Online Appendix

Incumbency Advantages: Price Dispersion, Price Discrimination and Consumer Search at Online Platforms

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

November 14, 2021

Abstract

This document contains additional results, which are not included in the main paper.

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

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

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

[¶]Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), mario.liebensteiner@fau.de

C A piece-wise linear search cost distribution

In this part of the online Appendix, we analyze the impact of changes of the search cost distribution on prices and the fraction of searchers for the piece-wise linear search cost distribution. We focus on parameter values such that $\widehat{s}_1 < \widetilde{s}_1 < \widehat{s}_2 < \widetilde{s}_2$, i.e., the consumer that is indifferent between two online offers is in the first interval of the search cost distribution, while the consumer that is indifferent between searching and not searching is in the second interval of the search cost distribution. With this formulation, an increase in z unambiguously leads the search cost distribution to have a larger fraction of consumers with lower search cost and a smaller fraction of consumers with intermediate search cost.

It is important to note that the number of active searchers $F(P_H^{I^*} - P_L^{I^e})$ is endogenously determined by the equilibrium prices. To dethermine the number of active searchers, we first determine the level of price discrimination $P_H^{I^*} - P_L^{I^*}$. Using (3) it is easy to see that for the case where the search cost distribution is piece-wise linear the equilibrium level of price discrimination equals

$$P_H^{I^*} - P_L^{I^*} = \frac{\widetilde{s}_2 - \widetilde{s}_1 - (z-1)\widetilde{s}_1\widetilde{s}_2}{2\left(\widetilde{s}_2 - z\widetilde{s}_1\right)} \tag{7}$$

and thus that the equilibrium fraction of online searchers equals

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

Applying the piece-wise linear search cost distribution to (1) and (2), it is easy to see that the relation between the equilibrium online prices is given by $P^{E^*} = \frac{1}{2}P_L^{I^*}$ so that

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

which implies that

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

and

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

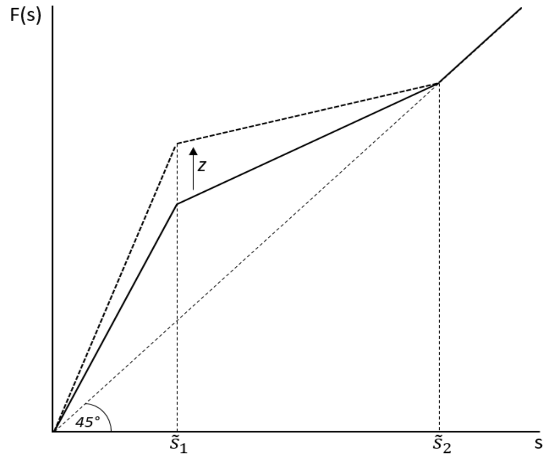
Using Proposition 1 and the fact that for a piece-wise linear distribution $f' = 0$ in the interior of the intervals, online equilibrium prices always change in the same direction and the level of online price dispersion $P_L^{I^*} - P^{E^*}$ positively correlates with both prices. Using the expressions for the different prices, it is easy to see that the condition $\widehat{s}_1 < \widetilde{s}_1 < \widehat{s}_2 < \widetilde{s}_2$ is satisfied if

$$\frac{\widetilde{s}_2 - \widetilde{s}_1 - \widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_1(5\widetilde{s}_2 - 6\widetilde{s}_1)} < z < \frac{\widetilde{s}_2^2 - (\widetilde{s}_2 - \widetilde{s}_1)(1 - \widetilde{s}_2)}{\widetilde{s}_1\widetilde{s}_2}. \tag{8}$$

The proposition below contains the comparative statics properties of our model in terms of price discrimination and dispersion using the piece-wise linear search cost distribution.

Proposition 2 (price levels). If (8) holds, then an increase in the fraction of online searchers $F(P_H^{I^*} - P_L^{I^*})$, initiated by an increase in z, coincides with a decrease in online prices P^{E^*} and $P_L^{I^*}$ if and only if $\widetilde{s}_2 - \widetilde{s}_1 > \widetilde{s}_2\widetilde{s}_1$, while it coincides with an increase in the baseline price $P_H^{I^*}$ if θ is

Figure C1: A piece-wise linear search cost distribution



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

small enough, z is large enough, or $\widetilde{s}_2 - \widetilde{s}_1$ is small enough.

Proof. It is clear that

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

From the expressions determining equilibrium prices, it follows that

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

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

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

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

These results demonstrate that price discrimination maximizes an incumbent firm's profits, as long as it is possible to charge searching and loyal consumers different tariffs. The results can be explained as follows. First, Proposition 1 already stated that price discrimination increases if the inverse hazard condition is satisfied, which is the case for the piece-wise linear distribution. Second, if $\widetilde{s}_2 - \widetilde{s}_1 > \widetilde{s}_2\widetilde{s}_1$, then $\widetilde{s}_1$ is relatively far away from $\widetilde{s}_2$. In this case, if z increases, there are relatively many online consumers that have a relatively low transaction cost to switch away from the incumbent. This gives the incumbent little market power on the online platform, resulting in a lower online price. The result then follows as Proposition 1 already indicated that online price dispersion is positively correlated with the incumbent's online price. Finally, overall price dispersion is closely related to the incumbent's baseline price. That price (and overall price dispersion) is increasing under two broad set of conditions. First, if θ is relatively small, there is fierce competition online and the more consumers search online, the more the incumbent wants to extract surplus from the consumers with high search costs. Second, if z is relatively large, or $\widetilde{s}_2 - \widetilde{s}_1$ is relatively small, there are relatively few consumers that have their decision on whether or not to search be influenced by the base line price, giving the incumbent an incentive to increase its baseline price.

Proposition 3 (price discrimination and dispersion). If (8) holds, then an increase in the fraction of online searchers $F(P_H^{I^*} - P_L^{I^*})$, initiated by an increase in z, coincides with (i) an increase in price discrimination $P_H^{I^*} - P_L^{I^*}$ and (ii) a decrease in online price dispersion $P_L^{I^*} - P^{E^*}$, if and only if $\widetilde{s}_2 - \widetilde{s}_1 > \widetilde{s}_1\widetilde{s}_2$ and (iii) an increase in overall price dispersion $P_H^{I^*} - P^{E^*}$ if θ is small enough, z is large enough, or $\widetilde{s}_2 - \widetilde{s}_1$ is small enough.

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

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

and

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

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

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

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

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

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

D Welfare Analysis of Price Discrimination

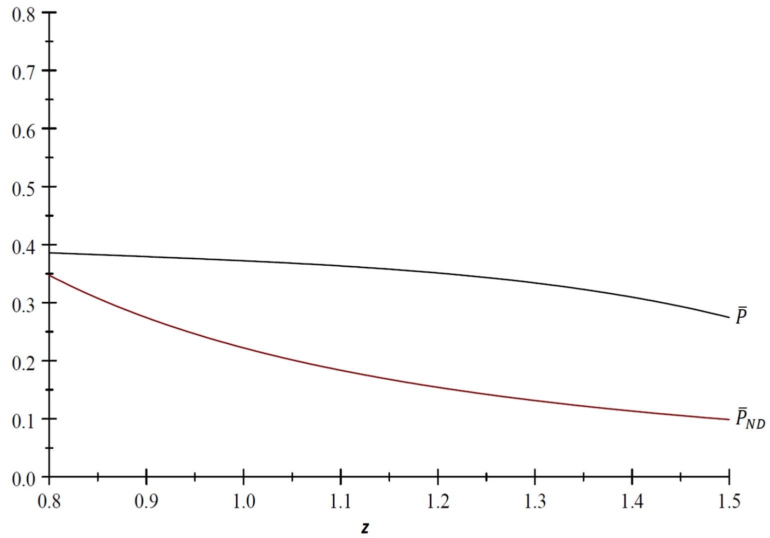
We will now analyze whether consumers are better off without price discrimination. To this end, we simply force $P_H^I = P_L^I$ (and denote this value by P^I) and solve for the equilibrium values, denoting the price choice of the entrant under "no discrimination" by P_{ND}^E (to distinguish it from the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

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

and

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

Figure D1: Average prices with and without price discrimination



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

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

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

and

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

Note that these conditions are very close to (1) and (2). In particular, it is clear that as $F(P^I_H - P^I_L; z) < 1$ in (2) in equilibrium $P^I_L < P^I$ and that because of the strategic complementarity of the price strategies, $P^E < P^E_{ND}$. Thus, searching consumers are better off with price discrimination. Intuitively, without price discrimination the incumbent has a larger share of "loyal" consumers it serves with the price P^I, compared to when it can price discriminate where P^I_L is meant to compete with the entrant's price and the large share of loyal consumers is "addressed" by P^I_H. Thus, with price discrimination, there is simply more online competition to attract searching consumers.

To compare P^I_H and P^I for the general case (and thus to make an overall comparison of the average price consumers pay[1]) is more difficult. Intuitively, though, it would be natural to have that $P^I_H > P^I$, as under price discrimination the incumbent does not need to directly compete with the entrant's price when setting P^I_H. This is easily confirmed for the uniform distribution of search costs, where $z = 1$. In that case $P^{E^*} = \frac{\theta}{6}, P^{I^*}_L = \frac{\theta}{3}$, and $P^{I^*}_H = \frac{1}{2} + \frac{\theta}{3}$, while $P^{E^*}_{ND} = \frac{\theta}{3}, P^{I^*} = \frac{2\theta}{3}$. As $\theta < 1$ it is easy to see that $P^{I^*} < P^{I^*}_H$.

For the case of the uniform distribution, it is also easy to calculate the average price con-

[1]One can also inquire into how the average price depends on the search intensity. The weighted average price is given by $(1 - (F(\widehat{s}_2))P^{I^*}_H + (F(\widehat{s}_2) - F(\widehat{s}_1))P^{I^*}_L + F(\widehat{s}_1)P^{E^*} = P^{I^*}_H - F(\widehat{s}_2)(P^{I^*}_H - P^{I^*}_L) - F(\widehat{s}_1)(P^{I^*}_L - P^{E^*})$.

sumers pay as $\frac{1}{2}\left(\frac{1}{2}+\frac{\theta}{3}\right)+\frac{1}{3}\frac{\theta}{3}+\frac{1}{6}\frac{\theta}{6}=\frac{1}{4}+\frac{11\theta}{36}$, for the case of price discrimination, while without price discrimination, the average price equals $\frac{2}{3}\frac{2\theta}{3}+\frac{1}{3}\frac{\theta}{3}=\frac{5\theta}{9}$. It follows that as $\theta<1$, on average, the effect of the higher baseline price $P_H^{I^*}$dominates and that consumers are worse off under price discrimination. Figure D1 confirms that, for the piece-wise linear distribution that we have considered above and for the same parameter values ($\widetilde{s}_2=3/5$ and $\widetilde{s}_1=1/5$ and $\theta=2/5$), on average consumers are worse off under price discrimination. This average hides, however, that searching consumers are better off, and that loyal consumers are considerably worse off, under price discrimination. In addition, with price discrimination, there will be fewer consumers switching to entrants than if price discrimination were banned. Note that the fraction of switchers is given by $F\left(\frac{P_L^I-P^E}{\theta};z\right)$ for the case of price discrimination and by $F\left(\frac{P^I-P_{ND}^E}{\theta};z\right)$ when price discrimination is banned. We have argued that $P_L^I<P^I$ and $P^E<P_{ND}^E$. As for the piece-wise linear distribution $P_{ND}^E=P^I/2$ and $P^E=P_L^I/2$, it follows that $\frac{P_L^I-P^E}{\theta}<\frac{P^I-P_{ND}^E}{\theta}$ so that fewer consumers switch under price discrimination.

E Sequential price setting game

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

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

and

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

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

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

and

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

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

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

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

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

For general distribution functions, it is not possible to solve these three equations in a meaningful way. Thus, in the rest of this appendix we consider the piece-wise linear distribution, where (as in the main text) we consider $\frac{P^I_L - P^E}{\theta} < \widetilde{s}_1 < P^I_H - P^I_L < \widetilde{s}_2$. As in the main text, the solution to (1) yields $P^E = P^I_L/2$, while in combination with (2) we have

$$\left(\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}\right) P^I_L = \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1} P^I_H + \frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}.$$

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

$$0 = -\frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}(P^I_H - P^I_L)\left(1 - \frac{\frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}{\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}\right) - \frac{z}{2\theta}\frac{\frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}{\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}} P^I_H$$
$$+ \left[\frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1} - \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}(P^I_H - P^I_L) - z\frac{P^I_L - P^E}{\theta}\right] \frac{\frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}{\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}$$
$$+ 1 - \frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1} - \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}(P^I_H - P^I_L),$$

or as

$$\left(\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}\right) P^I_L = \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1} P^I_H + \frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}$$

we have that

$$0 = -2\left(\frac{3z}{2\theta}P^I_L - \frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}\right) - \frac{z}{\theta}\frac{\frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}{\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}} P^I_L$$
$$+ 1 - \frac{(z-1)\widetilde{s}_1\widetilde{s}_2}{\widetilde{s}_2 - \widetilde{s}_1}\left(1 - \frac{\frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}{\frac{3z}{2\theta} + \frac{\widetilde{s}_2 - z\widetilde{s}_1}{\widetilde{s}_2 - \widetilde{s}_1}}\right),$$

which can be simplified to

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

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

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

so that

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

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

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

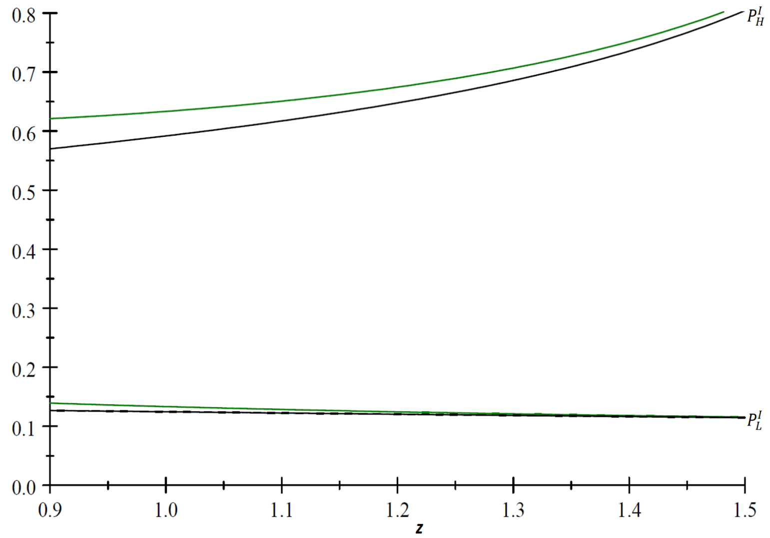
and

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

Figure F1 plots these prices under sequential price setting as a function of z together with the corresponding prices for the simultaneous move game analyzed in the main text. The figure shows that the two different analyses (simultaneous versus sequential choice of offline and online prices) show that equilibrium outcomes are very close to each other. The reason is twofold. First, as indicated above, for given and identical P_H^I, the online market is governed by the same incentives and F.O.C.s. Second, if in the sequential setting the incumbent wants to increase its baseline tariff compared to the simultaneous choice setting, the incumbent not only gains because all prices will increase, but also loses as more consumers will switch to the entrant instead of buying from the online incumbent's price. These opposing forces are such that the net effect is that the baseline price is almost identical in the two cases.

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

Figure F1: Price patterns: simultaneous versus sequential game



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

sion decreases.

F Additional Figures and Tables

Figure F1: Price zones of "Envia Mitteldeutsche Energie GmbH"

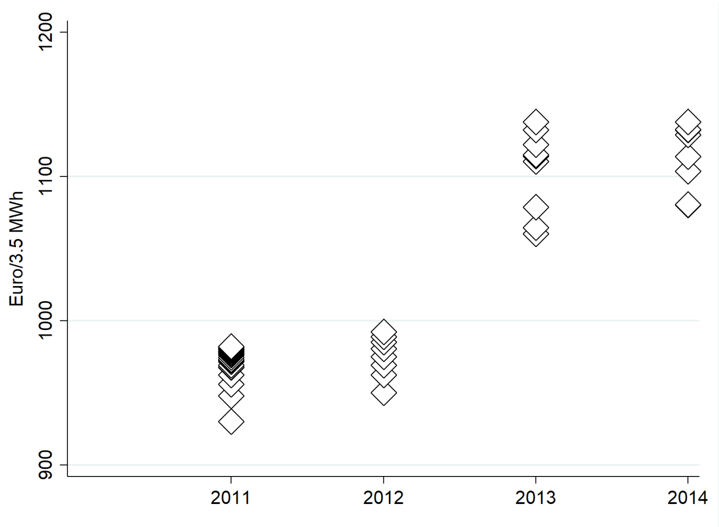
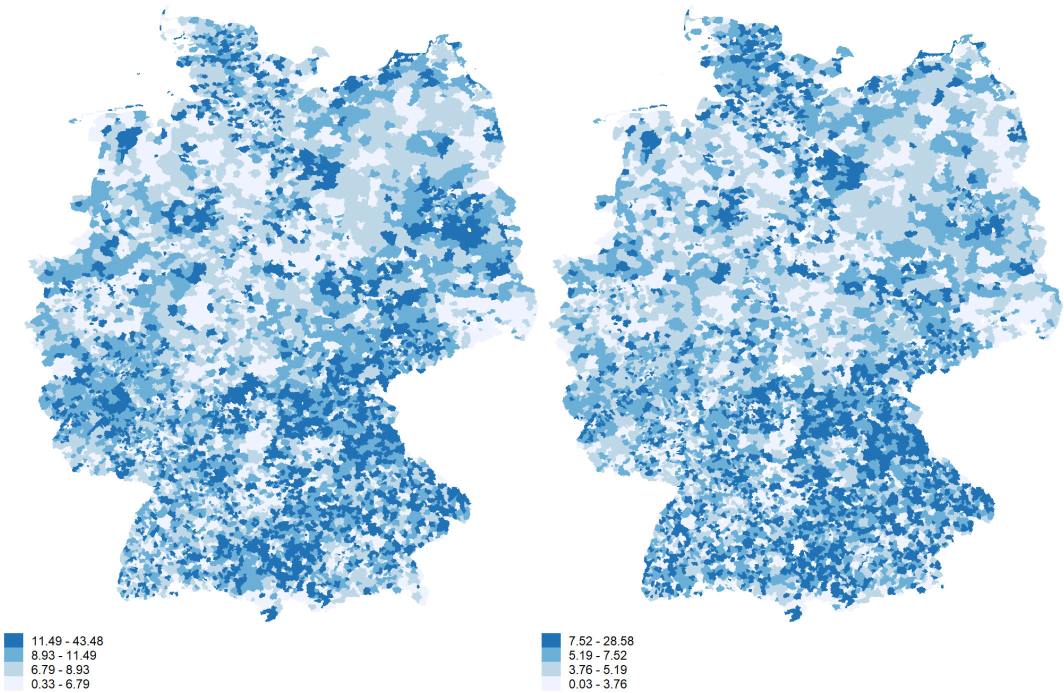
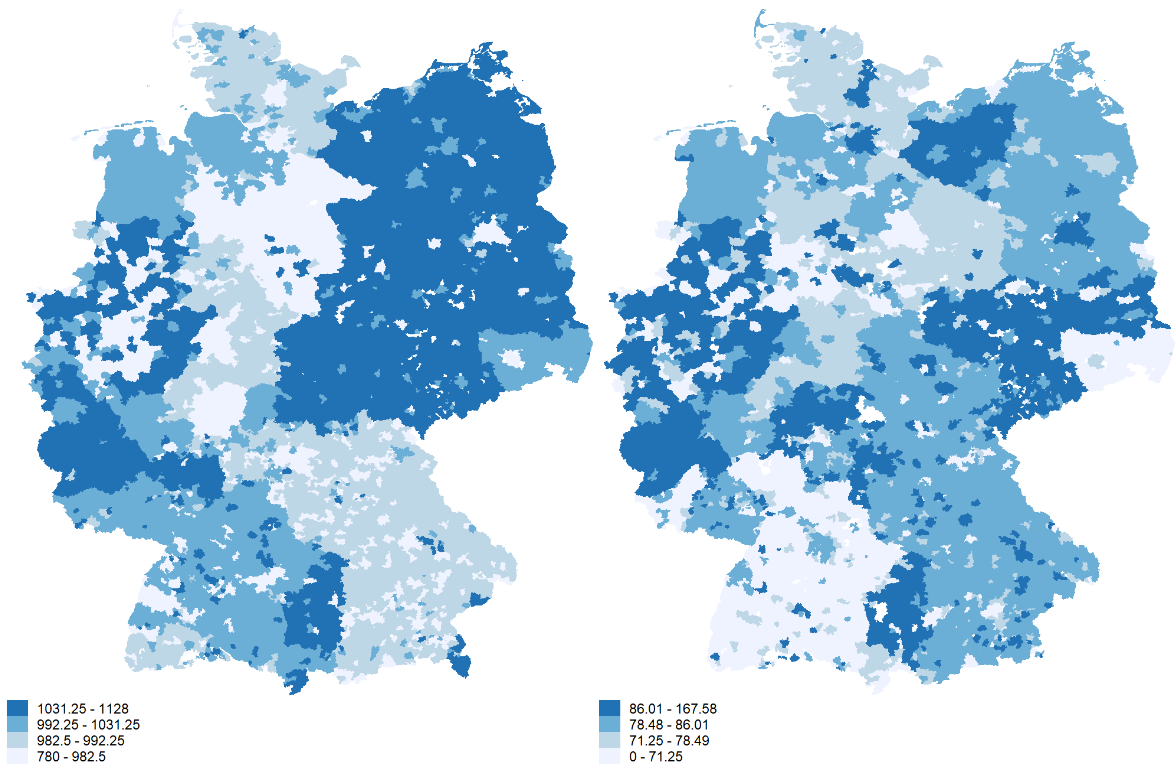


Figure F2: Between and within variation of consumer search intensity



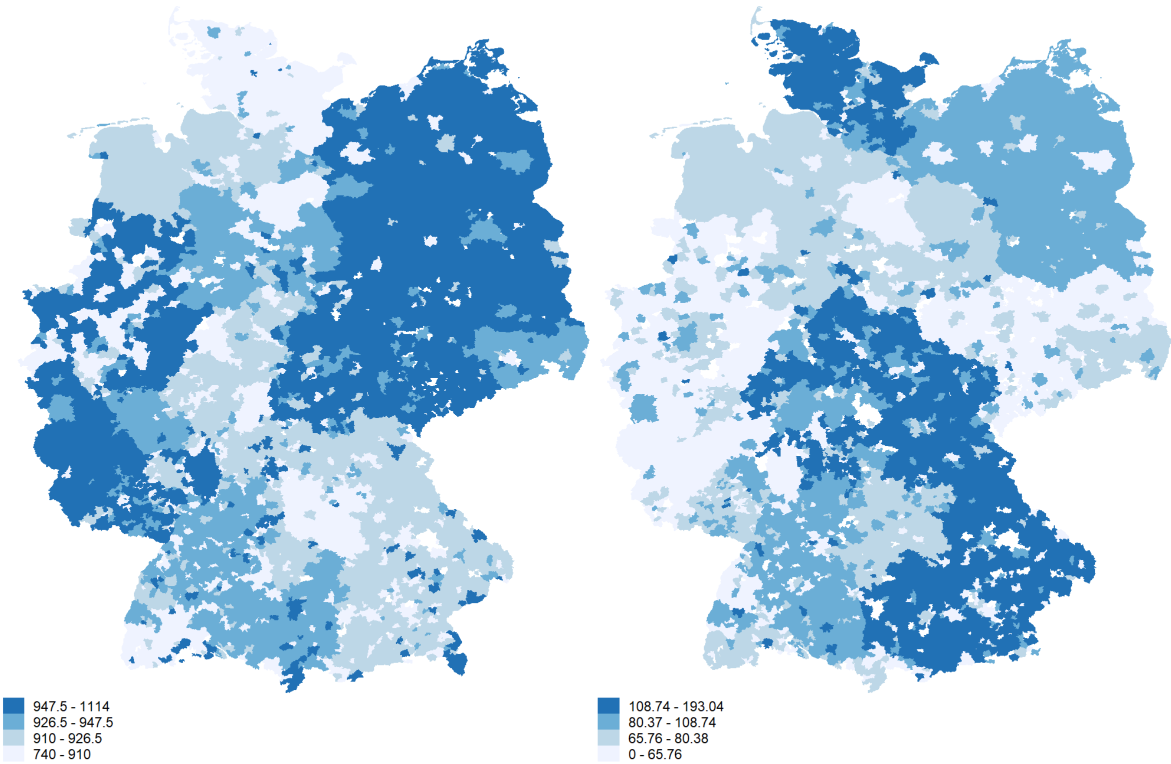
Notes: The left panel presents the between variation in consumer search intensity computed as the average search intensity per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

Figure F3: Between and within variation of incumbents' base prices



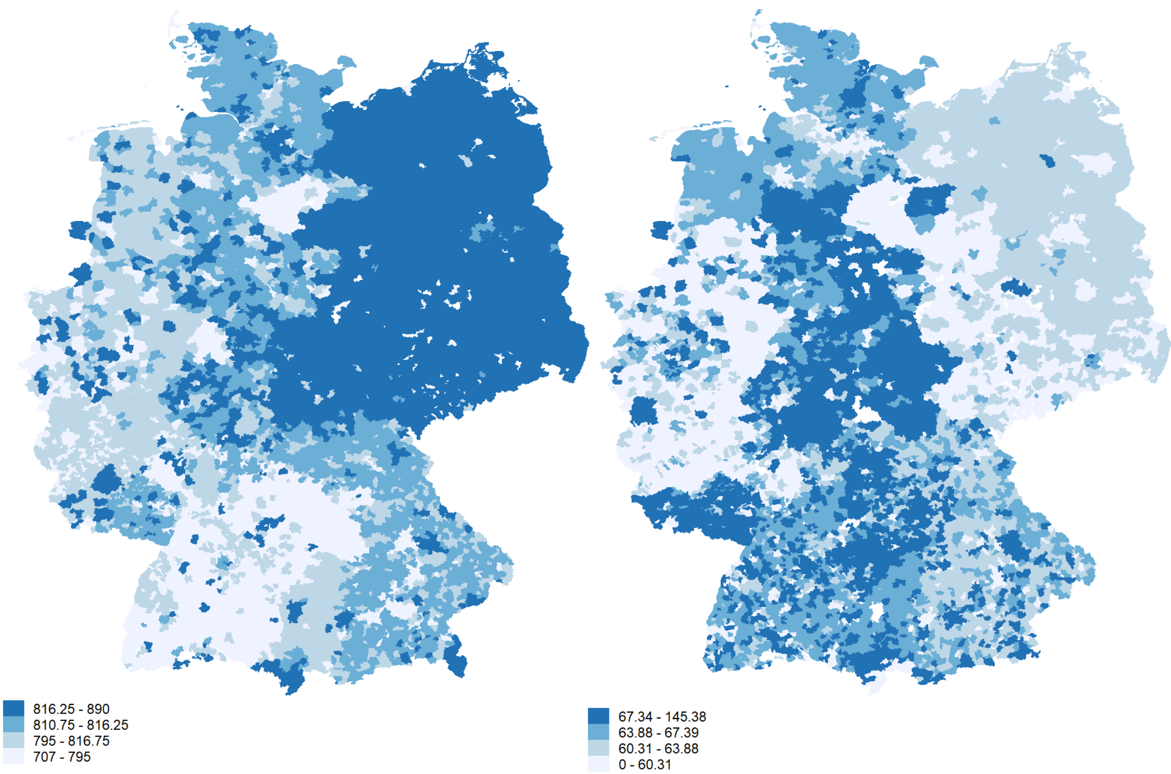
Notes: The left panel presents the between variation in incumbent base prices computed as the average incumbent base price per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

Figure F4: Between and within variation of incumbents' cheapest prices



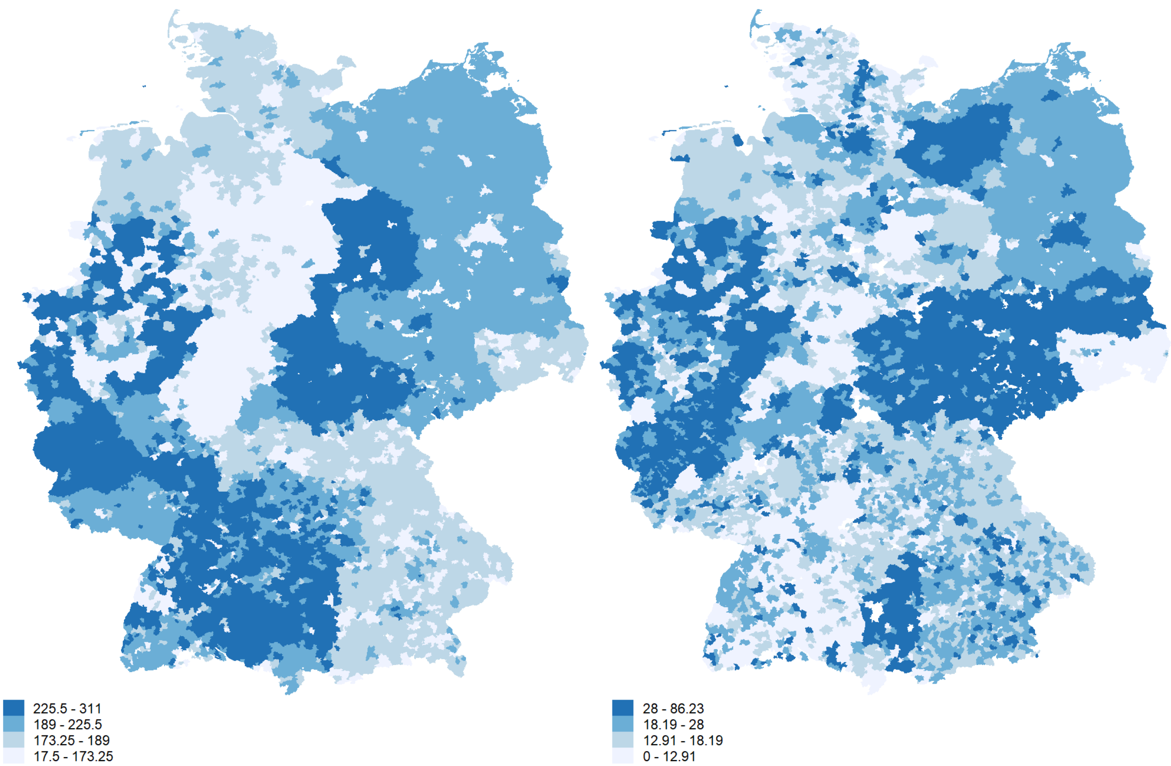
Notes: The left panel presents the between variation in incumbents' cheapest prices computed as the average incumbents' cheapest price per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

Figure F5: Between and within variation of the overall cheapest prices



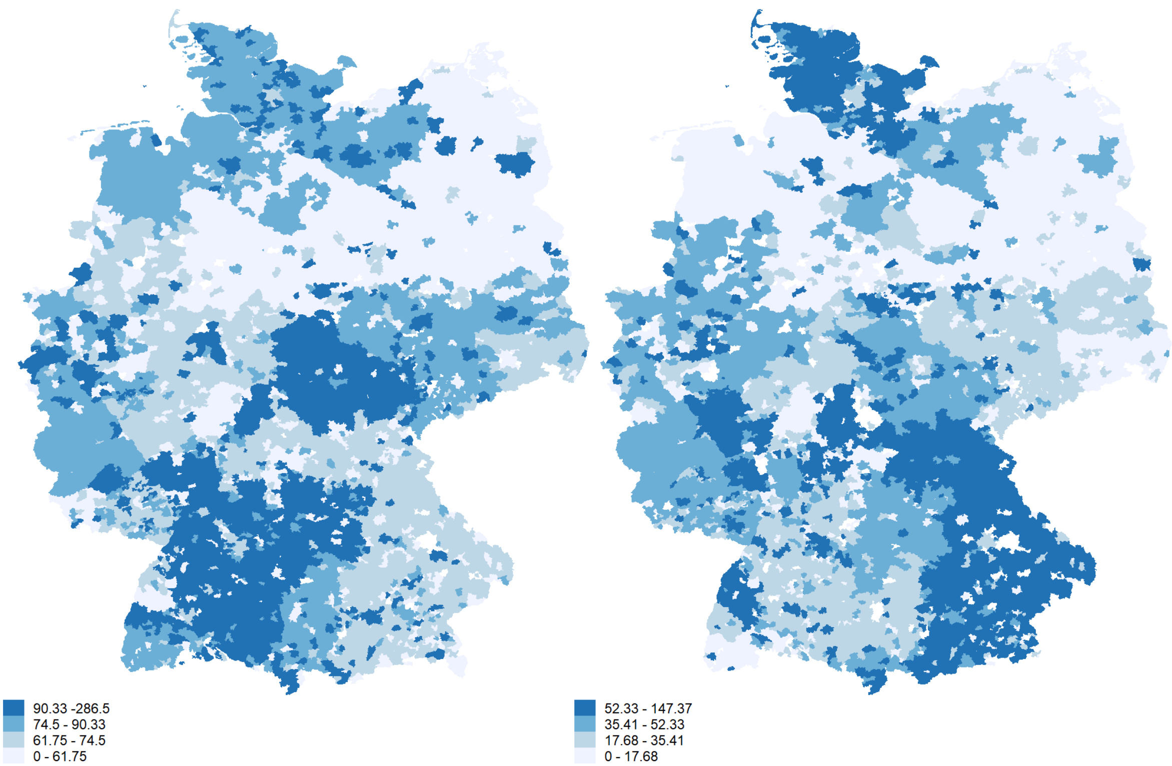
Notes: The left panel presents the between variation in the overall cheapest prices computed as the average cheapest price per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

Figure F6: Between and within variation of price dispersion



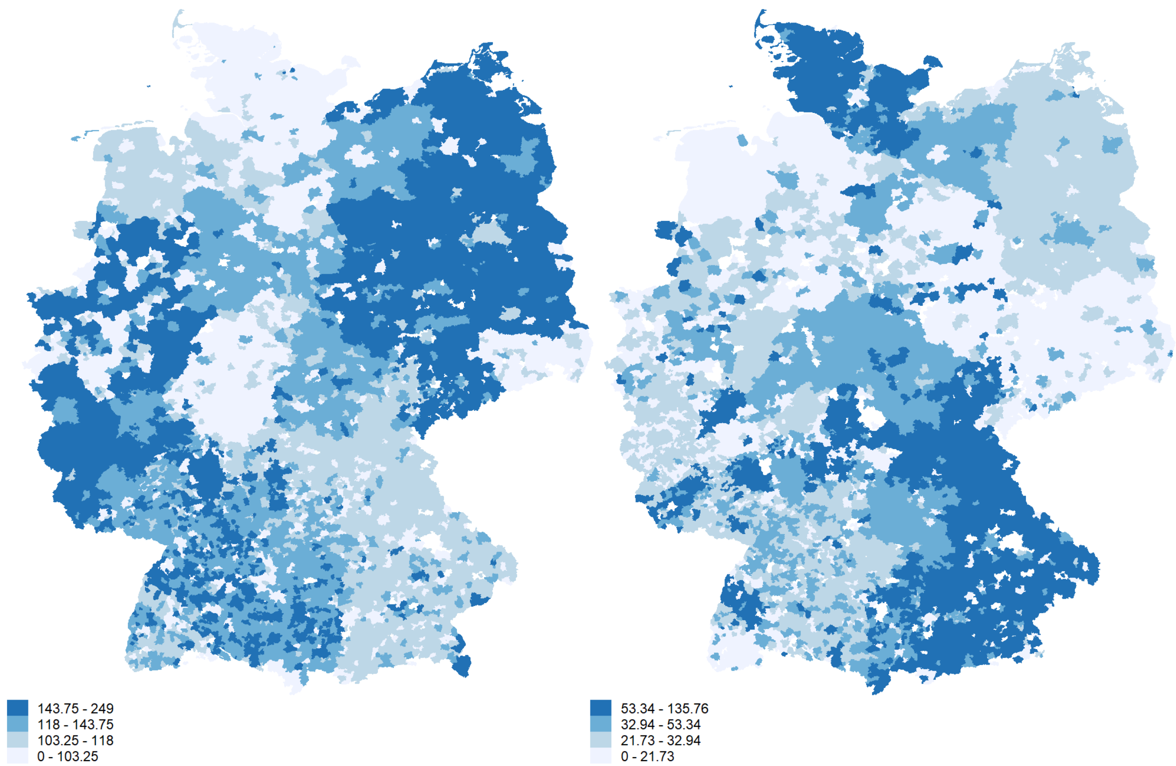
Notes: The left panel presents the between variation in the price dispersion computed as the average price dispersion per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

Figure F7: Between and within variation of price discrimination



Notes: The left panel presents the between variation in the price discrimination computed as the average price discrimination per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

Figure F8: Between and within variation of online price dispersion



Notes: The left panel presents the between variation in the online price dispersion computed as the average online price dispersion per zip code during our observation period. The right panel presents the within variation per zip code computed as the standard deviation per zip code during our observation period.

G Robustness

Estimation without covariates — All results stay robust when we drop all covariates, as shown in Table G1. That is, the instrumental-variables regression counteracts an omitted-variables bias.

Alternative Instruments I — Our two classes of instruments, broadband availability (BBA) and the share of young households (U40) in a zip code, arguably fulfill the exclusion restriction conditional on the included explanatory variables, since one may argue that these variables directly affect search intensity but do not directly affect prices – conditional on our control variables. To further increase the credibility of our identification strategy, we also apply Hausman-type instruments, as inspired by Hausman (1996) (see also Berry and Haile, 2015; Hausman et al., 1994; Nevo, 2000). Thus, we take averages of BBA and U40 in the 50 *surrounding* zip codes as instruments for search intensity in the focal zip code. In addition, we only include those surrounding zip codes when their prices differ from those in the focal zip code. The logic behind the validity of these instruments is that variation in these two sets of instruments proxy for changes in search intensity unrelated to electricity prices in the focal zip code. For example, if their variation is geographically correlated – so that broadband expansion and/or the age distribution are correlated across zip codes – the instruments are correlated with search intensity in the focal zip code.[2] The key assumption on excludability is that the search intensity in the focal zip code is (mean) independent of the error terms of the price equation conditional on the exogenous control variables. This would only fail if the price shocks in the focal zip code are correlated with broadband expansion or the age distribution in *surrounding* zip codes, which we deem highly unlikely.[3] Also, there is sufficient correlation between our Hausman-type instruments and our main instruments: 0.55 for BBA and 0.60 for U40. We find that the results stay robust to these alternative instruments, as shown in Table G2. Figure G1 compares the estimates of the baseline IV with the Hausman-type estimates and shows that there is significant overlap in the confidence intervals.

Alternative Instruments II — It may be that the IV results are largely driven by one of the instruments, either the broadband internet availability or the age structure. In a robustness exercise we employ either only broadband internet or age based instruments in the estimations. The results stay qualitatively robust and are reported in Table G3.

Level-Level Estimation — Our results remain also fully robust if we run level-level (instead of log-log) specifications of the models, as shown in Table G4.

[2]There is ample reason to believe that this is the case, e.g. firms expand their broadband network across many zip codes at once due to economies of scale, scope, and density of broadband expansion.

[3]We call these instruments "Hausman-type" and not "Hausman" instruments, because there is a difference in application which further strengthens our identification strategy. "Hausman instruments" in our context would be, for example, the average search intensity in surrounding ZIP codes. In this case, one may argue that – if there are problems of reverse causality, measurement error or omitted variables – that they also manifest themselves in the geographically related variables. For example, if an electricity retailer sets prices for several ZIP codes simultaneously, and if there is reverse causality, so that consumers change their search behavior as a reaction to a price change, the two error terms in the first stage and the outcome equation may be correlated and "Hausman" instruments may not fulfill the exclusion restriction. Our requirements are much more modest, since it is highly unlikely that a price change in the focal ZIP code change broadband expansion or the age distribution in surrounding ZIP codes.

Non-Linear Effect of Search on Prices — We also allow for a non-linear relationship between search and prices, by adding a μ^2 in Equation 4. We instrument for μ^2 by using the square of the first-stage estimate of μ from Equation 5 as the instrument for μ^2 (see Wooldridge (2010, p. 262) on this approach). The results remain robust and are reported in Table G5 in the Online Appendix.

Alternative Outcome Variable — We also estimate models using *Lerner Indices* as the dependent variables (Table G6). They are computed as the ratio of markups (i.e.the differences between (net) prices and costs) to prices. The results when using Lerner Indices as the dependent variables are as one would expect from the results of the price estimations.

Additionally Instrumenting for the Number of Competitors — The prices and the number of competitors may affect each other. In order to test if potential endogeneity of the number of competitors affects our results, we also instrument for the number of firms in a zip code by using the average number of firms in the surrounding 50 zip codes in the spirit of Hausman (1996) (for further details, see the discussion in Subsection "Alternative Instruments"). There is sufficient variation regarding the number of firms in the zip code itself and the surrounding 50 zip codes: we observe an absolute difference between the former and the latter of 7 and also a standard deviation of 7. The results only change marginally, as shown in Table G7 suggesting robustness of our results to instrumenting for the number of competitors.

Alternative Clustering of Standard Errors — Many incumbents operate only locally and 46% of the incumbents only have a single zip code in their incumbency area. These small incumbents are mostly municipal utilities. However, larger incumbents often have several zip codes in their incumbency area and charge locally differing baseline tariffs. The different *price zones* are not necessarily at the zip code level (see Section 2 for more details). Hence, as a robustness check, we cluster standard errors in two alternative ways. In the first version, we allow the residuals to correlate within a prices zone and cluster standard errors at that level, instead of the zip code level (see Table G8). In a second version, the clustering is at the incumbency area level (see Table G9). In both cases the results remain fully robust.

Table G1: IV estimates of the impact of consumer search on prices (log-log) – estimations without covariates

	ln Incumbent Base (P^I_H) (1)	ln Incumbent Cheapest (P^I_L) (2)	ln Overall Cheapest (P^E) (3)
ln Search (μ)	0.0431*** (0.0087)	-0.1151*** (0.0261)	-0.0296*** (0.0069)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
First stage effective F stat.	26.37	26.37	26.37
Obs.	24,175	24,175	24,175

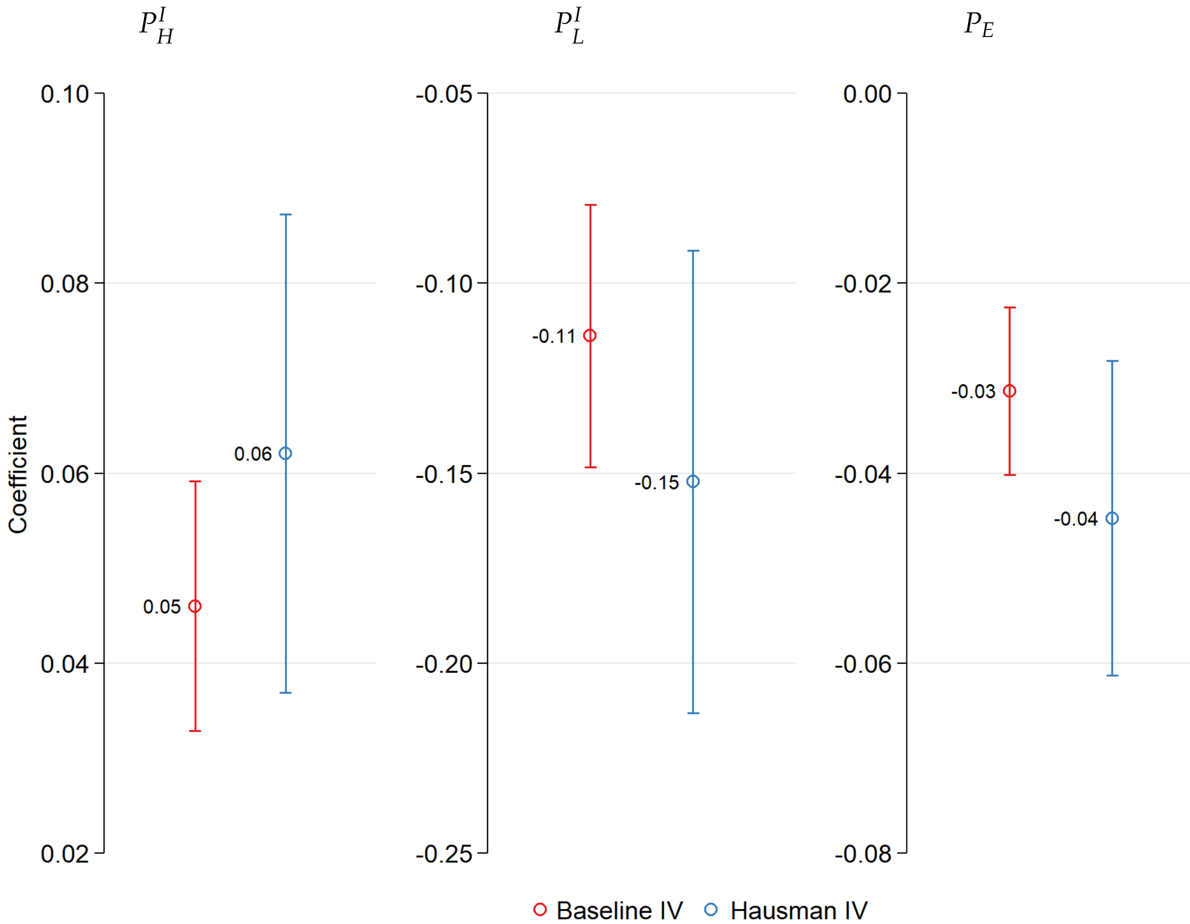
Notes: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBA*. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table G2: IV estimates of the impact of consumer search on prices (log-log) – Hausman-type instruments for search

	ln Incumbent Base (P_H^I) (1)	ln Incumbent Cheapest (P_L^I) (2)	ln Overall Cheapest (P^E) (3)
ln Search (μ)	0.0621***	-0.1522***	-0.0447***
	(0.0153)	(0.0370)	(0.0101)
No. competitors	0.0003***	0.0049***	0.0003***
	(0.0001)	(0.0002)	(0.0001)
No. households	0.0052***	-0.0065	-0.0032***
	(0.0019)	(0.0044)	(0.0012)
ln Costs	0.2054***	0.4770***	0.5221***
	(0.0142)	(0.0316)	(0.0103)
Average HH size	0.0390***	0.0124	-0.0091*
	(0.0082)	(0.0203)	(0.0054)
ln Purchase Power	0.0067	-0.0459***	-0.0116***
	(0.0075)	(0.0144)	(0.0034)
Income <25k €/a	-0.0002	-0.0015***	-0.0002***
	(0.0001)	(0.0003)	(0.0001)
Income 25-50k €/year	-0.0004**	-0.0008**	-0.0001
	(0.0002)	(0.0004)	(0.0001)
Lower class social status	0.0001***	-0.0000	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
Middle class social status	0.0001***	0.0001**	0.0000
	(0.0000)	(0.0000)	(0.0000)
High Urbanization	0.0216***	-0.0672***	-0.0162***
	(0.0057)	(0.0135)	(0.0039)
Unemployed	0.0204***	0.0523***	0.0068**
	(0.0053)	(0.0125)	(0.0034)
Population density	0.0004**	-0.0002	-0.0000
	(0.0001)	(0.0004)	(0.0001)
Social insurance	0.0012***	-0.0002	-0.0002
	(0.0002)	(0.0005)	(0.0002)
ln Lagged switching rate	0.0044***	0.0285***	0.0014**
	(0.0009)	(0.0024)	(0.0006)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
First stage effective F stat.	12.37	12.37	12.37
Obs.	24,175	24,175	24,175

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

Figure G1: Comparison of estimates from baseline IV and Hausman-type IV estimations



Notes: the red circles represent point estimates from our baseline IV regressions while the blue circles those from the Hausman-type IV regressions. Red and blue vertical lines represent the corresponding 90% CIs.

Table G3: Estimates of the impact of consumer search on prices (log-log) – instruments either only based on broadband availability or on age characteristics

	ln Incumbent base P_H^I		ln Incumbent cheapest P_L^I		ln Overall cheapest P^E	
	Age (1)	Internet (2)	Age (3)	Internet (4)	Age (5)	Internet (6)
ln Search (μ)	0.0478***	0.0598***	-0.1044***	-0.1533***	-0.0333***	-0.0149**
	(0.0086)	(0.0121)	(0.0227)	(0.0340)	(0.0062)	(0.0066)
No. competitors	0.0004***	0.0004***	0.0046***	0.0050***	0.0002***	0.0001***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0000)	(0.0000)
No. households	0.0037***	0.0046***	-0.0028	-0.0072*	-0.0022**	-0.0014
	(0.0014)	(0.0016)	(0.0035)	(0.0043)	(0.0010)	(0.0009)
ln Costs	0.2106***	0.2235***	0.4560***	0.5079***	0.5242***	0.5317***
	(0.0114)	(0.0130)	(0.0264)	(0.0324)	(0.0093)	(0.0085)
Average HH size	0.0407***	0.0382***	0.0156	0.0194	-0.0064	-0.0127***
	(0.0070)	(0.0083)	(0.0183)	(0.0225)	(0.0053)	(0.0045)
ln Purchase Power	0.0047	0.0103	-0.0356***	-0.0450***	-0.0099***	-0.0055**
	(0.0064)	(0.0079)	(0.0128)	(0.0165)	(0.0033)	(0.0028)
Income <25k €/a	-0.0001	-0.0002	-0.0016***	-0.0014***	-0.0002***	-0.0003***
	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0001)	(0.0001)
Income 25-50k €/year	-0.0003**	-0.0004**	-0.0009**	-0.0006	-0.0001	-0.0002*
	(0.0002)	(0.0002)	(0.0003)	(0.0004)	(0.0001)	(0.0001)
Lower class social status	0.0001***	0.0001***	-0.0000	-0.0000	-0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Middle class social status	0.0001***	0.0001***	0.0001***	0.0001**	0.0000*	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
High Urbanization	0.0185***	0.0221***	-0.0559***	-0.0663***	-0.0131***	-0.0079**
	(0.0047)	(0.0054)	(0.0111)	(0.0135)	(0.0035)	(0.0031)
Unemployed	0.0181***	0.0205***	0.0633***	0.0476***	0.0094***	0.0124***
	(0.0044)	(0.0051)	(0.0101)	(0.0127)	(0.0030)	(0.0028)
Population density	0.0003***	0.0003**	-0.0001	-0.0002	0.0000	0.0001
	(0.0001)	(0.0001)	(0.0004)	(0.0004)	(0.0001)	(0.0001)
Social insurance	0.0011***	0.0011***	0.0001	-0.0002	0.0000	-0.0000
	(0.0002)	(0.0002)	(0.0005)	(0.0006)	(0.0001)	(0.0001)
ln Lagged switching rate	0.0049***	0.0046***	0.0275***	0.0297***	0.0010*	0.0004
	(0.0007)	(0.0009)	(0.0021)	(0.0026)	(0.0006)	(0.0005)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	Yes	Yes
First stage effective F stat.	34.88	10.27	34.88	10.27	34.88	10.27
Obs.	24,175	23,910	24,175	23,910	24,175	23,910

Notes: Standard errors clustered at the zip code level in parentheses. Instruments for μ are either the local coverage of different broadband internet speeds (16mb/s, 6mb/s as well as mobile internet availaibility of 6mb/s and 1mb/s) or the share of young households (household head younger than 40) and middle aged households (household head between 40 and 60 years old). $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table G4: IV estimates of the impact of consumer search on prices (level-level)

	Incumbent Base (P_H^I)	Incumbent Cheapest (P_L^I)	Overall Cheapest (P^E)
	(1)	(2)	(3)
Search (μ)	2.7979***	-9.4572***	-1.7719***
	(0.7256)	(1.9140)	(0.4170)
No. competitors	0.4122***	4.3446***	0.2551***
	(0.0748)	(0.1907)	(0.0423)
No. households	2.4428	-12.4749***	-2.9812***
	(1.6704)	(4.3756)	(0.9868)
ln Costs	187.3952***	551.0023***	433.6734***
	(16.5061)	(41.9002)	(10.5287)
Average HH size	36.9131***	-43.1948**	-16.9805***
	(6.6625)	(17.7897)	(4.2414)
ln Purchase Power	-1.7343	-15.7611	-4.2896*
	(5.4960)	(13.5147)	(2.5991)
Income <25k €/a	0.2528**	-2.3856***	-0.3234***
	(0.1162)	(0.3046)	(0.0726)
Income 25-50k €/year	0.0833	-1.6828***	-0.2409***
	(0.1342)	(0.3552)	(0.0867)
Lower class social status	0.1205***	-0.0629	-0.0518***
	(0.0149)	(0.0397)	(0.0089)
Middle class social status	0.0829***	0.1117***	0.0155**
	(0.0123)	(0.0347)	(0.0077)
High Urbanization	23.4521***	-83.8896***	-14.2943***
	(6.0162)	(16.2818)	(3.8276)
Unemployed	-1.4004	90.4368***	11.8443***
	(4.1647)	(9.5017)	(2.3258)
Population density	0.3171***	-0.0360	0.0339
	(0.1060)	(0.3881)	(0.0786)
Social insurance	0.8248***	-0.3585	-0.0963
	(0.1891)	(0.4753)	(0.1119)
ln Lagged switching rate	3.8982***	28.8753***	1.5850***
	(0.9072)	(2.7828)	(0.5714)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
First stage effective F stat.	12.04	12.04	12.04
Obs.	24,175	24,175	24,175

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

Table G5: IV estimates of a non-linear impact of consumer search on prices (log-log)

	ln Incumbent Base (P_H^I) (1)	ln Incumbent Cheapest (P_L^I) (2)	ln Overall Cheapest (P^E) (3)
ln Search (μ)	0.0389***	-0.1534***	-0.0373***
	(0.0085)	(0.0224)	(0.0056)
$Search^2$	0.0000	0.0134***	0.0032***
	(0.0004)	(0.0010)	(0.0003)
No. competitors	0.0004***	0.0044***	0.0001***
	(0.0001)	(0.0002)	(0.0000)
No. households	0.0030**	0.0016	-0.0007
	(0.0013)	(0.0032)	(0.0008)
ln Costs	0.2172***	0.4059***	0.5030***
	(0.0105)	(0.0242)	(0.0085)
Average HH size	0.0418***	0.0272	-0.0075
	(0.0066)	(0.0181)	(0.0048)
ln Purchase Power	0.0051	-0.0447***	-0.0099***
	(0.0060)	(0.0118)	(0.0028)
Income <25k €/a	-0.0001	-0.0011***	-0.0001*
	(0.0001)	(0.0003)	(0.0001)
Income 25-50k €/year	-0.0003**	-0.0005	-0.0001
	(0.0001)	(0.0003)	(0.0001)
Lower class social status	0.0001***	-0.0000	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
Middle class social status	0.0001***	0.0001***	0.0000*
	(0.0000)	(0.0000)	(0.0000)
High Urbanization	0.0163***	-0.0444***	-0.0082***
	(0.0044)	(0.0100)	(0.0030)
Unemployed	0.0155***	0.0472***	0.0067**
	(0.0044)	(0.0100)	(0.0028)
Population density	0.0003**	-0.0001	0.0000
	(0.0001)	(0.0003)	(0.0001)
Social insurance	0.0010***	0.0002	0.0000
	(0.0002)	(0.0005)	(0.0001)
ln Lagged switching rate	0.0052***	0.0263***	0.0004
	(0.0007)	(0.0019)	(0.0005)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
Kleibergen-Paap F-stat.	28.29	28.29	28.29
Durbin-Wu-Hausman test	0.00	0.00	0.00
Obs.	24,175	24,175	24,175

Notes: Standard errors clustered at the zip code level in parentheses. Estimation is by GMM. Instrumented for μ by *U40* and *BBA* and for μ^2 by $\hat{\mu^2}$ – the square of the first stage predictions of μ. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table G6: IV estimates of the impact of consumer search on Lerner Indices (LI)

	LI Incumbent Base (1)	LI Incumbent Cheapest (2)	LI Overall Cheapest (3)
ln Search (μ)	0.0359***	-0.0925***	-0.0278***
	(0.0058)	(0.0164)	(0.0048)
No. competitors	0.0003***	0.0034***	0.0002***
	(0.0000)	(0.0001)	(0.0000)
No. households	0.0032***	-0.0033	-0.0023***
	(0.0010)	(0.0026)	(0.0008)
ln Costs	-0.5378***	-0.4016***	-0.3991***
	(0.0080)	(0.0209)	(0.0081)
Average HH size	0.0296***	0.0114	-0.0127***
	(0.0050)	(0.0143)	(0.0045)
ln Purchase Power	0.0051	-0.0290***	-0.0086***
	(0.0046)	(0.0097)	(0.0028)
Income <25k €/a	-0.0001	-0.0011***	-0.0001**
	(0.0001)	(0.0002)	(0.0001)
Income 25-50k €/year	-0.0002**	-0.0006**	-0.0001
	(0.0001)	(0.0003)	(0.0001)
Lower class social status	0.0001***	-0.0000	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
Middle class social status	0.0001***	0.0001***	0.0000
	(0.0000)	(0.0000)	(0.0000)
High Urbanization	0.0137***	-0.0476***	-0.0113***
	(0.0034)	(0.0088)	(0.0030)
Unemployed	0.0102***	0.0411***	0.0038
	(0.0032)	(0.0076)	(0.0026)
Population density	0.0002**	-0.0001	0.0000
	(0.0001)	(0.0003)	(0.0001)
Social insurance	0.0007***	-0.0000	-0.0001
	(0.0001)	(0.0004)	(0.0001)
ln Lagged switching rate	0.0037***	0.0230***	0.0009*
	(0.0005)	(0.0016)	(0.0005)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
First stage effective F stat.	33.96	33.96	33.96
Obs.	24,175	24,175	24,175

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

Table G7: IV estimates of the impact of consumer search on prices (log-log) – additionally instrumented for #Competitors

	ln Incumbent Base (P_H^I) (1)	ln Incumbent Cheapest (P_L^I) (2)	ln Overall Cheapest (P^E) (3)
ln Search (μ)	0.0414***	-0.1265***	-0.0309***
	(0.0079)	(0.0224)	(0.0054)
No. competitors	0.0010***	0.0065***	0.0001**
	(0.0001)	(0.0002)	(0.0001)
No. households	0.0040***	-0.0021	-0.0022**
	(0.0013)	(0.0036)	(0.0009)
ln Costs	0.2319***	0.5129***	0.5146***
	(0.0115)	(0.0291)	(0.0091)
Average HH size	0.0354***	-0.0014	-0.0101**
	(0.0068)	(0.0194)	(0.0050)
ln Purchase Power	-0.0004	-0.0586***	-0.0085***
	(0.0065)	(0.0138)	(0.0032)
Income <25k €/a	-0.0000	-0.0012***	-0.0002***
	(0.0001)	(0.0003)	(0.0001)
Income 25-50k €/year	-0.0003*	-0.0006*	-0.0001
	(0.0001)	(0.0004)	(0.0001)
Lower class social status	0.0001***	-0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
Middle class social status	0.0001***	0.0001**	0.0000**
	(0.0000)	(0.0000)	(0.0000)
High Urbanization	0.0190***	-0.0545***	-0.0125***
	(0.0045)	(0.0114)	(0.0033)
Unemployed	0.0130***	0.0479***	0.0088***
	(0.0043)	(0.0107)	(0.0029)
Population density	0.0003**	-0.0001	0.0000
	(0.0001)	(0.0004)	(0.0001)
Social insurance	0.0010***	-0.0001	-0.0000
	(0.0002)	(0.0005)	(0.0001)
ln Lagged switching rate	0.0058***	0.0302***	0.0009
	(0.0007)	(0.0022)	(0.0005)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
Kleibergen-Paap stat.	24.96	24.96	24.96
Obs.	24,175	24,175	24,175

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

Table G8: IV estimates of the impact of consumer search on prices (log-log) – alternative clustering of standard errors I

	ln Incumbent Base (P^I_H) (1)	ln Incumbent Cheapest (P^I_L) (2)	ln Overall Cheapest (P^E) (3)
ln Search (μ)	0.0486***	-0.1130***	-0.0315***
	(0.0119)	(0.0271)	(0.0082)
No. competitors	0.0004***	0.0046***	0.0002**
	(0.0001)	(0.0004)	(0.0001)
No. households	0.0037	-0.0033	-0.0022
	(0.0023)	(0.0042)	(0.0015)
ln Costs	0.2099***	0.4559***	0.5168***
	(0.0192)	(0.0450)	(0.0168)
Average HH size	0.0404***	0.0128	-0.0108
	(0.0100)	(0.0236)	(0.0069)
ln Purchase Power	0.0075	-0.0391***	-0.0092**
	(0.0076)	(0.0149)	(0.0043)
Income <25k €/a	-0.0001	-0.0015***	-0.0002**
	(0.0001)	(0.0003)	(0.0001)
Income 25-50k €/year	-0.0004**	-0.0008**	-0.0001
	(0.0002)	(0.0003)	(0.0001)
Lower class social status	0.0001***	-0.0000	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
Middle class social status	0.0001***	0.0001***	0.0000
	(0.0000)	(0.0000)	(0.0000)
High Urbanization	0.0186***	-0.0574***	-0.0124***
	(0.0049)	(0.0125)	(0.0036)
Unemployed	0.0168***	0.0582***	0.0083*
	(0.0056)	(0.0136)	(0.0044)
Population density	0.0003**	-0.0001	0.0000
	(0.0001)	(0.0004)	(0.0001)
Social insurance	0.0010***	0.0000	-0.0001
	(0.0002)	(0.0006)	(0.0002)
ln Lagged switching rate	0.0049***	0.0278***	0.0009
	(0.0009)	(0.0025)	(0.0008)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
First stage effective F stat.	42.71	42.71	42.71
Obs.	24,175	24,175	24,175

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

Table G9: IV estimates of the impact of consumer search on prices (log-log) – alternative clustering of standard errors II

	ln Incumbent Base (P_H^I) (1)	ln Incumbent Cheapest (P_L^I) (2)	ln Overall Cheapest (P^E) (3)
ln Search (μ)	0.0515*	-0.0930**	-0.0322*
	(0.0276)	(0.0396)	(0.0175)
No. competitors	0.0004	0.0036***	0.0002
	(0.0003)	(0.0010)	(0.0003)
No. households	0.0045	-0.0006	-0.0024
	(0.0033)	(0.0084)	(0.0024)
ln Costs	0.2020***	0.4984***	0.5145***
	(0.0663)	(0.1256)	(0.0575)
Average HH size	0.0424	-0.0035	-0.0101
	(0.0254)	(0.1321)	(0.0197)
ln Purchase Power	0.0089	-0.0357	-0.0095
	(0.0099)	(0.0303)	(0.0076)
Income <25k €/a	-0.0001	-0.0012**	-0.0002
	(0.0002)	(0.0006)	(0.0002)
Income 25-50k €/year	-0.0004	-0.0007	-0.0001
	(0.0002)	(0.0005)	(0.0002)
Lower class social status	0.0001*	-0.0000	-0.0001***
	(0.0000)	(0.0001)	(0.0000)
Middle class social status	0.0001***	0.0001	0.0000
	(0.0000)	(0.0001)	(0.0000)
High Urbanization	0.0195**	-0.0503***	-0.0127**
	(0.0080)	(0.0126)	(0.0053)
Unemployed	0.0179**	0.0449	0.0085
	(0.0088)	(0.0337)	(0.0087)
Population density	0.0003**	0.0001	0.0000
	(0.0001)	(0.0007)	(0.0001)
Social insurance	0.0011***	-0.0001	-0.0000
	(0.0004)	(0.0016)	(0.0003)
ln Lagged switching rate	0.0056***	0.0262***	0.0009
	(0.0017)	(0.0077)	(0.0016)
Year FE	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes
First stage effective F stat.	10.97	10.97	10.97
Obs.	24,175	24,175	24,175

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