DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

Market Structure and Market Performance in E-Commerce

by

Franz HACKL
Michael E. KUMMER
Rudolf WINTER-EBMER*)
Christine ZULEHNER

Working Paper No. 1404
January 2014

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

rudolf.winterebmer@jku.at
phone: +43 732 2468 8236

Market Structure and Market Performance in E-Commerce *

Franz Hackl
University of Linz

Michael E. Kummer
ZEW and CDSE, Mannheim

Rudolf Winter-Ebmer
University of Linz & IHS, Vienna

Christine Zulehner
Goethe University Frankfurt & WIFO, Vienna

January 31, 2014

Abstract

We analyze the interaction between market structure and market performance and how it varies over the product cycle. To account for the potential endogeneity in this relation, we use an instrumental variable approach. We combine data from the largest Austrian online market for price comparisons with retail data on wholesale prices provided by a major hardware producer for consumer electronics. Our results show that instrumenting is important for estimating the empirical effect of competition on the markup of the price leader. One more firm in the market is associated with a reduction of the price leader's markup which is equivalent to competition between existing firms for an additional three weeks in the product life cycle. Our results support search theoretic models and contradict models of monopolistic competition. Moreover our results support the existence of price dynamics over the product cycle. They also highlight the substitutability between newly innovated and old expiring technologies and how it varies with respect to competitors' and own brand innovations.

JEL-Classification: L11, L13, L81, D43

Keywords: Retailing, Product Life cycle, Market Structure, Market Performance, Markup, Price Dispersion

*Correspondence: Michael Kummer, University of Mannheim, Department of Economics and ZEW, L7, 1, 68131 Mannheim, Germany, Email: Kummer@zew.de. We thank Irene Bertschek, Sara Ellison, Avi Goldfarb, Gautam Gowrisankaran, Michelle Haynes, Jose Luis Moraga-Gonzalez, Ariel Pakes, Martin Peitz, Philipp Schmidt-Dengler and the participants of the SEEK 2010 Workshop on online markets for valuable comments. We are indebted to geizhals.at and a producer of consumer electronics for providing the data. Andreas Lumesberger and Sergiy Golovin provided invaluable research assistance. We would like to thank the editor and two anonymous referees for very helpful comments. This project was supported by the SEEK research program (ZEW).

1 Introduction

Under reasonably general conditions, the consequences of an increase in the number of market participants are lower prices and lower markups. The empirical assessment of this relation is however not an easy task. Markups are not readily available, and prices and market structure are endogenous: firms may enter in response to perceived profit opportunities or may exit in response to realized losses.[1,2]

In this paper, we use a novel instrumental variables strategy to investigate the interaction between market structure and market performance in e-commerce. We use data for digital cameras from an Austrian online price-comparison site (price search engine). We observe the firms' retail and input prices as well as all their moves in the entry and the pricing game. When we measure the rate at which markups decline towards zero, we account for the endogenous timing decision to list a specific product by using previous listing decisions as instruments. In addition, we include product fixed effects to capture unobserved quality and design features of the specific cameras as these might be correlated with both markups and firms' entry. To obtain a full picture of the underlying model of competition, we then follow Baye et al. (2004) and Haynes and Thompson (2008a) and analyze measures of price dispersion as well.[3]

We further analyze the relation of the number of firms and markups across the product life cycle. Products in e-commerce are very often only listed for a short time, which allows us to observe products from birth to death.[4] This is important for three reasons: i) Entry

[1]One way to account for endogeneity is developing a structural model of market structure, entry, and exit. The pioneering study on entry into local markets by Bresnahan and Reiss (1991) shows that the first two or three entrants have the largest impact on market price, and that later entrants do not significantly reduce market price any further.

[2]Experimental evidence of this relation goes back to Selten (1973), who coined the statement "four are few and six are many."

[3]For example, monopolistic competition predicts markups and price dispersion to go down when the number of firms increases (Perloff and Salop, 1985); while in a model with heterogeneity in consumers' search cost and producers' marginal cost the latter would go up (Carlson and McAfee, 1983).

[4]The average span of the product life cycle of digital cameras amounts to 167 days in our data.

in such a market is particularly easy because an existing firm only has to decide whether to list a new camera or not. This low entry cost makes the number of firms volatile and provides an optimal testing field.[5] ii) Several researchers have claimed that competition or the absence thereof is particularly important at the beginning of a product life cycle, while later on, competition may matter less.[6] In particular, when a new product emerges on the market and consumers are uncertain about their tastes, they may postpone their purchasing decision. Firms react to this uncertainty of demand and various price dynamics might be the consequence.[7] iii) Finally, we investigate the effect of substitutes on the markup over the product life cycle and are interested in differences between newly innovated and old expiring technologies as well as between own brand and rivals' brand products.[8]

We are not the first to investigate the relation of market structure and market performance in e-commerce. Previous studies such as Brynjolfsson and Smith (2000), Baye et al. (2009), Baye et al. (2003) and Haynes and Thompson (2008a), however, do not take the endogeneity of the number of sellers and product life cycle effects into account.[9] Baye et al. (2003) and Baye et al. (2004) look at price dispersion using various metrics. Baye et al. (2004), for example, analyze price dispersion measured by the relative price gap (the difference of the first and second price) and show that it decreases as a function of the

[5]In a recent survey, Martin (2012) argues that market structure may adapt only slowly to long-run equilibrium levels and many entering firms may be atypical fringe firms unable to influence market structure at the core. While this describes well-established markets with market leaders and high advertising requirements, market structure in e-commerce is different: due to the cheap and easy establishment of online shops, many such shops operate only online.

[6]Examples include Berry (1992), Campbell and Hopenhayn (2005), Carlton (1983), Davis (2006), Dunne, Roberts and Samuelson (1988), Geroski (1989), Mazzeo (2002), Seim (2006), and Toivanen and Waterson (2000, 2005). For a survey see Berry and Reiss (2007).

[7]See, for example, Bergemann and Välimäki (2006a,b), who analyze dynamic price paths in monopolistic settings and find that in mass markets prices should decrease over the product cycle.

[8]Klepper (2002, 1996) describes the evolutionary pattern of birth and maturity of technologically progressive industries and we apply and extend the predictions of his model to the market of consumer electronics.

[9]Barron et al. (2004) analyze the relationship of markups and price dispersion and the number of firms using data from gasoline retail markets. They find that both markups and price dispersion decrease as the number of firms increases and interpret this as evidence in favor of models of monopolistic competition.

number of firms, but not over time. Haynes and Thompson (2008a) use data on 400 digital cameras in the US and show that with more firms in the market prices go down and dispersion increases. Ellison and Ellison (2005, 2009) examine the competition of internet retailers and identify different strategies that are applied in online markets to cope with the increased price sensitivity.

The empirical literature investigating the market structure along the life cycle of a consumer product is rather small. Haynes and Thompson (2008b) take a first step to explain entry and exit behavior in a shopbot. To do so, they estimate an error-correction model and show that entry into and exit from a market are correlated with a measure of lagged price-cost margins and the number of competitors. Barron et al. (2004) mention the life cycle, but use it only as a control variable. In the marketing literature, Moe and Yang (2009) analyze the product life cycle in e-tailing. However, their data did not allow them to consider the endogeneity of entry and exit. Hitsch (2006) considers the dynamic decision problem of a single firm that is uncertain about the demand for a new product and shows that in the ready-to-eat cereal industry the value of reducing uncertainty is large. This indicates that there are product cycle effects which should be accounted for.

For e-commerce in Austria, we find a highly significant results of the number of firms on markups. *Ten* additional competitors in the market are associated with a reduction of the median markups by 0.23 percentage points and the minimum markup by 0.55 percentage points. However, accounting for the potential endogeneity of markups and the number of firms in the market, we see a substantially higher negative outcome: *ten* additional retailers tend to reduce the markup of the median firm by 0.95 percentage points and the markup of the cheapest firm by 1.24 percentage points. We also find that having one more firm in the market apparently reduces the markup of the price leader by the same amount as the competition between existing firms in a period of three additional weeks in the product life cycle. If we abstract from any dynamic or product life cycle effects, our results support the validity of search theoretic models such as Carlson and McAfee (1983) or Baye and Morgan (2001) and contradict models of monopolistic competition.

We use firms' past listings decisions as an instrument. We argue that this is a valid instrument as products offered in different markets some time ago should have no direct influence on prices and sales of current products. Potential threats to this identification strategy are the timing of past listing decisions and the similarity of products. We thus run robustness checks on the instrument by varying the timing of firms' past behavior and using markets farther away in terms of time or model specification from our chosen product market. The results of these robustness checks show qualitatively and quantitatively very similar results to our preferred specification.

Furthermore, we find that there is a highly significant age effect, i.e., the longer a product is on the market the lower are markups. Although Bergemann and Välimäki (2006a) only consider monopolistic markets, this result could be interpreted in support of their price dynamics when consumers are uncertain about their tastes of a product which newly emerged on the market and when there is social learning.

Our main results on the effects of competition are robust to the inclusion of varying numbers of substitutes over the life cycle of the product. For products with a higher number or substitutes we measure lower markups, in particular if the substitutes are newer products and come from rivals rather then from the same brand.

The remainder of the paper is organized as follows. We present the theoretical predictions derived in the literature in Section 2 and describe the data as well as the empirical strategy in Section 3. We discuss our estimation results in Section 4 and conclude in Section 5.

2 Theoretical Predictions

Our paper aims at explaining the effect of market structure on markups and on price dispersion. While under reasonably general conditions the direction of the effect is clear, search cost may complicate things. Therefore, we now discuss the various models and potential hypotheses. As our empirical analysis also includes the effects over the product life cycle, we discuss the literature on price dynamics and industry dynamics with overlapping technologies as well.

As earlier studies have argued, models allowing for price dispersion in a homogenous market have been classified into three groups:[10] i) First, search-theoretic models (Varian (1980), Rosenthal (1980)), which evoke price dispersion by introducing heterogeneity in the search costs of consumers. These models predict that an increased number of sellers results in a larger price dispersion and – somewhat counter-intuitively – a higher average price. Baye and Morgan (2001) is an example that directly considers e-commerce. They get rid of the counterintuitive result on market structure by allowing for different groups of consumers (a price-insensitive group and price-sensitive consumers who take advantage of zero online search costs) which makes the retailers' randomization over prices rational. ii) Models of monopolistic competition (Perloff and Salop, 1985) can account for price dispersion when extended by asymmetries across firms such as heterogeneous producer cost or heterogeneous producer demand (Barron et al., 2004). These models would predict that a larger number of sellers is associated with a lower average price and smaller price dispersion. iii) Carlson and McAfee (1983) present a search-theoretic model that accommodates two sources of heterogeneity by assuming a non-degenerate distribution of producers' marginal cost and heterogeneous visiting cost of consumers. Here, the prediction is that average prices would go down while price dispersion would rise.

[10]See, for example, Barron et al. (2004) or Haynes and Thompson (2008a).

In the purely search theoretic approach as it was coined by Varian (1980) two types of consumers are present in the market. One type has very low search cost and will hence always buy at the cheapest shop, while the other one with high search cost will buy from a random shop. As a consequence shops have to strike a balance between aiming to be the cheapest shop and selling to all the price-sensitive consumers or confining themselves to their share of price-insensitive consumers but selling to them with a higher markup. In such a setup, everything else being equal an increased number of sellers results in a larger price dispersion and – somewhat counter-intuitively – a higher average price.

Baye and Morgan (2001) theoretically investigate the market for information gatekeepers. They analyze the behavior of firms listed on a price-comparison site as well as the behavior of the monopolistic shopbot. Shops, which have a local monopoly in their town, have to choose between serving only the uninformed population of their own town or advertising on the price-comparison site to potentially serve informed customers in all other towns as well. Consumers, on the other hand, have the option to subscribe to the price-information site or to remain uninformed. In the first case, they can choose from among all shops, but in the latter case they can only buy locally. Given the site's behavior and the share of consumers using the site, the model predicts that the shops will randomize over prices in the price setting equilibrium. They do so in order to maintain positive markups without being undercut by their opponents *with certainty*. Thus, they generate price dispersion in the market for this homogenous product. The impact of more competition on the platform is not explicitly analyzed in the model. Yet it is relatively easy to see that the minimum price (the lower bound of the support of the price distribution in their model) is decreasing in the number of firms, whereas the range of the distribution (price dispersion) increases with an increasing number of firms.

In models of monopolistic competition it is assumed that consumers perceive products to be different across sellers. If all sellers have the same marginal cost, each consumer draws her valuation for the good offered by each seller from a common distribution, and demand is symmetric, then Perloff and Salop (1985) show that an increase in the number of sellers yields an increase in the price elasticity of individual firm demand, lowers the markup and the equilibrium price. If demand is asymmetric and the number of different seller types is constant, an extension of the analysis (Barron et al., 2004) indicates that the increase in the number of sellers of each type will yield a reduction of markups and prices through an increased price elasticity across sellers. Because of the reduction in markups for sellers and common marginal cost, the variance in markups decreases with an increase in competition. As a consequence the price dispersion diminishes.[11]

Carlson and McAfee (1983) assume monopolistic competition with heterogeneous firms in market where heterogenous buyers search. Buyers look for the best price until their expected return from visiting one more shop is smaller than their search cost. Shops, which have heterogeneous marginal cost, use pricing rules which depend on the average price in the market. While price dispersion is due to the heterogeneous marginal cost, they explicitly conjecture, that reputation or heterogeneous visibility in advertisement would also generate equilibrium price dispersion. In equilibrium all shops in the market make positive profits (those who do not are predicted to leave the market), yet the most efficient ones make larger profits than the others. An increased number of shops, all else equal, leads to lower prices and a modest increase in price dispersion, which is bounded from above by the heterogeneity of the shops. As far as the informational requirements of the model are concerned the model's assumption that every change of a shop's pricing policy is perceived fits well with the setting on a price comparison site. However, the reasoning of sequential search is somewhat at odds with such a market setting.

The models discussed so far are static models. In a dynamic context, Bergemann

[11]See Barron et al. (2004) for an extensive discussion on the models of monopolistic competition and their predictions.

and Välimäki (2006a,b) investigate the intertemporal incentives of a new buyer who is uncertain about her tastes for an experience good. Their model of optimal pricing assumes a monopolist that sells a new experience good over time to a population of heterogeneous buyers. While oligopolistic competition is not analyzed, these models provide insights into the intertemporal pricing effects per se. For example, Bergemann and Välimäki (2006a)'s results show that markets can be classified into mass and niche markets. The dynamic equilibrium prices of mass markets decrease over time and buyers purchase in all periods. In a niche market, however, not all consumers buy at the static monopoly price. Therefore, the monopolist initially offers low prices to capture a larger share of consumers. This is at the expense of targeting the more solvent consumer group of the market. The results also hold in the context of social learning which fits our market of digital cameras of one brand producer best. From their analysis we expect price paths that could be either decreasing or increasing, at least at the beginning of the product cycle. As we investigate the effect of the number of firms over the product life cycle, our empirical results may also provide stylized facts theoretical models could incorporate.

In his seminal work Klepper (1996) describes the evolutionary pattern of birth and maturity of technologically progressive industries in form of overlapping product life cycle: Innovation, entry, growth, decline and exit are driven by the way new technologies evolve over time.[12] Consumers are assumed as the driving force behind this evolvement: as they experiment with alternative technological variants they form a view on their preferred variation and decide on the success and failure of different offered technologies. As a consequence overlapping product life cycles emerge in which existing technology rivals with newly innovated and old expiring technologies. Although Klepper has its focus on major technological innovations his model can also be applied to our product markets on

[12] Further literature concentrating on market growth as the major driving force that explains market entry and exit are Schmalensee (1989), Scherer and Ross (1990), Shaked and Sutton (1987), Sutton (1991), Sutton (1998), Dasgupta and Stiglitz (1980), Bresnahan and Reiss (1991) and Asplund and Sandin (1999).

consumer electronics in which the product innovations manifest as additional or increasingly powerful camera features. Our empirical model differs from the work of (Klepper, 1996) and (Klepper, 2002) in various ways. Whereas Klepper focuses on the innovational process of manufacturers our focus lies on the retailers' markup. Moreover, Klepper does not explicitly distinguish whether the manufacturers' innovation (new product) competes with the competitors' or the own products' life cycle. While it typically always makes sense to skim off the rivals' rents by introducing new and better products, an early launch of new product generations might cannibalize the sales of the the firms' own and earlier introduced products.

3 Data and Empirical Strategy

Price search engine: In our analysis we use data from the only Austrian price comparison site, www.geizhals.at. At the time of our analysis Geizhals.at listed on average price offers from 1,200 firms for 200,000 products.[13] The business model of Geizhals is as follows: the retailers have to pay a fixed fee for each referral request of a customer to the respective e-shop.[15] If the retailer agrees to embed the Geizhals logo and link on its website, a reduced fee is paid. If the total of these click-dependent fees does not exceed a certain limit, the online store has to pay a flat fee. The electronic retailers can list as many products they want and can change the prices as often they want, free of charge. There is also no cost if retailers decide to suspend a certain price quote temporarily. Hence, apart from a relatively small flat rate and the click-dependent fee retailers are not confronted with entry or exit

[13]Recently, Geizhals has expanded to other European countries, including Germany, Poland, and the United Kingdom.[14] This recent internationalization led to a substantial expansion of products and supplying retailers. At present, the website offers more than 723,000 products with 64 million price quotes that can could be actualized by the retailers several times per day. In January 2012, Geizhals registered 3.1 million "unique clients." The number of unique clients is calculated from the number of different terminal devices (PCs, PDAs, etc.) used to access a website.

[15]A referral request is a click by the customer on the link of an online shop at Geizhals.at. After the click, the online shop of the retailer opens in a new browser window.

costs in the different product markets.[16] By this construction, Geizhals has the incentive to increase its profits by permanently extending the number of retailers. However, at least in Austria, Geizhals has already acquired such a strong reputation among customers that online stores cannot afford not to be listed at geizhals.at as the market is dominated by this price search engine. Hence, our data cover the whole electronic retailing market in Austria.

Available data: For the study in this paper, we use *daily* data on 70 items (mainly digital cameras) from a major hardware manufacturer,[17] which were listed during the period from January 2007 through December 2008.[18] We define a camera's birth by its appearance on Geizhals.at. The cameras were offered by up to 203 sellers from Austria and Germany. Fixed effects for the different products will control for the unobserved heterogeneity of goods and traders on the varying product markets.

For time t (measured in days), we observe for each product i and retailer j the $price_{ijt}$,[19] the $shipping\ cost_{ijt}$ posted at the website,[20] and the $availability_{ijt}$ of the product.[21] Additionally, we observe the customers' referral requests ($clicks_{ijt}$) from the Geizhals.at website to the retailers' e-commerce website as a proxy for consumer demand.[22] Customers have the option to evaluate the *(service) quality* of the firms on a five-point scale, the average of which is listed together with the price information on Geizhals.at. *Wholesale prices* for

[16] Of the 1,200 retailers at the time of our analysis, only a very small number of retailers have other contracts with Geizhals.at, e.g., they pay only for products actually sold.

[17] The hardware manufacturer is a multinational corporation specializing in manufacturing electronic equipment in several areas. The manufacturer asked to keep its name anonymously.

[18] For our instrumentation strategy, we use also the product life cycle of cameras entering the market starting from May 2006.

[19] We would like to stress that in our market transaction prices are equal to posted prices. Consumers of digital cameras are most often no firms and have no bargaining power.

[20] Shipping cost is the only variable that has to be parsed from a text field. We use the information on "cash in advance for shipping to Germany," which is the type of shipping cost most widely quoted by the shops. Missing shipping costs are imputed with the mean shipping cost by the other retailers and controlled for with a dummy for imputed values.

[21] We coded two dummies whether the product was available immediately (or at short notice) or within 2-4 days.

[22] See Dulleck et al. (2011) for a description of the data.

each product i at time t were obtained from the Austrian representative of the international manufacturer. We do not claim that these wholesale prices correspond perfectly to the retailers' marginal cost. Even though the manufacturer's distribution policy indicates that the retailers should be served by the local representative, it may happen that single retailers procure commodities from, for instance, the Asian market. Moreover, the local representative might offer special promotions including lower wholesale prices in exceptional cases, e.g. if a retailer commits to promoting the manufacturer's good in a special way. Finally, it must be mentioned that the retailers in e-commerce might have additional costs each time they order in addition to the wholesale price. We assume this additional cost to be constant for all stores. Despite these qualifications, our measures are a very good proxy for the actual marginal cost of the retailers.[23] $Price_{ijt}$ and $wholesale\ price_{it}$ were used to calculate the firms' $markup_{ijt}$[24] according to the Lerner index and the markets' $price\ dispersion_{it}$.

Organization of data: We reorganized the data in such a way that the product life cycles of all digicams start at the same day 1. Hence, we have shifted the product life cycles of the digicams so that we can analyze the impact of market structure on markup and price dispersion in each of the different stages of the product life cycle[25]. Our panel consists of 70 product life cycles. We define the end of a product life cycle as the point when the number of referral requests diminishes to less than 500 *clicks*. Thus, we use a daily unbalanced panel with information on the *products' age*, the *number of firms*, *average*

[23]According to the Austrian distributor the Austrian and German lists of wholesale prices are almost identical. Note the manufacturer's incentive to keep cross-border sales between distributors and retailers as low as possible - an argument which supports the reliability of our wholesale prices as indicator for marginal cost.

[24]To account for the problem that high markups might be economically irrelevant, we run all our regressions also weighted by $clicks_{ijt}$. We do not observe quantities sold, only the clicks when a consumer goes from the Geizhals.at website to the retailers' e-commerce website. The clicks do, of course, not reflect the actual demand, but we think is a valid indicator.

[25]We control for the contemporaneous structure of calendar time with dummies for each different month in the dataset and with the number of substitutes at the respective calendar times.

markups, the *markup of the price leader*, different measures for price-dispersion, and the *number of clicks* for product market i over time t.

Descriptives: Table 1 contains summary statistics of the data. Our dataset includes 70 complete product life cycles. The average length of a model's life cycle is 240 days with a standard deviation 133 days. Each observation in the descriptives refers to a single product i at a given day t in the product life cycle. We will use the markup (Lerner index) and the price dispersion as endogenous variables. Whereas the median markup amounts to 18% on average, the mean markup for price leaders is only 4.8%. These numbers are of comparable size to those in Ellison and Snyder (2011), who report an average markup of 4% for memory modules on Pricewatch.com. We use different measures for price dispersion: the *coefficient of variation* and the *standard deviation of the distribution of prices*, as well as the *absolute price gap between the price leader and the second cheapest price*. On average, a product life cycle amounts to 166 days with a mean of 104 firms offering the digicams.

Figure 1 shows that the estimated markup declines with age, and, more importantly, as the number of firms increases. The decline in markups with the number of firms is a rather smooth phenomenon, and not as quick as one might expect in perfectly transparent e-commerce markets. Even with 70 and more firms in the market, there is a positive markup. In the top left panel, the $median\ markup_{it}$ is scattered against the $number\ of\ firms_{it}$ in the corresponding market; the top right panel shows the average pattern. The number of firms ranges from 0 to more than 200 and the median markup ranges from 0 to 35%.

There are also some negative markups, especially for the minimum price firms, where the average markup of the price leader is only 4.8%. In our dataset, we observe negative markups for 26.9% of all best-price offers. This is in line with Ellison and Snyder (2011), who also report a substantial number of price offers with negative markups in the case of Pricewatch.com. Negative markups might have several possible causes: they might simply point to sellouts after overstocking, they might hint to cases where retailers are not procuring products via official retail channels, or they might indicate the use of loss leader strategies where a digicam is offered at a price below marginal cost to attract new customers or to make profits with complementary goods.

In the middle row, the median markup is plotted against the age of the product. We typically observe a camera between seven and 15 months. Again, the markets' median markups fall with the duration of the product life cycle. Our assumption of constant variable costs for e-tailers squares well with the flattening of the markup after 3-4 months. In the lower row, the number or retailers is plotted against the age of the camera: there is a steep increase in the number of listing firms at the very beginning of the life cycle, whereas after 12 months the number of firms is declining again. This average pattern hides some heterogeneity, which can be seen on the left-hand side. Some cameras are listed only by a small number of firms (between 20 and 60). Then there is a group of products which is offered by about 60 shops, and finally, the third and largest group of cameras is listed by roughly 150 shops and more. This segmentation can be explained by a specialization of shops on certain product categories. Whereas some shops are focused in their assortment on mass products (simple digital cameras) others restrict their range of products on highly specialized digital SLR cameras for professionals. This heterogeneity over the products provides a first indication on the importance of product-fixed effects in our estimations.

Empirical strategy: To estimate the impact of market structure on markups and price dispersion, we estimate the following fixed-effects regression as our baseline model:

$$markup_{it} = \gamma + \alpha_1 \ age_{it} + \alpha_2 \ age_{it}^2 + \beta_1 \ numfirms_{it} + \beta_2 \ numfirms_{it}^2 + \omega_i + \tau_t + \epsilon_{it}$$

This model will be estimated for the market's i *median markup* and the *minimum* *markup*; a similar strategy is used for price dispersion which is measured as the *coefficient of variation* and other measures. Life cycle effects are captured by a quadratic age trend. In a later specification, we compute separate splines for each phase of the life cycle to capture varying competition effects over the life cycle of the product. We included month-specific dummies to account for calendar-time fixed effects (τ_t) and product fixed effects (ω_i). Note that the product fixed effects control for observable and unobservable product characteristics on the basis of which the e-tailers build their profit expectations and thus make their decisions on the offered product portfolio. These product fixed effects filter out product specific characteristics which remain constant over the lifespan of the camera [26]. Hence, for our analysis we exploit only the time varying information within the life cycle of each product. However, we have to control for potential endogeneity issues which are based on the time dimension of firms' listing decisions over the product life cycle.

Sources of endogeneity: In all markets, particularly in an e-tailing shopbot market, it is important to treat market structure as endogenous; due of simple and low-cost market entry and exit, e-tailers can easily adapt to changing circumstances by listing a particular product or not. If, for example, unobserved factors temporarily drive up markups for some item, shops that did not sell the item before might move into this market. Thus, we would expect to observe more shops in markets where higher markups can be reaped and vice versa: reverse causation. This, in turn, will result in an upward bias: an estimated OLS coefficient showing the correlation of the number of firms with markups will be less negative than the true causal parameter.

On the other hand, an OLS estimate might suffer from omitted variables bias: variables related to demand, like consumer preferences or actual sales are unobserved, but might be correlated with both prices and market structure. Again, a positive correlation between

[26]For example, they control for the shops' expectations in best-seller products with a large number of sellers versus non-selling products which will be listed by only few shops

demand and market structure and at the same time a positive correlation between sales and prices will lead to an upward bias of an OLS estimate.

In order to overcome both problems, we suggest an instrumentation strategy that can explain market structure but which is both unrelated with demand and has no direct influence on prices.

Instrumentation strategy: In the Geizhals.at data, we observe the complete life cycle of many products together with the firms' decisions to carry the products in the shop. Therefore, we use previous listing decisions as instrument to cope with the endogeneity of the number of firms offering a specific product at time t during the life cycle. There are two dimensions for the endogeneity problem: One dimension refers to the timing decision over the product life cycle. Does a shop list these products from the beginning or at a later point in time? For markets with brand names, part of the listing decisions can be explained by common patterns, such as an established supply relationship, a shop might have with a producer or a wholesale importer, or variations in the availability of the product. These patterns will be independent from the type of camera the shop may supply. Thus we use the timing of previous listing decisions of e-tailers for other brand products of our manufacturer as an instrument for current listing decisions. This is a strong instrument: statistically it does influence current listing decisions strongly. For our instrument to be valid, an exclusion restriction must hold: the listing decision of a series of different products in the past will have no direct impact on markups for another product in the future. This is a plausible assumption because we are using very different markets. We will discuss some threats to this identification strategy below.

Another dimension refers to the assortment the e-tailer offers. There might be shops, which specialize only on products that promise ex ante large markups (or variables correlated with the markup like sales). There might be some heterogeneity of the cameras in the market in terms of aspects as quality and design features that might be correlated with both markups and entry of firms. In our estimation, we will use product fixed effects to capture these unobserved features of the specific cameras. To cope with this dimension of endogeneity we control throughout the paper with product fixed effects for observable and non-observable product features which might have influence on the assortment decision. Given these product fixed effects on the first and second stages of the IV regressions potential selection effects by varying product assortments are controlled for. Therefore, in our analysis we are only exploiting markup variations over the individual product life cycle.

The instrumentation strategy is illustrated with an example in Figure 2. The figure shows the product life cycles for cameras A to H introduced at different points in time. In a first step we are only interested in the listing decision of a single e-tailer, which we shall call an E-shop for the sake of illustration. The vertical dashes indicate the listing decisions (either zero or one) on the third (tenth) day of product life cycle of camera D (camera F). Let us consider whether our E-shop will list a product D on the third day after introduction. This decision is represented by the encircled line on item D. We predict the probability of this event by the E-shop's general probability of listing a similar item that has been on the market for three days. We consider only the last three items that have been introduced *before* product D has entered the market. We then calculate how many of those items were listed by the E-shop on the third day after they appeared. Taking the share gives us an estimate of E-shop's probability of listing product D on its third day of existence. In a second step, we aggregate these probabilities across shops to obtain the predictor of the number of shops that will offer item D on a given day.

This strategy can easily be extended for each day in the product life cycle, giving us a predicted market structure for each day of the product cycle. For the E-shop's listing decision of product F on the tenth day, for example, we use the respective decisions on the tenth day for products E, G, and H. To guarantee the validity of the instrument, we use only products that were introduced before the introduction of the camera in question. Note that for instrumenting F, we ignored products A through D because those cameras lay too far in the past. When calculating the instrument, we fixed the number of earlier introduced cameras to a constant number of three products.[27] In contrast to a constant time interval our approach of fixing a constant number of products guarantees valid standard errors that can be calculated without bootstrapping methods.

First-stage regressions: As we use the time patterns of previous listing decisions in completely different markets, our instrument should not have a direct causal implication for today's markups and price dispersion. Moreover, the listing of a particular type of camera on a particular day in the past should have no influence on sales today. Therefore, the instrument will comply with the necessary exclusion restrictions. Table 2 presents the first stage regressions and shows that the instrument is strong enough to explain the market's actual entry decisions, which are depicted by the number of firms at each point in the product life cycle. Columns (1) and (2) compare the contribution of the instrument to explaining the number of firms, and columns (3) and (4) show the contribution to its quadratic term. The instruments are strongly significant and have the expected sign. The marginal R^2 – due to the large number of fixed effects – amounts from 0.0016 up to 0.028 with F-values close to or above 400.

Could past listing decisions have a direct impact on markups of current products and thereby threaten our identification mechanism? Potential threats concern the fact that markets for past products might be close – either in calendar time or in the type of product – to the current product. Therefore, we tested variations of the instrument to see whether our findings are robust. We applied three systematic variations: i) We vary the number of previous products forming the baseline of previous listing decisions from 3 to 5 or 8. This changes also the length of time used for previous listings. ii) We use the listing decisions

[27] In the next subsection robustness checks on the first-stage regressions are discussed.

of different brand names – instead of our brand name. iii) Our cameras can be roughly divided into several subsubcategories[28]: simple digicams and SLR cameras. We introduce instruments which are only from the same subsubcategory as our product or only from the other subsubcategory. Both previous listings from other brand names as well as those from other subcategories should be seen as exerting even less influence on demand for the current product and, therefore, the exclusion restriction should be easier fulfilled. All these variations do not change our results much.[29]

Caveats with respect to the instrumentations strategy are dynamic aspects of pricing decisions of firms: if there is a remaining correlation between the market structure of past products at a particular day in their life cycles and current pricing decisions of the firm on the same days of the life cycle, then our identification would fail and we would get biased IV coefficients.

4 Results

4.1 Market Structure and Market Performance

Tables 3 and 4 show our basic results for the interrelation of market structure and markups. These baseline specifications are parsimonious, as they consider only the number of firms on the market – either linearly or in quadratic terms – and the product life cycle. We also account for calendar time and product fixed effects. Columns 1 and 3 show OLS estimations, whereas in Columns 2 and 4, our instrumental variables approach is used.

Our results indicate a highly significant and relatively strong correlation of the number of firms with markups. Not accounting for the endogeneity of the number of firms and using OLS, we would estimate the effect of ten additional competitors in the market to

[28]The categorization has been done by geizhals.at.
[29]Results are available from the authors.

reduce minimum markup by 0.55 percentage points and median markups by 0.23. The cheapest firm would react significantly more strongly than the median firm, which might be explained by the high frequency with which prices are changed in online markets, where the cheapest price is a focus of considerable attention from both consumers and firms.

If we instrument for the number of firms, we see a substantially larger negative effect: 10 additional retailers tend to reduce the markup of the cheapest firm by 1.24 percentage points and the markup of the median firm by 0.95 percentage points. These figures are large in economic terms considering the standard deviation of 57 firms in our sample. As discussed above, OLS is likely to underestimate the true absolute effect of an additional firm on the markup, as it does not account for the fact that attractive items also attract more firms. Again, the cheapest firm reacts considerably stronger than the median firm.

In Columns 3 and 4 we use a quadratic specification of the number of firms: it turns out that there is a solid negative but decreasing influence of the number of retailers on markup, both for the cheapest and the median firm. There is virtually no turnaround given the maximum number of firms in our sample of 203. For the minimum markup, the negative influence of the number of firms ceases at 375, and that for the median markup with 195 firms.

Looking at the impact of the product cycle on markups, we observe that markups decrease with time in all regressions. In the 2SLS regressions markups decrease more slowly at product introduction, then more steeply as the product life cycle advances. This decline is stronger for minimum markups – relative to median markups. One reason for this phenomenon could be market saturation.

To investigate the impact of the number of sellers on price dispersion we concentrate on the coefficient of variation (Table 5). While the OLS regressions show a small negative relation between the number of firms and price dispersion, in the 2SLS results (Columns

2 and 4), we see a positive relationship. In the linear case, increasing the number of firms by 10 increases the coefficient of variation by 0.011. For the quadratic case (Column 4) we observe an even stronger increase in price dispersion.

Ignoring dynamic effects, the combined results on markups and price dispersion are compatible with the search theoretic model of Carlson and McAfee (1983) that accommodates two sources of heterogeneities by assuming a non-degenerate distribution of producers' marginal cost and heterogeneous visiting costs of consumers. In addition, the augmented search theoretic model by Baye and Morgan (2001), which features the firms' randomization over prices as a consequence of different user groups, are well in line with our findings. The other search theoretic models are, however, not in line with our findings of a decreasing median markup. Moreover, models of monopolistic competition predict a decreasing price dispersion, a hypothesis which is not supported by our data.

Baye et al. (2004) analyze price dispersion, using a very similar dataset to the one we use here. They focus on the relative price gap (the difference of the first and second price) and show that it decreases as a function of the number of firms, but not over time. Given this finding and a brief analysis of the average price they use a calibration to discriminate the predictions of several clearing house models. We build on their findings, but estimate both markups and price dispersion. Moreover, we analyze the lifecycle dynamics and we can instrument the number of firms based on the lifecycle of previous models. Our OLS estimates corroborate their finding of a decreasing average price, since even the median markup is decreasing in the number of firms. The negative relationship is even stronger, when accounting for the endogeneity in the number of firms. When instrumenting for the number of firms, we do not find a statistically significant decreasing relationship between the number of firms and the coefficient of variation (price dispersion).[30] Note however, that their baseline measure of dispersion is the price difference between the first two offers,

[30]This finding persists even after introducing a click-weighted measure of price dispersion.

whereas we focus on the coefficient of variation (a measure they used in robustness checks).

Shipping cost: While sellers are ranked at Geizhals.at according to prices *net* of shipping costs by default, the price ranking including shipping cost is only accessible via a detour. Figure 5 in the appendix shows a screenshot of the price comparison site. As can be seen there, a quick and easy comparison of shipping costs alone is not possible, because shipping costs can be reported in different waysand there is no automatic ranking possible.[31] As shipping costs are often used as part of an obfuscation strategy (Ellison and Ellison, 2009) it is interesting to see whether shipping costs react to market pressure as well. In Table 6 we report the effects of the market structure on shipping costs divided by median price.[32] As there are different shipping costs available, we concentrate on those mostly observed in the data: shipping costs to Germany when paying cash in advance. Interestingly, the IV patterns are largely the same as in Tables 3 and 4. While OLS predicts a positive relationship between shipping costs and the number of firms, Columns 2 and 4 reveal a robust negative relationship with a small and positive quadratic term.

It is remarkable that more competition seems to decrease also shipping costs. Our estimation shows that ten more firms actually decrease average shipping cost in that market by 62 cents (the mean of the shipping cost is 7.7 Euro).This market structure effect of shipping costs is economically significant, but smaller than the effects on markups: around a quarter of the effect on the median and one-eighth of the effect on minimum markups. These results confirm the visibility argument, that consumers have a much harder time comparing shipping costs than actual prices.[33] This raises interesting questions about whether and

[31] Different possibilities of shipping costs are e.g. standard shipping, shipping to Germany or Austria, and different shipping costs depending on the payment options.

[32] In order to make the results comparable with the percentage values of the markup we are also using a percentage value for the shipping cost.

[33] In a robustness check we computed the total markup including shipping costs to Germany when paying cash in advance. In line with the separate results for markups and shipping cost, we find an even higher effect of market structure on "gross markups": both the minimum and median markups tends to decline significantly with the number of firms. The dispersion of "gross prices" increases.

how different market structures may result in a different role for price transparency. As this is beyond the scope of the paper we left it for further research.

4.2 Life Cycle Effects

In this section we investigate whether the profit-squeezing effect of a higher number of firms is the same in different phases of the product life cycle. Several authors claim that competition might be particularly important at the beginning of the life cycle of a product (e.g. Toivanen and Waterson (2005) and Berry (1992)). On the other hand, at the beginning of the product life cycle pioneer consumers might react less to prices and therefore a higher markup can be achieved. If they are uncertain about their preferences, the opposite may also hold (see Bergemann and Välimäki (2006a)).

To check for different effects of market structure on markups over the product cycle we extend our linear baseline model with crossterms. These crossterms interact the number of firms with four dummy variables for the life cycle of the product (Phase1: days 0-45, phase 2: days 46-105, phase 3: 106-225, phase 4: days 226-800). For ease of interpretation of the coefficients, in Figures 3 and 4 we plot these results for the minimum and median markup, respectively.

In these plots, each line represents a product of a certain age; we plot the curve for products right after their introduction and after 1, 2, 5, and 9 months on the market[34]. Our plots show a consistent pattern. In Figure 3, we observe the pattern for minimum markups. Throughout the life cycle of the product, an increasing number of firms is associated with a fairly similar reduction in markups. The picture is similar for median markups in Figure 4, but here the reaction to an increasing number of firms tends to fall over the life cycle - the splines become flatter.

[34]The dots represent the median of the empirically observed distribution of the number of firms within each phase.

A simple Cournot model would predict that the markup is inversely related to the number of firms, which would lead to a flattening out reaction to increased competition. Our detailed analysis of competition effects over the life cycle shows a more nuanced pattern. For median markups, we do see some flattening out: after the first months of product introduction, median markup reacts still negative, but somewhat less to an increased number of firms. For minimum markups, this is not the case: regardless of the phase of the life cycle of the product, the reaction to more intense competition is the same. This may be due to the higher importance of minimum mark-ups (prices) for consumer demand in online price-comparison sites. Actual transactions are much more concentrated towards the lowest prices. An intuitive argumentation would be that in online price-comparison sites, where prices are very transparent, it does not make sense for newcomers to start with median prices. Only very low prices will catch the attention of customers. In addition online stores on Geizhals.at can follow the prices of their competitor over the platform. Although we do not know whether the retailers actually follow all their competitors, we are convinced that they know when their lost their leading position and to whom. An increased competition effect might be the results. In contrast to existing literature stressing that the first two or three entrants have the largest impact on markets prices (e. g. Bresnahan and Reiss (1991)) our empirical results show a different picture: enforced by the transparency of e-commerce markets additional firm entries in all phases of the product life cycle have the same effect on markups of the price-leader.

4.3 Substitutes over the Product Life Cycle

So far, we considered all markets for cameras independently, implicitly assuming no interrelations between cameras of different type which are offered at the same time and on the same platform. This simplification allowed us to describe the market in general terms.

In this section, we enlarge our empirical model by allowing for substitute products. The availability of substitutes may offer an additional channel for competition in such a market; not considering it may seriously bias measured market structure or competition effects.

On the one hand, the availability of substitutes may drive down profits and markups as such; on the other hand, our measure of competition, the number of firms offering the *same* camera, may be misleading: the number of competitors offering a *similar* camera may be important as well. As the technology of these cameras is quickly improving over time, it is important to distinguish between substitutes with an older technology – which we define as products brought earlier on the market – and a newer technology, i.e. cameras which are introduced later. Moreover, the brand of the camera may be decisive: cameras from a rival producer may be stronger substitutes as compared to new cameras from the same producer. The former may target a new camera towards successful rival products, whereas in the latter case firms may fear cannibalization effects in the introduction of successor products and may, therefore, be more careful in the choice of design or timing of a new product introduction.

As there is no natural definition of substitutes, we identify substitutes by a conclusive behavior of searchers on the price comparison site. We will follow the general idea to identify and analyze different search spells (search cluster) for each user of the website geizhals.at. Each search spell should represent the customers' search and information process during the purchase of a specific product. We assume customers to consider all clicked products during the search spell as potential substitutes. The analysis of frequencies and the identification of most frequently clicked pairs of products over all customers give us a statistical foundation for the identification of substitutes. See Appendix A for a description how we have calculated substitutes in our data.

Table 7 extends our 2SLS estimations for minimum markups[35] (compare the benchmark

[35]Similar results can be obtained for the median markup.

case from Table 4, Column 4) by controlling for available substitutes in its different forms. Note, that the number of substitutes is time varying as it counts the number of available substitutes during each day of the life cycle separately. As expected, the coefficient for the number of firms is still significantly negative, but a little bit smaller than in the simple specification. Moreover, the quadratic terms disappears. This may be due to the additional competition effect coming from the substitutes. As compared to the simple specification, the life cycle effect is completely unchanged.

In general, an additional substitute product reduces the markup of the cheapest firm by 0.77 percentage points. Note, that we have defined substitutes in a very narrow sense, with a mean of only 0.63. This markup-reducing effect of additional substitutes is substantial, its economic importance is difficult to judge, though. Bringing a new substitute to the market may open up a new field of competition, which is not easily compared with an additional number of firms: Many firms may offer this new substitute, but the new product does not operate at exactly the same market. To make the effect of substitutes comparable with the direct competition effect we can increase substitutes and number of firms each by 10%: increasing the number of firms by 10% will reduce minimum markup by 1.19 percentage points, whereas 10% more substitutes will reduce markup only by 0.04 percentage points.

Splitting up substitutes into older and younger ones we see our presumption confirmed that technologically more advanced (newer) products represent a larger threat to the minimum markup compared to older substitutes. Furthermore, substitutes of rival brands have a substantially larger negative outcome – minus 2.3 percentage points of the markup. On the contrary, same brand substitutes even have a positive association with the minimum markup. One explanation for this result might be found in the fact that we observe the retailers' and not the manufacturer's markup (although we would assume a high correlation between both). When introducing new products manufacturers apparently leave retailers larger margins if potential and good running older products from the same brand are still on the market. In that way manufacturers might convince retailers to better promote the new brand product with new technology or features.

A look at the most detailed level in Column (4) confirms this presumption. Newer substitutes from competitors have a larger negative effect on the markup than older and more outdated products of rivals. Although we do not measure any effect of newer substitutes of the same brand – as manufacturers may understand their business not to cannibalize the old products' rents – we measure a significant positive association with the markup if older same brand substitutes are still seen as potential substitutes by the customers.

4.4 Robustness

We perform several robustness checks. First, we test the robustness of the basic results by using varying definitions of price dispersion and by using other definitions of markups, Second, we account by weighting for the fact that some of the price offers may attract less attention from potential buyers. Finally, we consider characteristics of shop specialisation and quality.

We experiment with different definitions of price dispersion: apart from the coefficient of variation (the benchmark case from Table 5, Column 4) we use the standard deviation of prices and a coefficient of variation calculated in such a way that the prices are weighted with the number of clicks received. All these variations show a similar pattern: price dispersion increases with the number of firms, in some cases with a decreasing rate.[36]

We investigate further whether our results are influenced by the fact that we treat all product offers symmetrically in our regressions. In particular, in questions of price dispersion researchers typically mistrust the validity of price offers that are much too high (cf. Baye et al. (2004)). This suggests weighing price offers with the number of clicks they

[36]Results are available on request.

receive to give low-ranked and perhaps less reliable price offers less weight. We do this in Table 8, which weighs each offer by how often it was clicked. This also implies that any offers that did not attract clicks by consumers do not enter this specification at all. Our main results are confirmed in this specification. In the case of the minimum markup weighting reduces significance of the linear term to some extent while the squared term gains importance.

Finally, we want to see whether our results are due to changes in the composition of the shops offering an item over the life cycle. In particular, the presence of larger shops, cheaper, or more reliable shops or a higher presence of shops that sell not only online but also have a brick and mortar outlet might affect the outcomes. Therefore, in Table 9 we include the composition of shops in the regression. In this table the first column shows the benchmark estimation (compare Table 4, Column 4). We then add the share of firms that have the item stocked (i.e., immediately available), the share of firms with low reputation (measured by customer feedback), the share of low-price firms (firms offering generally lower prices in other markets), the share of large firms, and the share of shops with a brick and mortar facility. All of these shares are scaled on a range from 0 to 100: for example, if, in Table 9, the share of large firms increases by 10%, this is associated with a drop in median markups by 0.55 percentage points.

When we introduce these measures of market heterogeneity one by one, because they are to a large extent multicollinear, we find our general results completely unaffected. Both, minimum and median markups fall as the share of larger firms (Column 5) and the share of firms with low reputation (Column 3) increase. The sign of the statistically significant variables are reasonable: we would expect lower markups if there are more firms with low reputation and stronger competition in case of an increasing share of large firms in which undercutting of prices might have a larger impact. The decrease of markups is more pronounced for the size of the firms. The share of low-price firms is not correlated with the median markup (Column 4). Finally, the share of firms that have the item in stock (Column 2) and the share of firms also having a brick-and-mortar facility (Column 6) are related to an increase in markups; again, the positive signs confirm the expected price-setting behavior.[37]

5 Conclusions

In this paper, we investigate the interaction between market structure and market performance in e-commerce using detailed data for digital cameras from an Austrian online price-comparison site. We analyze the empirical association of competition with the markup of the price leader and of the median firm. We account for potential endogenous timing decisions to list a specific product by using previous listing decisions as instruments and include product fixed effects to capture the products' unobserved quality and design features. We further investigate the relation of market structure and measures of price dispersion as well as the development of markups over the product cycle.

Our estimation results show a significant empirical association of markups and the number of retailers in the market. Median markups are lower by 0.23 percentage points and the minimum markup by 0.55 percentage points once ten additional competitors have entered the market. We also find that instrumenting is important for estimating the relation between competition and the markup and we see a substantially higher negative effect. With ten additional retailers, the markup of the median firm is reduced by 0.95 percentage points and the markup of the cheapest firm by 1.24 percentage points. Ignoring dynamic pricing effects, we may interpret our results as support for search theoretic models (Carlson and McAfee, 1983). They contradict models of monopolistic competition (Perloff

[37]Analogous regressions for minimum markups show the same signs, except for the presence of low-price shops, having a positive effect.

and Salop, 1985).

Our results are also in line with the theoretical predictions of Baye and Morgan (2001) as well as the results of recent empirical papers by Haynes and Thompson (2008a) for the US online market for cameras and by Campbell and Hopenhayn (2005) for the US brick-and-mortar retail industry. In both cases, the competitive effects of an increasing number of firms persist in a homogenous goods market. Even with more than one hundred retailers we find markups still decreasing.

The analysis of markups over the product cycle further shows significantly lower markups the longer a product is on the market. Our results refer to e-tailing in the presence of a price-search engine with very narrowly defined products. In such a situation, consumers can very easily collect information about prices and seller reliability. Still, it takes a large number of sellers and a relatively long time for firm markups to dissipate. We may interpret this result in support of price dynamic models when consumers are uncertain about their tastes of a product newly emerged on the market and there is social learning (Bergemann and Välimäki, 2006a).

The markup of the price leader diminishes over the life cycle of the product. This allows us to compare the competitive effect of the number of firms to the effect of time: having one more firm in the market reduces the markup of the price leader by the same amount as three additional weeks in the product life cycle. In other words, by waiting three more weeks a consumer will get the same price reduction she would get if she went to a market with one additional firm, ceteris paribus. In reality, waiting longer will typically also increase the number of firms, thus increasing the advantage of waiting.

Finally, our results also indicate support for substitutability between newly innovated and old expiring technologies. The inclusion of potential substitutes in our estimations reveals interesting stylized facts. The amount of substitutes tends to reduce the firms' mark ups. We distinguish between older and younger substitutes as well as own brand or competitors' brand products. Newer substitutes by competitors are associated with larger reductions in markups compared to older substitutes by competitors. Whereas an increasing amount of older substitutes of the same brand leads to higher markups for the younger products,[38] we do not observe changing markups for the older substitutes if more new own brand products are introduced. Lacking testable theoretical hypothesis we do not account for the price setting game which might be involved in the listing behavior of online shops but just focus on the quantity of competing products.

Our results highlight the usefulness of this very specific market for consumer electronics, where product life cycles are particularly short and thus can be fully observed. Thus, analyses of such environments have great potential to shed light on phenomena of markups over the product life cycle, early adopters, and inter-temporal price discrimination.

[38]Manufacturers might, for example, incite retailers with higher markups on the new product if a high number of older own brand substitutes are on the market

References

Asplund, M. and R. Sandin, "Competition in Interrelated Markets: An Empirical Study," *International Journal of Industrial Organization*, 1999, *17* (3), 353–369.

Barron, J.M., B.A. Taylor, and J.R. Umbeck, "Number of sellers, average prices, and price dispersion," *International Journal of Industrial Organization*, 2004, *22* (8-9), 1041–1066.

Baye, M.R. and J. Morgan, "Information gatekeepers on the internet and the competitiveness of homogeneous product markets," *American Economic Review*, 2001, *91* (3), 454–474.

_ , _ , and P. Scholten, "The value of information in an online consumer electronics market," *Journal of Public Policy & Marketing*, 2003, *22 (1)*, 17–25.

_ , _ , and _ , "Price dispersion in the small and in the large: Evidence from an internet price comparison site," *The Journal of Industrial Economics*, 2004, *52* (4), 463–496.

_ , J.R.J. Gatti, P. Kattuman, and J. Morgan, "Clicks, discontinuities, and firm demand online," *Journal of Economics & Management Strategy*, 2009, *18* (4), 935–975.

Bergemann, D. and J. Välimäki, "Dynamic price competition," *Journal of Economic Theory*, 2006a, *127* (1), 232–263.

_ and _ , "Dynamic Pricing of New Experience Goods," *Journal of Political Economy*, 2006b, *114* (4), 713–743.

Berry, S., "Estimation of a model of entry in the airline industry," *Econometrica*, 1992, *60* (4), 889–905.

_ and P. Reiss, "Empirical models of entry and market structure," *in: M. Armstrong & R. Porter (ed.), Handbook of Industrial Organization, Volume 3, Chapter 29, Elsevier*, 2007, pp. 1845–1886.

Bresnahan, T.F. and P.C. Reiss, "Entry and competition in concentrated markets," *Journal of Political Economy*, 1991, *99* (5), 977–1009.

Brynjolfsson, E. and M.D. Smith, "Frictionless commerce? A comparison of Internet and conventional retailers," *Management Science*, 2000, *46* (4), 563–585.

Campbell, J.R. and H.A. Hopenhayn, "Market size matters," *Journal of Industrial Economics*, 2005, *53* (1).

Carlson, J.A. and R.P. McAfee, "Discrete equilibrium price dispersion," *The Journal of Political Economy*, 1983, *91* (3), 480–493.

Carlton, D.W., "The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables," *Review of Economics and Statistics*, 1983, *65* (3), 440–449.

Dasgupta, P. and J. Stiglitz, "Industrial Structure and the Nature of Innovative Activity," *Economic Journal*, 1980, *90* (358), 266–293.

Davis, P., "Spatial competition in retail markets: Movie theaters," *Rand Journal of Economics*, 2006, *37* (4).

Dulleck, U., F. Hackl, B. Weiss, and R. Winter-Ebmer, "Buying online: An analysis of shopbot visitors," *German Economic Review*, 2011, *12* (4), 395–408.

Dunne, T., M.J. Roberts, and L. Samuelson, "Patterns of firm entry and exit in U.S. manufacturing industries," *Rand Journal of Economics*, 1988, *19* (4), 495–515.

Ellison, G. and S.F. Ellison, "Lessons about markets from the Internet," *Journal of Economic Perspectives*, 2005, *19* (2), 139–158.

_ **and** _ , "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 2009, *77* (2), 427–452.

Ellison, S.F. and C.M. Snyder, "An empirical study of pricing strategies in an online market with high frequency price information," *MIT, Department of Economics, Working Paper 11-13*, 2011.

Geroski, P.A., "The effect of entry on profit margins in the short and long run," *Annales d'Economie et de Statistique*, 1989, *15-16*, 333–353.

Haynes, M. and S. Thompson, "Price, price dispersion and number of sellers at a low entry cost shopbot," *International Journal of Industrial Organization*, 2008a, *26* (2), 459–472.

_ **and** _ , "Entry and exit behavior at a shopbot: E-sellers as Kirznerian entrepreneurs," *Unpublished Working Paper*, 2008b.

Hitsch, G., "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty," *Management Science*, 2006, *25* (1), 25–50.

Klepper, S., "Firm survival and the evolution of oligopoly," *RAND Journal of Economics*, 2002, *33* (1), 37–61.

Klepper, Steven, "Entry, Exit, Growth, and innovation over the Product Life Cycle," *American Economic Review*, 1996, *86* (3), 562–583.

Martin, S., "Market structure and market performance," *Review of Industrial Organization*, 2012, *40* (2), 87–108.

Mazzeo, M.J., "Product choice and oligopoly market structure," *RAND Journal of Economics*, 2002, *33* (2), 221–242.

Moe, W.W. and S. Yang, "Inertial disruption: The impact of a new competitive entrant on online consumer search," *Journal of Marketing*, 2009, *73* (1), 109–121.

Perloff, J.M. and S.C. Salop, "Equilibrium with product differentiation," *The Review of Economic Studies*, 1985, *52* (1), 107.

Rosenthal, R.W., "A model in which an increase in the number of sellers leads to a higher price," *Econometrica*, 1980, *48* (6), 1575–1579.

Scherer, F.M. and D. Ross, *Industrial Market Structure and Economic Performance, Third Edition*, Boston: Houghton Mifflin Company, 1990.

Schmalensee, R., "Inter-Industry Studies of Structure and Performance," *in: R. Schmalensee & R. Willig (ed.), Handbook of Industrial Organization, Edition 1, Volume 2, Chapter 16, Elsevier*, 1989, pp. 951–1009.

Seim, K., "An empirical model of firm entry with endogenous product-type choices," *The RAND Journal of Economics*, 2006, *37* (3), 619–640.

Selten, R., "A simple model of imperfect competition, where 4 are few and 6 are many," *International Journal of Game Theory*, 1973, *2* (1), 141–201.

Shaked, A. and J. Sutton, "Product Differentiation and Industrial Structure," *Journal of Industrial Economics*, 1987, *36* (2), 131–146.

Sutton, J., *Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration*, Cambridge, MA: MIT Press, 1991.

_ , *Technology and Market Structure*, Cambridge, MA: MIT Press, 1998.

Toivanen, O. and M. Waterson, "Empirical research on discrete choice game theory models of entry: An illustration," *European Economic Review*, 2000, *44* (4-6), 985–992.

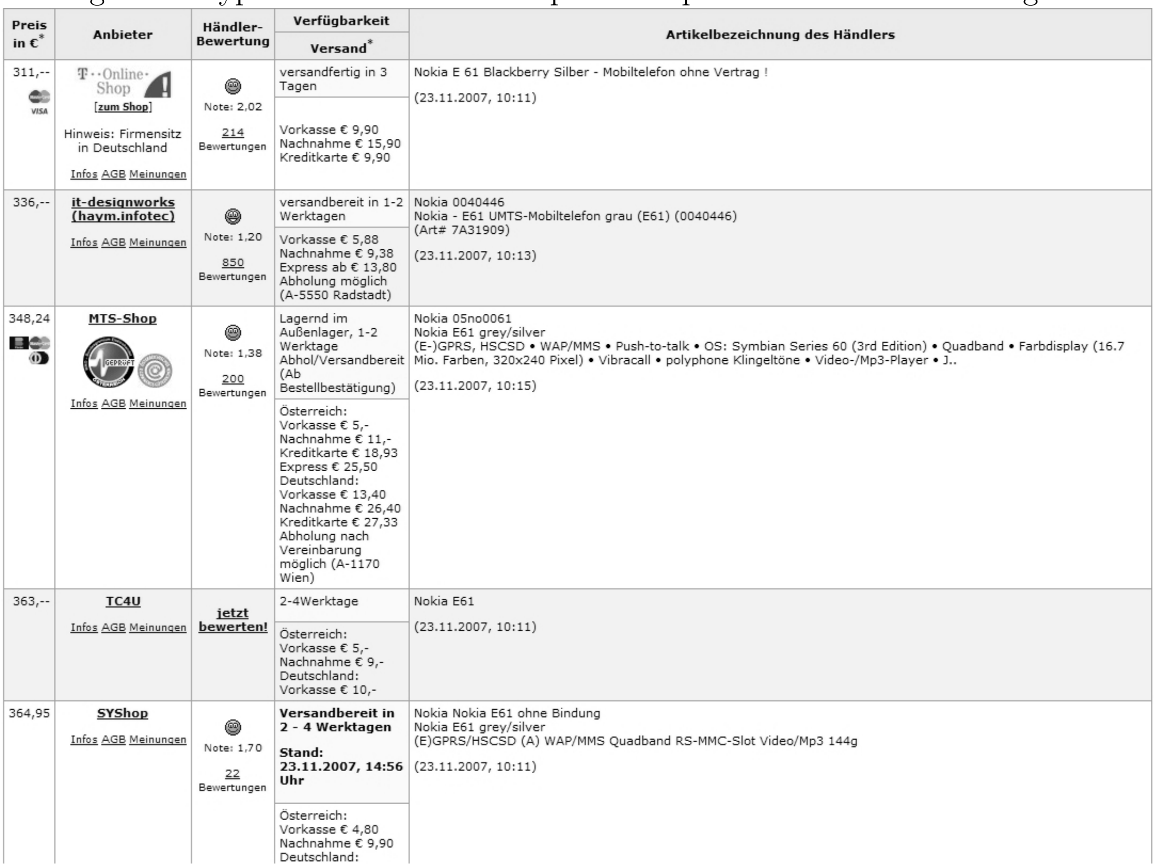
_ **and** _ , "Market structure and entry: Where's the beef?," *Rand Journal of Economics*, 2005, *36* (3), 680–699.

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

A Screenshot of the Price Comparison Site

Figure 5 displays a screenshot of the price-comparison site www.geizhals.at. By default the shops are sorted according to their price (the lowest price listed on the first place). The first column ("Preis in EURO") shows the price, followed by the shop's name ("Anbieter") the average consumer ratings for the shop ("Händlerbewertung"). Shipping cost and how readily the item is available ("Verfügbarkeit Versand") are indicated in the fourth column. Several prices are usually quoted, depending on destinations and method of payment. Consumers can easily see how much they will be charged for shipping (shipping costs are quoted on the same line as prices). The screenshot also shows that consumers themselves have to build the sum of price and the relevant shipping cost, if they use the standard settings. In the rightmost column, most shops provide additional information on the product ("Artikelbezeichnung des Händlers").

Figure 5: Typical screenshot of the price comparison that is shown at geizhals.at



Preis in €*	Anbieter	Händler-Bewertung	Verfügbarkeit / Versand*	Artikelbezeichnung des Händlers
311,--	T··Online· Shop [zum Shop] Hinweis: Firmensitz in Deutschland Infos AGB Meinungen	Note: 2,02 214 Bewertungen	versandfertig in 3 Tagen Vorkasse € 9,90 Nachnahme € 15,90 Kreditkarte € 9,90	Nokia E 61 Blackberry Silber - Mobiltelefon ohne Vertrag ! (23.11.2007, 10:11)
336,--	it-designworks (haym.infotec) Infos AGB Meinungen	Note: 1,20 850 Bewertungen	versandbereit in 1-2 Werktagen Vorkasse € 5,88 Nachnahme € 9,38 Express ab € 13,80 Abholung möglich (A-5550 Radstadt)	Nokia 0040446 Nokia - E61 UMTS-Mobiltelefon grau (E61) (0040446) (Art# 7A31909) (23.11.2007, 10:13)
348,24	MTS-Shop GEPRÜFT Infos AGB Meinungen	Note: 1,38 200 Bewertungen	Lagernd im Außenlager, 1-2 Werktage Abhol/Versandbereit (Ab Bestellbestätigung) Österreich: Vorkasse € 5,- Nachnahme € 11,- Kreditkarte € 18,93 Express € 25,50 Deutschland: Vorkasse € 13,40 Nachnahme € 26,40 Kreditkarte € 27,33 Abholung nach Vereinbarung möglich (A-1170 Wien)	Nokia 05no0061 Nokia E61 grey/silver (E-)GPRS, HSCSD • WAP/MMS • Push-to-talk • OS: Symbian Series 60 (3rd Edition) • Quadband • Farbdisplay (16.7 Mio. Farben, 320x240 Pixel) • Vibracall • polyphone Klingeltöne • Video-/Mp3-Player • J.. (23.11.2007, 10:15)
363,--	TC4U Infos AGB Meinungen	jetzt bewerten!	2-4Werktage Österreich: Vorkasse € 5,- Nachnahme € 9,- Deutschland: Vorkasse € 10,-	Nokia E61 (23.11.2007, 10:11)
364,95	SYShop Infos AGB Meinungen	Note: 1,70 22 Bewertungen	**Versandbereit in 2 - 4 Werktagen** **Stand: 23.11.2007, 14:56 Uhr** Österreich: Vorkasse € 4,80 Nachnahme € 9,90 Deutschland:	Nokia Nokia E61 ohne Bindung Nokia E61 grey/silver (E)GPRS/HSCSD (A) WAP/MMS Quadband RS-MMC-Slot Video/Mp3 144g (23.11.2007, 10:11)

NOTES: The Figure shows which information is shown to consumers on the price comparison web site. The shops' price quotes for a specific item are ordered by price net of shipping cost. The shops' consumer rating, the shipping cost, availability and the shop's own information about the item are provided with the price.

B Appendix: Construction of Substitutes

Based on the idea to identify the most frequently clicked pairs of products during the customers' search processes our calculation of substitutes is operationalized with the following steps:

(1) First, we identify the different consumers according to their individual geizhals.at-cookie which was downloaded to the users' computer by the Website and which uniquely identifies the user of the price search engine.

(2) Some users might search for very different products at the same time (e. g. a vacuum cleaner and a digicam), others are interested in similar products at different points in time (e. g. a consumer buys another camera one year later). We therefore have to sort the customers' click sequence (=referral requests) on geizhals.at into different search clusters. (a) A search cluster includes only clicks in the same subsubcategory[39] and have to contain at least three clicks. To detect whether consumers search identical products at different points in time we apply Grubbs' outlier detection test with a significance level of 95% to identify separate time-separated search cluster. Moreover, we define a period of at least one week as a minimum time span between two clicks to separate a sequence of clicks into two different search clusters.

(3) All clicked products within such a search cluster are considered by the customer as potential substitutes. We exploit this information by measuring the incidence how often certain product pairs are clicked together. The resulting frequency tables for each subsubcategory depict that some product pairs are clicked very frequently together and other pair combinations can be observed only very rarely. We define the list of potential substitutes as the top two percent percentile of these frequency tables. This is a rather conservative measure of potential substitutes, which gives us a relatively low number of substitutes. Repeating the exercise with a lower cutoff point resulted in very similar results.

[39] Geizhals maps its products hierarchically into categories, subcategories and subsubcategories to describe the similarity between the goods. Using the lowest hierarchical level in our analysis guarantees that only very similar products get into the customers' search spells.

Table 1: Summary statistics of the collapsed two-dimensional panel-data with the info on the level of goods and time

	count	mean	sd	min	p10	p90	max
Average price in EUR	15827	949	1399	99	153	2101	7864
Median price in EUR	15827	938	1387	98	151	2085	7990
Minimum price in EUR	15827	853	1294	78	127	1928	7085
Number of sellers	15827	104	58	1	11	168	203
Age in days	15827	166	110	1	35	328	444
Age at death of model in days	70	240	133	35	89	439	444
Wholesale price in EUR	15827	764	1114	79	123	1715	5801
1=item was clicked at least once	15827	.91	.29	0	1	1	1
Aggregate clicks at product i	15827	27	40	0	1	65	646
Average clicks per shop offering product i	15827	.33	.66	0	.0079	.76	21
Markup of price leader (percent)	15827	4.8	7.9	-26	-5	16	35
Median markup for product i in (percent)	15827	18	3	0	15	21	35
Markup of price leader incl. shipping cost in EUR	15445	7.8	6.8	-21	-.16	16	36
Median markup incl. shipping cost* in EUR	15445	19	3.7	-3.8	15	23	36
Coefficient of variation of prices	15735	.087	.16	0	.047	.11	5
Standard deviation of prices in EUR	15735	67	170	0	14	130	6051
Coefficient of variation of prices incl. shipping cost*	15120	.09	.16	0	.048	.12	4.9
Absolute price gap between best and second price in EUR	15735	11	27	-.000098	0	29	516
Average shipping cost in EUR	15445	7.7	1.4	0	6.2	9.2	24
Average reputation on a scale from 1.0 (best) to 5.0 (worst)	15744	1.7	.16	1.1	1.6	1.8	3.5
Average availability (1: in stock, 2: within 2-4 days)	15142	1.5	.15	1	1.4	1.7	2
Share of German shops	15827	.65	.11	0	.57	.74	1
Share of shops with brick and mortar facility	15806	.48	.11	0	.41	.57	1
Click-weighted markup of price leader (percent)	14339	4.9	8	-26	-4.8	15	44
Click-weighted median markup for product i (percent)	14339	9.1	7.3	-26	0	17	96
Click-weighted coefficient of variation of the prices	13577	.054	.091	0	.014	.087	4.1
Substitutes younger same	15827	.011	.1	0	0	0	1
Substitutes older same	15827	.44	1.4	0	0	1	8
Substitutes younger competitors	15827	.014	.12	0	0	0	1
Substitutes older competitors	15827	.087	.41	0	0	0	4
All substitutes	15827	.63	2.1	0	0	1	14

NOTES: The unit of observation is product i at time t (product-time panel). The time variable is days since market introduction.

Table 2: First stage regressions for instrumenting the number of firms

	(1)	(2)	(3)	(4)
VARIABLES	(# of firms/10)	(# of firms/10)	(# of firms/10)^2	(# of firms/10)^2
Instrument for number of firms/10	0.19***	0.85***	1.31***	15.27***
	(0.012)	(0.028)	(0.241)	(0.542)
Instrument for number of firms/10 squared		-0.05***		-0.99***
		(0.002)		(0.035)
Age in months	1.96***	1.75***	34.59***	30.13***
	(0.072)	(0.071)	(1.404)	(1.378)
Age in months squared	-0.12***	-0.11***	-2.36***	-2.17***
	(0.002)	(0.002)	(0.037)	(0.037)
Constant	0.59	-0.70	-52.64***	-79.82***
	(0.451)	(0.445)	(8.768)	(8.602)
Monthly Dummies	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes
Observations	15,893	15,893	15,893	15,893
Marginal R^2	0.0073	0.0283	0.0016	0.0276
Number of goods	70	70	70	70
F-test	420.7	417.6	388.1	392.5

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; The table shows the first stage regressions for instrumenting the number of firms by the instrumental variable that is based on listing behavior of shops over the life-cycles in earlier product markets. Columns (1) and (2) show the first stage regressions for instrumenting the number of firms. Columns (3) and (4) show the corresponding estimation for instrumenting the *square* of the number of firms. Note that the coefficients are much larger in these columns, since the variance of the predicted variable is much larger after taking the square. The R^2 of the baseline regression without the instrument (not included in the table) amounts to 0.4867. The F-Statistics amount to 231,1 for testing Column (1) against the baseline model and to 400,7 for testing Column (4) against Column (3).

Table 3: Minimum Markup

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Number of firms/10	-0.548***	-1.243***	-0.768***	-1.366***
	(0.012)	(0.110)	(0.030)	(0.159)
Number of firms/10 squared			0.012***	0.018**
			(0.002)	(0.009)
Age in months	-3.702***	-2.140***	-3.661***	-2.517***
	(0.109)	(0.274)	(0.109)	(0.166)
Age in months squared	0.027***	-0.068***	0.028***	-0.040***
	(0.003)	(0.015)	(0.003)	(0.008)
Constant	-0.017	1.416*	0.963	2.430**
	(0.672)	(0.773)	(0.682)	(0.981)
Monthly Dummies	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes
Observations	15893	15893	15893	15893
Number of products	70	70	70	70
Adj. R^2	0.559		0.561	

NOTES: The unit of observations is the outcome of product i on day t. The first two columns show the results without a squared term, Columns C and D include the squared number of firms. Columns A and C show OLS panel regressions with product fixed effects. In columns B and D the number of firms has been instrumented. The dependent variable is shown above the columns. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Median Markup

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Number of firms/10	-0.232***	-0.951***	-0.596***	-1.159***
	(0.007)	(0.076)	(0.018)	(0.097)
Number of firms/10 squared			0.021***	0.030***
			(0.001)	(0.005)
Age in months	-2.008***	-0.392**	-1.941***	-1.028***
	(0.065)	(0.189)	(0.064)	(0.101)
Age in months squared	-0.005***	-0.103***	-0.004**	-0.057***
	(0.002)	(0.011)	(0.002)	(0.005)
Constant	8.544***	10.026***	10.164***	11.734***
	(0.398)	(0.533)	(0.398)	(0.597)
Monthly Dummies	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes
Observations	15893	15893	15893	15893
Number of products	70	70	70	70
Adj. R^2	0.274		0.296	

NOTES: The unit of observations is the outcome of product i on day t. The first two columns show the results without a squared term, Columns C and D include the squared number of firms. Columns A and C show OLS panel regressions with product fixed effects. In columns B and D the number of firms has been instrumented. The dependent variable is shown above the columns. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Coefficient of Variation

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Number of firms/10	-0.003***	0.011***	-0.006***	0.016**
	(0.000)	(0.004)	(0.001)	(0.007)
Number of firms/10 squared			0.000**	-0.001
			(0.000)	(0.000)
Age in months	0.024***	-0.008	0.024***	0.002
	(0.005)	(0.010)	(0.005)	(0.007)
Age in months squared	-0.000	0.002***	-0.000	0.001***
	(0.000)	(0.001)	(0.000)	(0.000)
Constant	0.404***	0.372***	0.416***	0.338***
	(0.028)	(0.030)	(0.028)	(0.041)
Monthly Dummies	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes
Observations	15801	15801	15801	15801
Number of products	70	70	70	70
Adj. R^2	0.042		0.043	

NOTES: The unit of observations is the outcome of product i on day t. The first two columns show the results without a squared term, Columns C and D include the squared number of firms. Columns A and C show OLS panel regressions with product fixed effects. In columns B and D the number of firms has been instrumented. The dependent variable is shown above the columns. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Shipping Cost

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Number of firms/10	0.009***	-0.104***	-0.065***	-0.240***
	(0.002)	(0.015)	(0.004)	(0.023)
Number of firms/10 squared			0.004***	0.014***
			(0.000)	(0.001)
Age in months	0.018	0.250***	0.025*	0.019
	(0.014)	(0.035)	(0.014)	(0.023)
Age in months squared	-0.001*	-0.014***	-0.000	0.002**
	(0.000)	(0.002)	(0.000)	(0.001)
Constant	1.303***	1.583***	1.618***	2.383***
	(0.084)	(0.104)	(0.084)	(0.129)
Monthly Dummies	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes
Observations	15441	15441	15441	15441
Number of products	70	70	70	70
Adj. R^2	0.207		0.227	

NOTES: The unit of observations is the outcome of product i on day t. The first two columns show the results without a squared term, Columns C and D include the squared number of firms. Columns A and C show OLS panel regressions with product fixed effects. In columns B and D the number of firms has been instrumented. The dependent variable is shown above the columns. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: The Importance of Substitutes over the Life Cycle

	Minimum Markup			
	(1)	(2)	(3)	(4)
		older vs.	own brand vs.	all
	all	newer	other brands	interactions
Number of firms/10	-1.143***	-1.102***	-0.972***	-0.922***
	(0.147)	(0.151)	(0.149)	(0.154)
Number of firms/10 squared	0.005	0.003	-0.005	-0.009
	(0.008)	(0.008)	(0.008)	(0.009)
Age in months	-2.562***	-2.532***	-2.566***	-2.523***
	(0.163)	(0.165)	(0.164)	(0.165)
Age in months squared	-0.042***	-0.043***	-0.042***	-0.044***
	(0.008)	(0.008)	(0.008)	(0.008)
All substitutes	-0.765***			
	(0.112)			
Substitutes newer		-1.496***		
		(0.279)		
Substitutes older		-0.618***		
		(0.121)		
Substitutes same brand			0.383**	
			(0.173)	
Substitutes other brands			-2.276***	
			(0.236)	
Substitutes newer same brand				0.077
				(0.502)
Substitutes newer other brands				-3.912***
				(0.578)
Substitutes older other brands				-2.038***
				(0.253)
Substitutes older same brand				0.512***
				(0.196)
Constant	1.783*	1.436	0.693	0.438
	(0.951)	(0.971)	(0.967)	(0.998)
Monthly Dummies	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes
Observations	15827	15827	15827	15827
Number of products	70	70	70	70

NOTES: The table is based on the main results in the paper, but includes the number of substitutes for the product. The unit of observations is the outcome of product i on day t. The first columns shows the results when including all substitutes. Columns B differentiates between newer and older substitutes. Column C shows the results for distinguishing same brand substitutes from competitors' substitutes. Column D distinguishes the substitutes along both dimensions. In all columns the number of firms has been instrumented. The dependent variable is the minimum markup. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Markup and price dispersion weighted by clicks.

	(1) clw. min. markup	(2) clw. med. markup	(3) clw. coeff. var.
Number of firms/10	-0.228	0.758***	0.019***
	(0.241)	(0.284)	(0.006)
Number of firms/10 squared	-0.030**	-0.052***	-0.000*
	(0.012)	(0.014)	(0.000)
Age in months	-3.104***	-3.545***	-0.016***
	(0.192)	(0.227)	(0.004)
Age in months squared	-0.015*	0.014	0.001***
	(0.009)	(0.010)	(0.000)
Constant	-0.639	-0.848	0.035
	(1.276)	(1.508)	(0.029)
Monthly Dummies	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes
Observations	14401	14401	13639
Number of products	70	70	70

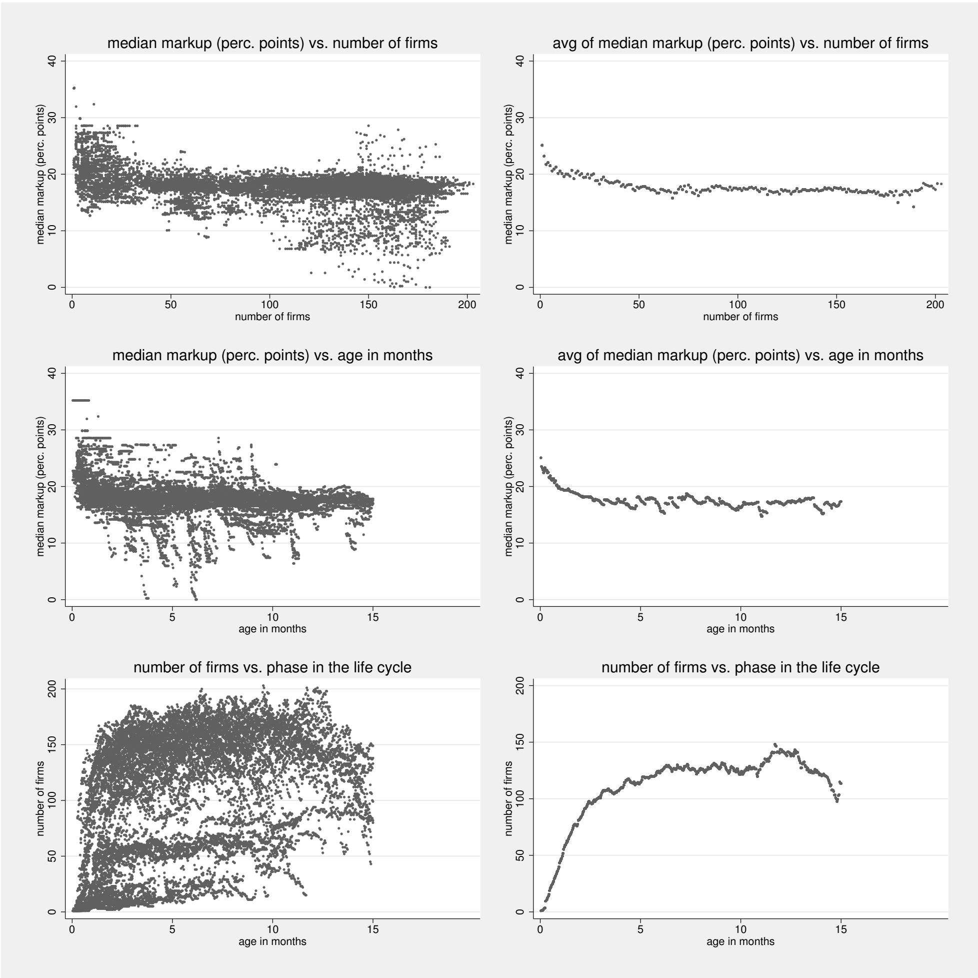
NOTES: The table weighs prices by clicks on the respective product before computing the moments of the price distribution to reproduce the main results in the paper. The unit of observations is the outcome of product i on day t. Each column shows the same estimation with a different dependent Variable. The dependent variable is shown above the columns. The number of firms has been instrumented. Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9: Median markup and the composition of shops

	(1) benchmark	(2) availability	(3) reputation	(4) price level	(5) size	(6) brick/mortar
Number of firms/10	-1.159***	-1.323***	-1.128***	-1.146***	-1.197***	-1.128***
	(0.097)	(0.136)	(0.097)	(0.119)	(0.092)	(0.105)
Number of firms/10 squared	0.030***	0.032***	0.028***	0.030***	0.038***	0.028***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)
Age in months	-1.028***	-0.977***	-1.022***	-1.043***	-1.197***	-1.007***
	(0.101)	(0.112)	(0.101)	(0.115)	(0.095)	(0.099)
Age in months squared	-0.057***	-0.063***	-0.058***	-0.056***	-0.044***	-0.059***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)
Share on stock		0.034***				
		(0.008)				
Share low rep			-0.010***			
			(0.002)			
Share low price				-0.004		
				(0.008)		
Share larger shops on					-0.055***	
					(0.003)	
Share brick/mortar						0.011***
						(0.003)
Constant	11.734***	10.883***	12.090***	11.909***	14.190***	11.096***
	(0.597)	(0.539)	(0.599)	(0.468)	(0.547)	(0.729)
Monthly Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Product Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15893	15893	15893	15893	15893	15872
Number of products	70	70	70	70	70	70

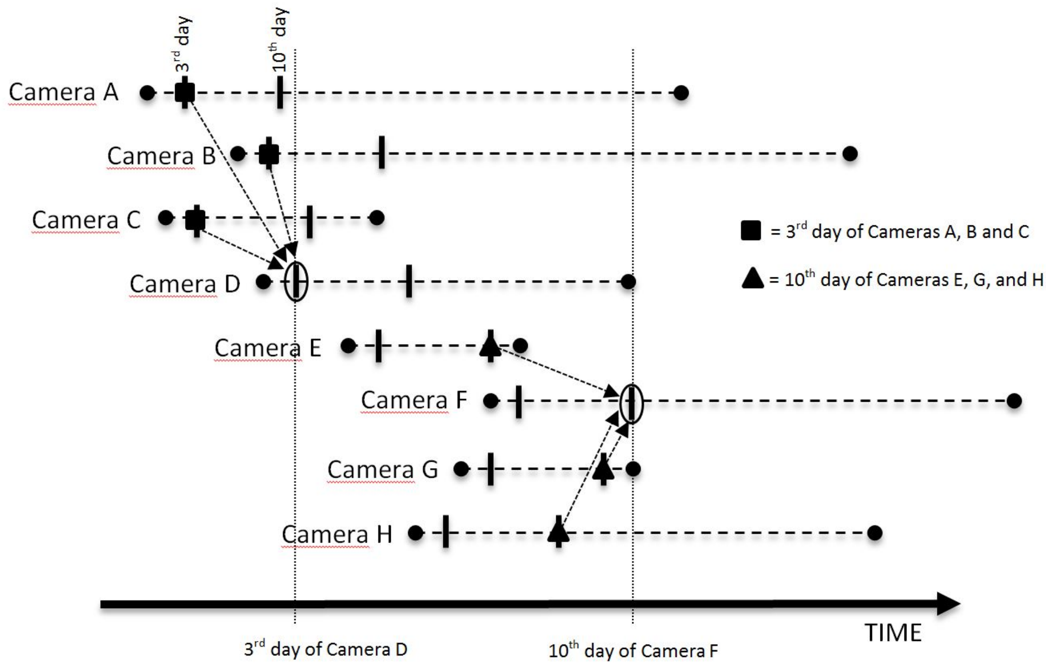
NOTES: The unit of observations is product i on day t. Each column shows the same estimation with a different variable capturing shop composition. Other than in the previous tables, the dependent variable in all columns is the median markup, as it is more relevant in the context of shop composition. The number of firms has been instrumented. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Median markup plotted against the number of firms and age of product



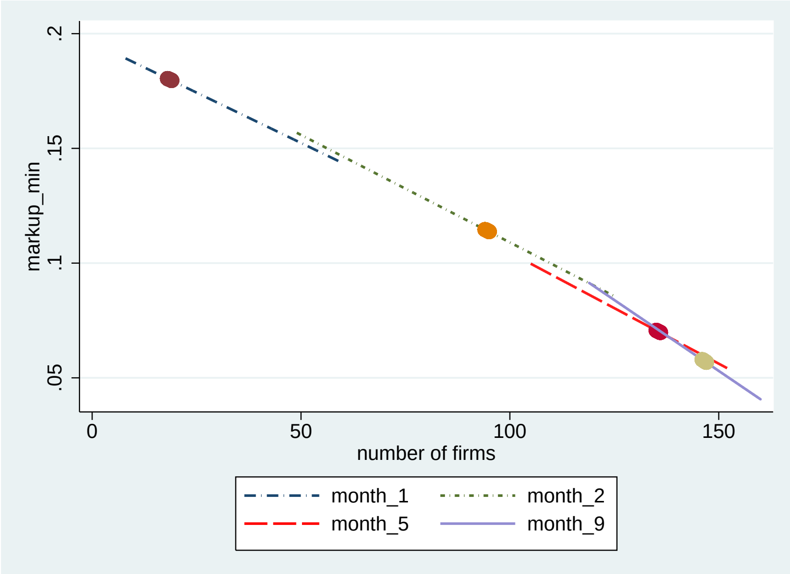
NOTES: The plot shows the empirically observed distributions of the number of firms, age and median markup plotted against each other. In the top left panel, the $median\ markup_{it}$ is scattered against the $number\ of\ firms_{it}$ in the corresponding market and the right column shows the corresponding averages. In the middle row the $median\ markup_{it}$ is plotted against the $age\ of\ the\ product$. In the lower row, the $number\ of\ firms$ is plotted against the $age\ of\ the\ product$ (in months).

Figure 2: Instrument using firm's listing behavior in earlier lifecycles



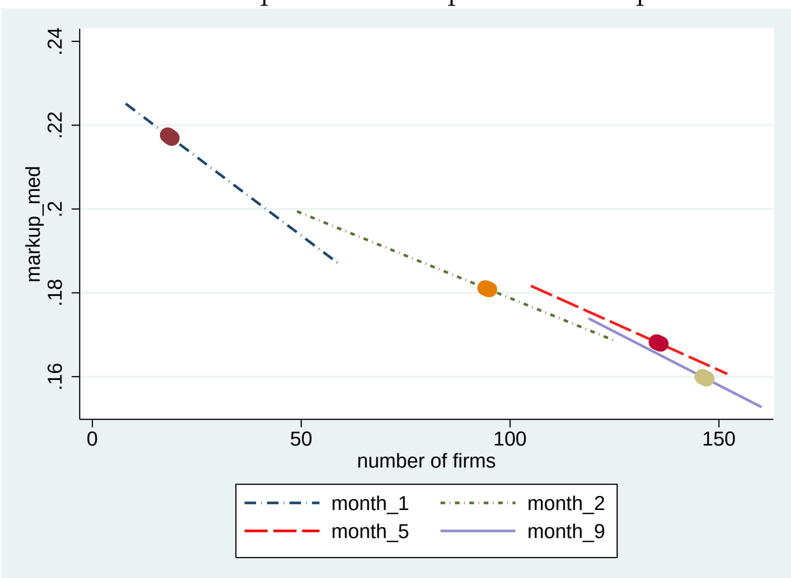
NOTES: If we want to predict how many shops will list a product on any given day q after introduction we use a shop's general probability of listing one of the three items that entered the market before product j, q days after they were introduced. Examples: to predict listing behavior for camera D on day 3 (encircled dash), we would use information on the three cameras A, B, and C on their respective third days of existence (black squares). However, we would not use the information from the cameras that saw light after D was introduced. To predict how many shops listed camera F on day 10 (encircled dash), we would use the information from the three cameras E, G, H which saw light before F was introduced. Cameras A, B, C (on day 10) are ignored for the computations of the instrument for camera F on day 10. These considerations can be applied to each day in the product lifecycle of a camera.

Figure 3: Minimum markup in different phases of the product life cycle



NOTES: Each spline shows the estimated relationship of number of firms and markup at a different point in time (after 1,2,5 and 9 months). The curves are plotted on the range from the 33rd to 67th percentile and the dots represent the median of the empirically observed distribution of the number of firms at the point in time it corresponds to.

Figure 4: Median markup in different phases of the product life cycle



NOTES: Each spline shows the estimated relationship of number of firms and markup at a different point in time (after 1,2,5 and 9 months). The curves are plotted on the range from the 33rd to 67th percentile and the dots represent the median of the empirically observed distribution of the number of firms at the point in time it corresponds to.