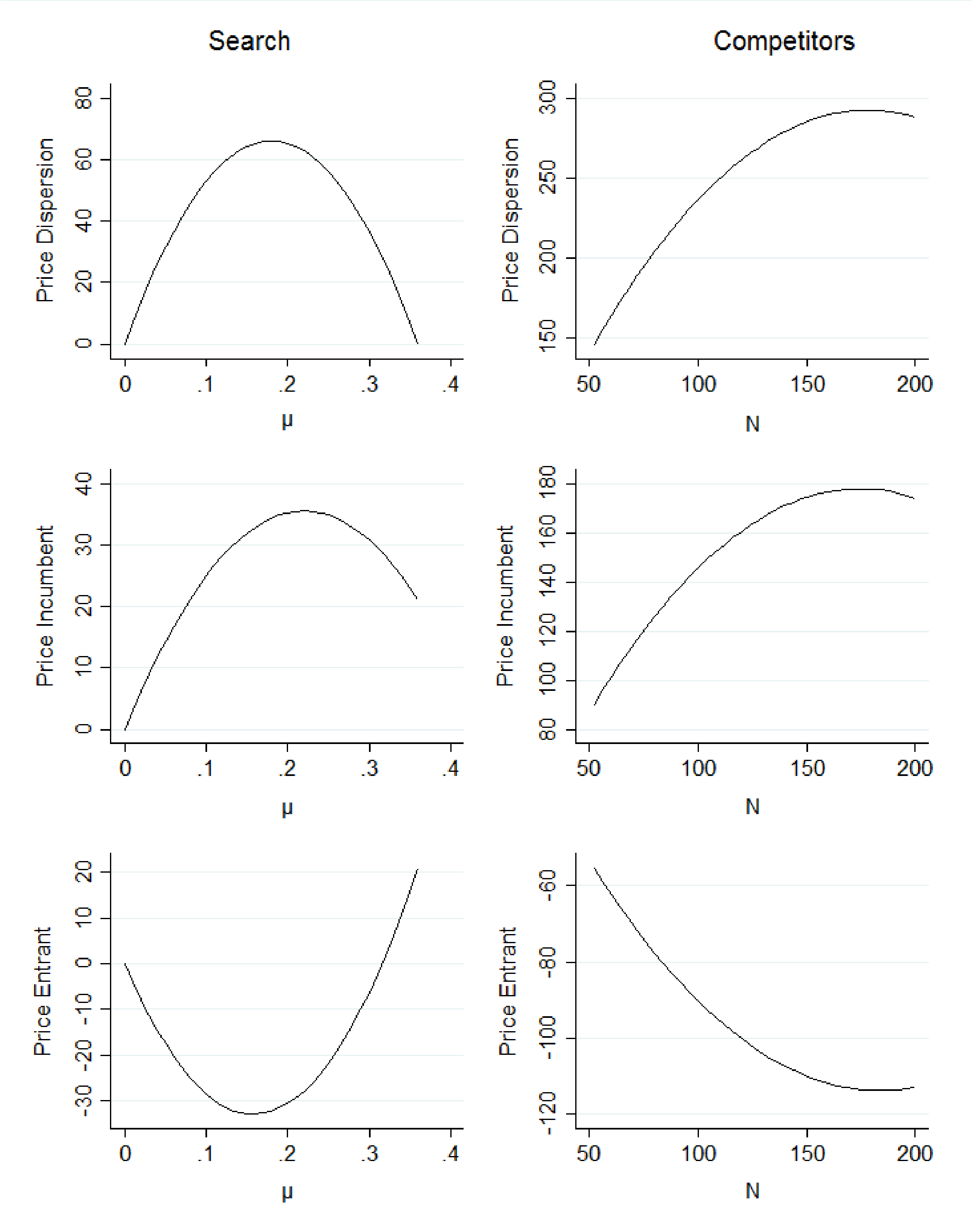
Table 6: Price estimates for the entrant

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information	12.219***	-9.040**	-112.127***	-416.885***
	(4.255)	(4.195)	(35.907)	(43.997)
Information^2	-0.355	54.907***	577.612***	1319.628***
	(11.427)	(11.118)	(95.099)	(115.958)
#Competitors	-0.197***	-0.236***	-0.752***	-1.241***
	(0.030)	(0.036)	(0.100)	(0.202)
#Competitors^2	0.001***	0.001***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)	(0.001)
Fixed Effects	NO	YES	NO	YES
Year Dummies	YES	YES	YES	YES
Extreme point (Inf.)	17.221	0.082	0.097	0.158
U-shape test (Inf.)	1.000	0.000	0.001	0.000
Extreme point (Comp.)	165.762	222.133	184.792	183.431
U-shape test (Comp.)	0.000	1.000	0.115	0.059
Kleibergen Paap F Stat.	-	-	85.879	56.582
Hansen J stat. (p-val.)	-	-	0.000	0.415
Wu-Hausman test (p-val.)	-	-	0.000	0.003
Adj. R^2	0.969	0.973	-	-
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Figure 2 provides a graphical illustration of the effects of both consumer search and competition on price dispersion, prices of incumbents, and prices of entrants according to the IV FE estimates.

Figure 2: Impact of Search intensity and competitors on price dispersion and prices.



Note: Curves represent estimates from columns 4 in Table 4 to Table 6 (IV with Fixed Effects). The image does not allow inference on the actual price and price dispersion levels as the constant is not estimated due to fixed effects.

Robustness

Even though we believe that the chosen modelling approach reflects the market best (also indicated by the highest degree of variance explanation by the instruments) our results are surprisingly robust to a variety of specifications. The main results remain largely unchanged regardless whether we change from incumbent fixed effects to zip code fixed effects. In addition, the findings stay robust if we apply as alternative dependent variables the markup dispersion or the Lerner index dispersion (reported in Table 8 – Table 13 in the Appendix). Also, the results are valid for other household sizes (i.e. other consumption levels) like one-, three-, and four-person households and different fixed effect specifications, i.e. zip code level.[43]

Moreover, the estimates are also confirmed in a semi-parametric context. With semi-parametric estimation, we allow for more flexible functional forms with regards to consumer search and the number of competitors. We apply Yatchew's (1997, 1998) differencing based semiparametric partial linear method.[44] Search intensity and number of competitors, respectively, are estimated non-parametrically by a local weighted scatterplot smoothing function (LOWESS) while the control variables from equations 3 enter the model linearly. We also include control functions, i.e. first stage residuals from equations 1 ($\hat{u}_{it1}$) and 2 ($\hat{u}_{it2}$), respectively, in order to consider the endogeneity issue (two-stage residual inclusion, 2SRI).[45]

The corresponding model specifications are

$$y_{it} = \alpha_{01} + f(\mu_{it}) + \alpha_{11}N_{it} + \alpha_{21}N_{it}^2 + X_{it}\theta_4 + \alpha_{31}\hat{u}_{it1} + \alpha_{41}\hat{u}_{it2} + \gamma_{t4} + \rho_{j4} + e_{it1} \tag{4}$$

and

$$y_{it} = \alpha_{02} + f(N_{it}) + \alpha_{12_{it}}\mu_{it} + \alpha_{22}\mu_{it}^2 + X_{it}\theta_5 + \alpha_{32}\hat{u}_{it1} + \alpha_{42}\hat{u}_{it2} + \gamma_{t5} + \rho_{j5} + e_{it2} \tag{5}$$

[43] Results are available from the authors upon request.

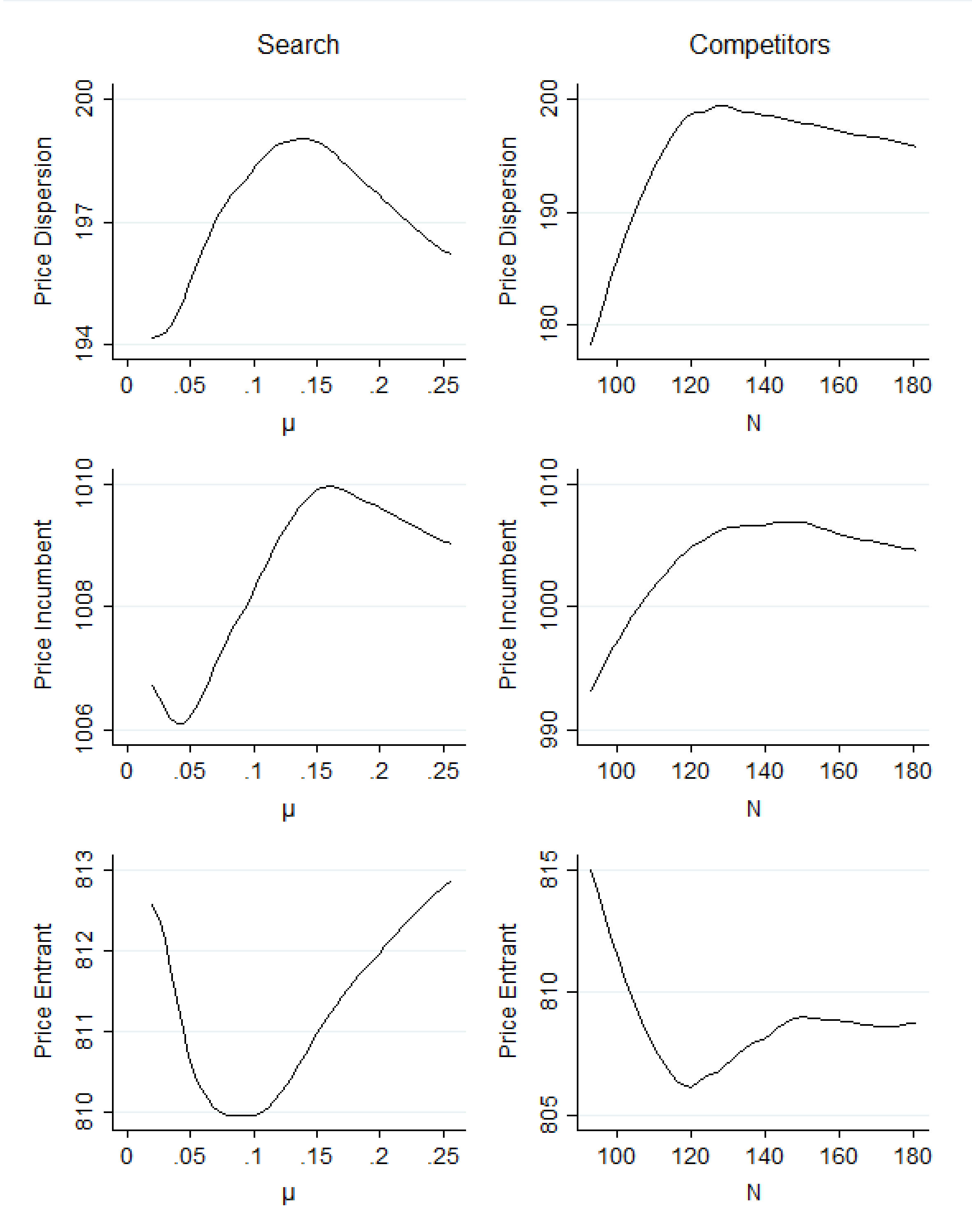
[44] Semiparametric estimators tend to be noisy in sparse regions and we therefore trim the data slightly by cutting the highest and lowest three per cent of the observations for search intensity and number of competitors, respectively, in the respective specifications.

[45] See Blundel and Powell (2004) or Wooldridge and Imbens (2007) on endogeneity in semiparametric models.

We start the description of the semiparametric estimates with the findings for the impact of consumer search. With regard to price dispersion, the parametric estimates are particularly confirmed as the shape of the non-parametrically estimated search function strongly suggests an inverted U-shaped impact of search intensity on price dispersion. Also, the entrants' price adjustments follow a U-shape when search intensity increases. The non-parametrically estimated pricing patterns of the incumbents somehow confirm the parametric estimates, except for low levels of μ (below 5% consumer information).

Turning to the impact of the number of competitors on prices and price dispersion the results particularly confirm the parametric results. Price dispersion increases at a decreasing rate when the number of competitors gets higher and the same is true for the incumbents' prices. Entrants decrease their prices as a reaction to increased competition. However, with about 120 competitors in the market, the effect becomes less pronounced.

Figure 3: Non-parametric fit from semiparametric estimates. Yatchew's (1997) method.



Note: Above figure illustrates the non-parametrically estimated functions of search intensity (left panel) and number of competitors (right panel) from a semiparametric partial linear model as suggested by Yatchew (1998). Nonparametric part estimated by a local polynomial function (LOWESS). Differencing is of order 10. Endogeneity considered through the inclusion of control functions in the parametric part. Fixed Effects models.

7. Conclusion

In this paper we investigate pricing strategies of local incumbents and entrants with respect to consumer information and competition in Germany's electricity retail market. Standard economic theory suggests that prices for a homogenous good, such as electricity, should converge to "the law of one price", as consumers may always switch to the cheapest tariff. Nonetheless, despite the market liberalization 17 years ago, switching rates are low and price dispersion has been persistent. A great share of customers stays with the former monopoly supplier at a relatively high tariff. For a typical two-person household with a yearly consumption level of 3,500 kWh the average incumbent supplier charges EUR 1,004 while the cheapest entrant charges on average EUR 808, implying a relatively large mean price dispersion of EUR 196 or almost 20%.

We stress that significant search frictions (see Stigler, 1961) are present in the market. This is even true despite the recent development of online search platforms, which are likely to have significantly lowered search costs for consumers. Obfuscation due to in transparent tariffs might be a potential explanation (Ellison and Ellison, 2009). A online comparison portal provides consumers with all available electricity retail tariffs in his/her zip code and, thus, serves as a clearinghouse. We therefore follow the relevant theoretical literature of non-sequential fixed sample search models in a clearinghouse environment, to derive hypotheses of how both consumer information and competition influence price dispersion.

In this study, we provide several novelties compared to the existing literature and, therefore, extend our knowledge on the nexus between search frictions, competition and pricing behavior in homogenous goods markets. (1) From our data we are able to construct a *direct measure of consumer information* that is rarely observable to researchers. (2) We infer about *firm heterogeneity* with respect to consumer information, since we distinguish incumbents, which are bound to "non-shoppers" who do not search for alternative tariffs, from entrants, which attract "shoppers" who engage in search. (3) We measure consumer information, competition and the tariffs of the cheapest entrant and the local incumbent in each German zip code for the annual period 2011–2014. This allows for inference about a *very large sample of 8,224 individual markets* (i.e. zip codes) over 4 years. (4) For us *market delineation is not an issue* as in other studies, where assumptions have to be made about relevant markets. (5) We derive empirical results for the effects of *both consumer search and competition* on prices and price dispersion altogether. Hence, we provide a fuller picture on the intricate relations.

The empirical results correspond surprisingly well with hypotheses from the theoretical literature. We estimate an inverse U-shape effect of consumer information on price dispersion, as in Stahl (1989). That is, an increase in consumer information leads to an increase in the incumbent's price at a decreasing rate (inverse U-shape), while the entrant's price declines at a decreasing rate (U-shape). Moreover, price dispersion follows an inverted U-shape when competition in the market increases. This is in line with theoretical findings by Janssen and Moraga-González (2004). The inverse U-shape effect of competition on price dispersion, again, stems from opposing pricing reactions of incumbents and entrants. Incumbents tend to increase their prices at a decreasing rate when competition intensifies. On the other hand, entrants' prices drop at a decreasing rate with additional firms in the market.

A variety of empirical tests support the validity of our results. Initially, we truncate the 4% search request for tariffs with eco labels to avoid product differentiation effects. Including these requests does not alter the findings. Also, we estimate the models for different household sizes (e.g. 1-, 3-, and 4-person households) at different consumption levels and yield robust estimates. Our results are generally unchanged regardless whether we include incumbent (or distribution grid) fixed effects or zip code fixed effects. Changing the dependent variables from price dispersion to markup dispersion or Lerner index dispersion reaches similar conclusions. Besides, we allow for flexible functional forms of consumer search and competition by semi-parametric estimation and find supportive evidence for our parametric estimates. Last but not least, our results stay valid when we apply the estimations on monthly level.

Our results show that the assumption of equal shares of informed versus uninformed consumers across firms – as generally assumed in the theoretical and empirical literature – is not necessarily tenable and – if relaxed – substantially changes results. Price setting strategies may not only differ across firms, e.g. incumbents versus entrants, but can even be opposed according to their "endowment" of informed and uninformed consumers. That is, incumbent firms, which are associated with non-shoppers who do not search for alternative tariffs, follow different pricing strategies with respect to consumer information and competition, compared to entrants, which generally attract "shoppers" who engage in search. Hence it may be either beneficial to raise prices in order to extract rents of consumers who exhibit high search costs ("surplus appropriation"), or conversely, to enlarge market shares by lowering prices to attract shoppers with low search costs ("business stealing"). These two forces eventually cause price dispersion as long as there is heterogeneity in consumer search behavior, so that neither everyone is a “shopper” nor that everyone is a “non-shopper” (i.e. $0<\mu<1$).

References

Barron, J.M., Taylor, B.A., Umbeck, J.R. (2004) Number of Sellers, Average Prices, and Price Dispersion. *International Journal of Industrial Organization*, 22, 1041–1066.

Baye, M.R., Morgan, J., Scholten, P. (2006) *Information, Search, and Price Dispersion.* Handbook of Economics and Information Systems. Elsevier, T. Hendershott, ed.

Baye, M.R., Morgan, J., Scholten, P. (2004) Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site. *Journal of Industrial Economics*, 52(4), 463–496.

Blundell, R. and Powell, J. (2004). Endogeneity in semiparametric binary response models. *Review of Economic Studies*, 71, 655–679.

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. and Smith, M. (2001) Consumer Decision-Making at an Internet Shopbot: Brand Still Matters. *Journal of Industrial Economics*, 49(4), 541–558.

Bundesnetzagentur (Federal Network Agency) and Bundeskartellamt (Federal Cartel Office (2014) *Monitoringreport.*

Chandra, A., Tappata, M. (2011) Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets. *RAND Journal of Economics*, 42(2), 681–704.

De los Santos, B., Hortacsu, A., Wildenbeest, M.R. (2012) Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior. *American Economic Review*, 102(6), 2955–2980.

Ericson, K.M.M. (2014) Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange. *American Economic Journal: Economic Policy*, 6(1), 38–64.

Giulietti, M., Waterson, M., Wildenbeest, M. (2014) Estimation of Search Frictions in the British Electricity Market. *Journal of Industrial Economics,* 62, 555–590.

Hong, H., Shum, M. (2006) Using Price Distributions to Estimate Search Costs. *RAND Journal of Economics*, 37(2), 257–275.

Honka, E. (2014) Quantifying Search and Switching Costs in the US Auto Insurance Industry. *RAND Journal of Economics*, 45(4), 847–884.

Hortacsu, A., Madanizadeh, S.A., Puller, S. (2015) Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market. NBER Working Paper No. 20988.

Imbens, G. and Wooldridge, J. (2009) Recent developments in the econometrics of program
evaluation. *Journal of Economic Literature*, 47:5–86.

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 (2014) Search for Differentiated Products: Identification and Estimation. Rand Journal of Economics, 45(3), 553–575.

Lewis, M. (2008). Price Dispersion and Competition with Spatially Differentiated Sellers. *Journal of Industrial Economics*, 56(3), 654–678.

Lind, J.T., Mehlum, H. (2010), With or Without U? The Appropriate Test for a U-Shaped Relationship. *Oxford Bulletin of Economics and Statistics*, 72, 109–118.

Montgomery, A., Hosanagar, K., Krishnan, R. Clay K. (2004) Designing a Better Shopbot, *Management Science*, 50 (2), 189-206.

Morgan, J., Orzen, H., Sefton, M. (2006) An Experimental Study of Price Dispersion. *Games and Economic Behavior*, 54, 134–158.

Varian, H.R. (1980) A Model of Sales. *American Economic Review*, 70, 651–659.

A.T. Kearney (2012) Der Strom- und Gasvertrieb im Wandel, Presentation, Berlin.

https://www.atkearney.de/documents/856314/1214638/BIP_Der_Strom_und_Gasvertrieb_im_Wandel.pdf/ee091e7c-9406-4b23-b5b3-608f936cbecc

Kleibergen, F. and Paap, R. (2006) Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics*, 133,97-126.

Koulayev, S. (2014) Search for Differentiated Products: Identification and Estimation. *RAND Journal of Economics*, 45 (3), 553–575.

Pennersdorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., Yontcheva, B. (2014) Information and Price Dispersion: Evidence from Retail Gasoline. University of Vienna, Working Paper.

Sengupta, A., Wiggins, S.N. (2014) Airline Pricing, Price Dispersion, and Ticket Characteristics On and Off the Internet. *American Economic Journal: Economic Policy*, 6(1), 272–307.

Sorensen, A.T. (2000). Equilibrium Price Dispersion in Retail Markets for Prescription Drugs. *Journal of Political Economy*, 108(4), 833–850.

Stahl, D.O. (1989) Oligopolistic Pricing with Sequential Consumer Search. *American Economic Review*, 79(4), 700–712.

Stigler, G.J. (1961) The Economics of Information. *Journal of Political Economy*, 69(3), 213–225.

Stock, J.H. and Yogo, M. (2005) Testing for Weak Instruments in Linear IV Regression. NBER Technical Working Paper 284.

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.

Wilson, C. M., Waddams Price, C. (2010) Do Consumers Switch to the Best Supplier?. *Oxford Economic Papers*, 62, 647–668.

Yatchew, A. 1(997) An elementary estimator of the partial linear model. *Economic Letters,* 57, 135–143.

Yatchew, A. (1998) Nonparametric regression techniques in economics. *Journal of Economic Literature,* 36, 669–721.

Appendix

Figure 4: Distribution of consumer search intensity before and after dropping the highest 2% of the observations.

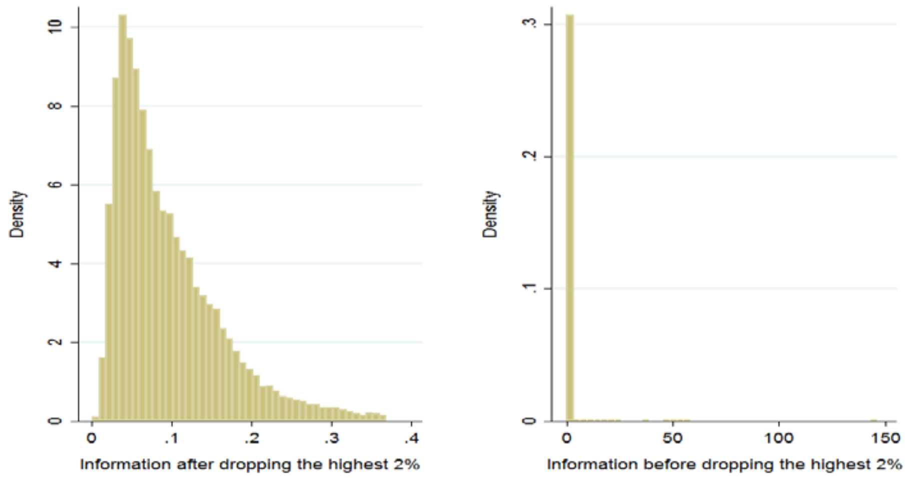


Figure 5: Distribution of number of competitors

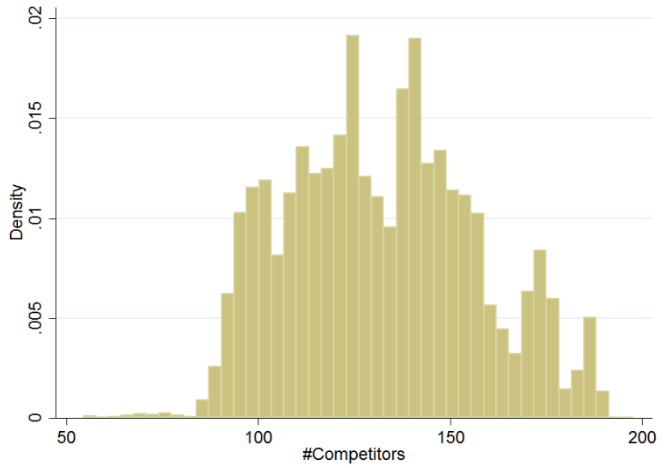


Figure 6: Distribution of price dispersion

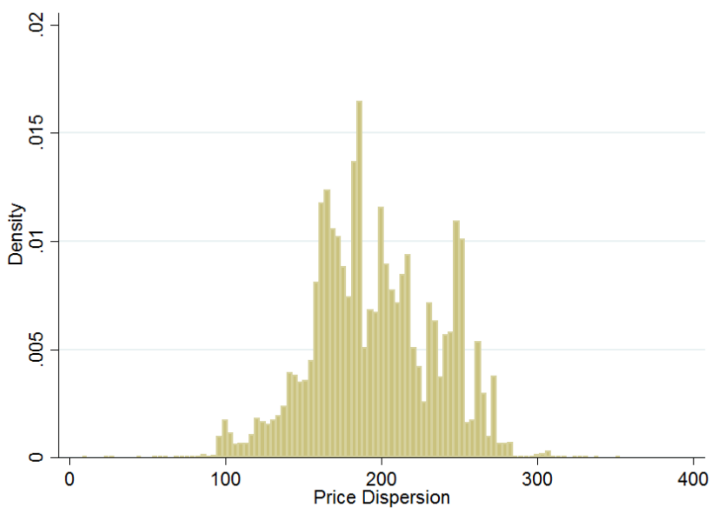


Figure 7: Spatial distribution of price dispersion (2014)

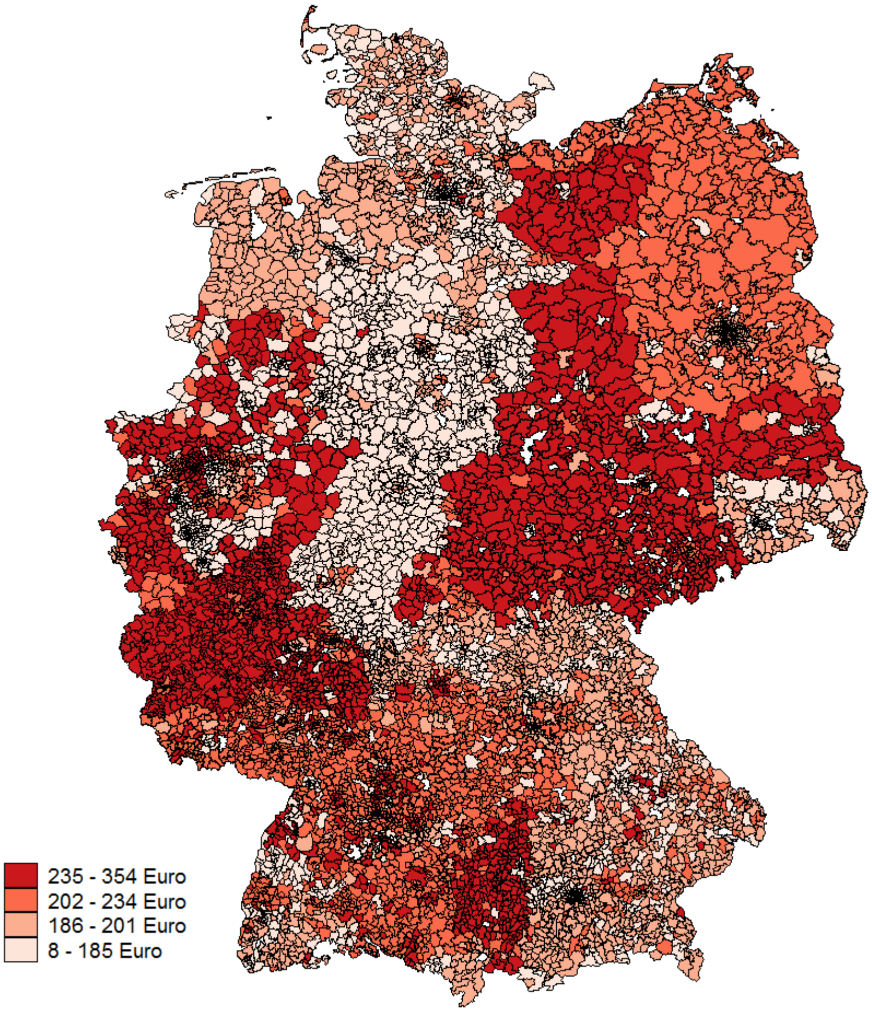


Figure 8: Spatial distribution of search intensity (2014)

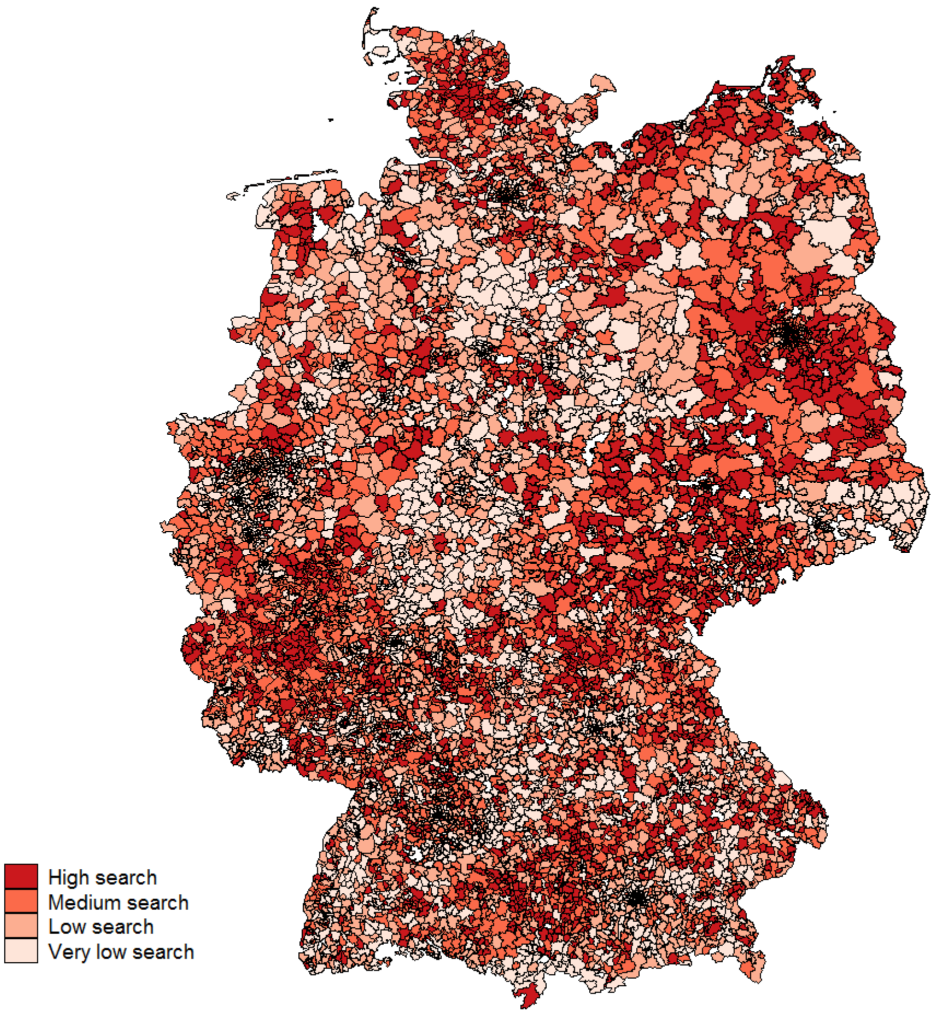


Figure 9: Spatial distribution of electricity retailers (2014)

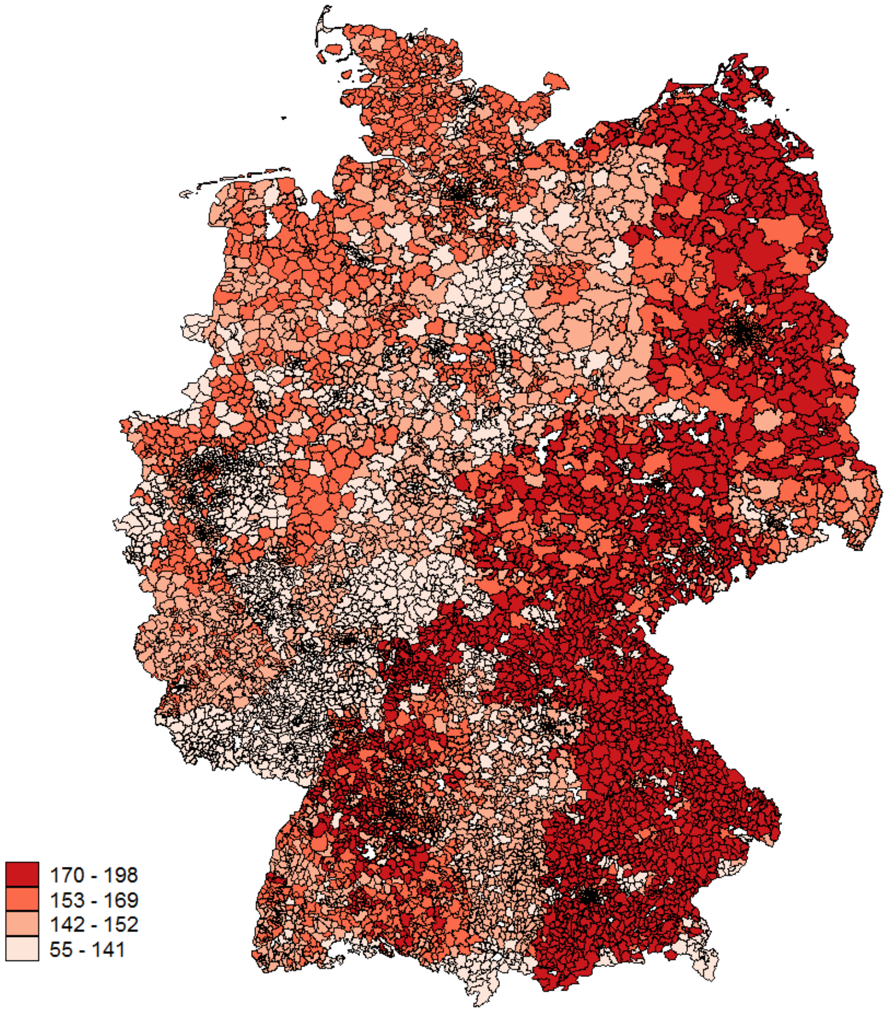


Table 7: First-stage regressions of equations 1 and 2

	(1) Information	(2) Information	(3) #Competitors	(4) #Competitors
New HH	0.006	0.035	-140.121***	-92.721**
	(0.035)	(0.037)	(12.552)	(41.252)
Moved HH	-0.125***	-0.111***	42.874***	47.229*
	(0.033)	(0.036)	(10.654)	(24.376)
Head of HH below age of 40	0.045***	0.038*	66.550***	13.223
	(0.007)	(0.021)	(2.248)	(14.750)
#Households	-0.029***	-0.052*	12.584***	-4.867
	(0.004)	(0.028)	(1.426)	(5.906)
Average HH size	0.048***	0.045***	27.361***	20.660***
	(0.003)	(0.009)	(0.800)	(7.036)
Income < 25k Euro/year	-0.029***	-0.052*	12.584***	-4.867
	(0.004)	(0.028)	(1.426)	(5.906)
Income 25-50k Euro/year	0.021**	0.001	-39.541***	-20.346
	(0.011)	(0.034)	(3.120)	(13.148)
Grid Charge (Variable Part)	0.006***	0.005**	4.406***	-0.893**
	(0.000)	(0.003)	(0.118)	(0.430)
Grid Charge (Fixed Part)	0.000***	0.000	0.086***	0.108
	(0.000)	(0.000)	(0.008)	(0.096)
Concession Fee	0.013***	0.002	0.629	2.433
	(0.001)	(0.005)	(0.452)	(1.791)
Fixes Effects	NO	YES	NO	YES
Year Dummies	YES	YES	YES	YES
First-stage *F* test	219.26	214.59	193.27	70.84
R^2	0.583	0.592	0.671	0.835
#Obs.	30968	30968	30968	30968

Constant term and year dummies not reported. Robust standard errors in parentheses significant for * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The first-stage F test is the Angrist-Pischke F-test for the excluded instruments. *Instruments in italics.*

Table 8: Markup Dispersion Estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	368.179***	59.456***	928.974***	587.085***
	(11.716)	(6.030)	(93.819)	(69.763)
Information^2	-922.534***	-157.194***	-2555.076***	-1633.464***
	(32.908)	(16.177)	(247.805)	(183.957)
#Competitors (N)	1.677***	0.922***	0.580**	2.787***
	(0.075)	(0.058)	(0.242)	(0.304)
#Competitors^2	-0.006***	-0.002***	-0.004***	-0.008***
	(0.000)	(0.000)	(0.001)	(0.001)
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Table 9: Markup Incumbent Estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	374.803***	45.948***	837.476***	245.476***
	(10.824)	(4.721)	(90.883)	(50.768)
Information^2	-913.770***	-96.764***	-2099.244***	-555.553***
	(31.446)	(12.334)	(239.456)	(134.572)
#Competitors (N)	1.450***	0.556***	0.010	1.584***
	(0.067)	(0.038)	(0.229)	(0.217)
#Competitors^2	-0.005***	-0.001***	-0.003***	-0.005***
	(0.000)	(0.000)	(0.001)	(0.001)
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Table 10: Markup Entrant Estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	6.624*	-13.508***	-82.830***	-343.046***
	(3.865)	(3.658)	(31.599)	(40.906)
Information^2	8.764	60.430***	432.093***	1083.128***
	(10.256)	(9.720)	(83.353)	(107.304)
#Competitors (N)	-0.226***	-0.366***	-0.598***	-1.245***
	(0.028)	(0.039)	(0.093)	(0.179)
#Competitors^2	0.001***	0.001***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)	(0.001)
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Table 11: Lerner Dispersion Estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	0.280***	0.048***	0.819***	0.678***
	(0.010)	(0.006)	(0.083)	(0.074)
Information^2	-0.717***	-0.143***	-2.377***	-1.977***
	(0.029)	(0.017)	(0.221)	(0.196)
#Competitors (N)	0.002***	0.001***	0.001***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
#Competitors^2	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Table 12: Lerner Incumbent Estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	0.295***	0.037***	0.720***	0.280***
	(0.009)	(0.004)	(0.077)	(0.047)
Information^2	-0.719***	-0.078***	-1.853***	-0.692***
	(0.026)	(0.011)	(0.203)	(0.126)
#Competitors (N)	0.001***	0.000***	-0.000	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
#Competitors^2	-0.000***	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Table 13: Lerner Entrant Estimates

	(1) OLS	(2) FE	(3) IV	(4) IV FE
Information (μ)	0.014***	-0.012**	-0.093**	-0.399***
	(0.005)	(0.005)	(0.040)	(0.051)
Information^2	-0.002	0.065***	0.509***	1.291***
	(0.013)	(0.012)	(0.106)	(0.134)
#Competitors (N)	-0.000***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
#Competitors^2	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
#Obs.	30,968	30,968	30,968	30,968

Note: Description as in Table 4.

Table 14: Semiparametric estimates. Search is nonparametric. Estimation by Yatchew's (1997) method.

	(1) Price Dispersion	(2) Price Incumbent	(3) Price Entrant
Information (μ)	Non-parametric	Non-parametric	Non-parametric
Control Function for μ	28.3661***	17.2032***	-11.1629***
	(4.9754)	(4.0361)	(2.7352)
#Competitors (N)	1.5565***	1.3132***	-0.2433***
	(0.0574)	(0.0466)	(0.0315)
#Competitors^2	-0.0048***	-0.0043***	0.0005***
	(0.0002)	(0.0002)	(0.0001)
Control Function for N	0.7114***	0.7214***	0.0100
	(0.0327)	(0.0265)	(0.0180)
R^2	0.8312	0.9680	0.9745
#Obs.	29100	29100	29100

Table 15: Semiparametric estimates. Number of competitors is nonparametric. Estimation by Yatchew's (1997) method.

	(1) Price Dispersion	(2) Price Incumbent	(3) Price Entrant
#Competitors (N)	Non-parametric	Non-parametric	Non-parametric
Control Function for N	0.6311***	0.5474***	-0.0837***
	(0.0344)	(0.0286)	(0.0188)
Information (μ)	46.6875***	44.5950***	-2.0925
	(6.7593)	(5.6165)	(3.6873)
Information^2	-108.7234***	-98.0224***	10.7011
	(19.2810)	(16.0211)	(10.5181)
Control Function for μ	16.0103***	10.9717***	-5.0386**
	(4.5947)	(3.8179)	(2.5065)
R^2	0.8140	0.9274	0.9329
#Obs.	29104	29104	29104