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
University of Linz & WIFO, Vienna

Preliminary

October 15, 2011

Abstract

We investigate the effect of market structure on market performance in the market for consumer electronics. We exploit product life cycle information to build an instrumental variable for the number of firms in a market, a variable which hitherto had to be treated as exogenous in comparable studies on seller-behavior in e-commerce.

We combine data from Austria's largest online site for price comparisons with retail-data on whole sale prices provided by a major hardware producer for consumer electronics. We observe input prices of firms, and all their moves in the entry and the pricing game. Using this information for 80 digital cameras, we generate instrumental variables based on the shops' entry decisions in the past. We find that instrumenting is particularly important for estimating the effect of competition on the markup of the price-leader.

*Correspondence: Michael Kummer: Universitaet Mannheim, Department of Economics and ZEW, L7, 1, 68131 Mannheim, Germany, Email: Kummer@zew.de.

This publication is the result of a project sponsored within the scope of the SEEK research programme which was carried out in cooperation with Johannes Kepler Universitaet (JKU) Linz, Austria and the Zentrum fuer Europaeische Wirtschaftsforschung GmbH (ZEW), Mannheim, Germany

1 Introduction

Analyzing the link between market structure and market performance is of central importance in the field of industrial organization. In particular, antitrust and regulatory authorities are interested to know how many firms it takes to sustain competition in a market. For example, the expected relation between the number of firms in the market and market outcomes such as prices or qualities is at the core of merger assessments.

We investigate the interaction between market structure and market performance in e-commerce using data from Austria's largest online site for price comparisons combined with retail-data on whole sale prices provided by a major hardware producer for consumer electronics. We observe firms' prices as well as their input prices, and all their moves in the entry and the pricing game. To measure the rate at which oligopoly margins decline toward zero, we analyze how quickly the break-even price-cost margins fall as the number of market participants increases from one to two firms, two to three firms, and so on. We also look at the impact of market structure on market performance over the product life cycle, as other studies focusing on market structure find that entry has, especially at the beginning of the life cycle, a significant impact on prices.[1]

The use of data from online markets has been pioneered by the studies of Brynjolfsson and Smith (2000) and Baye et al. (2004) who used online data to study the relationship of prices and competition on online markets. They also analyzed the distribution of prices and found that price dispersion increases with the number of competitors. Since then, a large variety of issues have been studied with online data. Baye et al. (2009) have examined the Kelkoo price comparison site and noted that there is a big discontinuity in clicks at the top listed products, a result which can be explained with clearinghouse models. Ellison and Ellison (2005) or Ellison and Ellison (2009) have examined the competition of Internet retailers and have identified different internet-specific firm strategies which are applied in online markets to cope with the increased price sensitivity of online markets.

The prime advantage of e-commerce is the easy availability of large amounts of data on retail-prices at very little cost for the researcher. Moreover it is generally possible to observe all the prices and price changes of the firms and to reconstruct the sequence in which they react to each other. More than that, it is possible to obtain data for many different markets, be it books or consumer electronics. However, the researcher faces the disadvantage that he does not always observe the entire market, but only a segment, and he usually cannot tell whether a posted price has also induced a transaction or not.

Our point of departure is the study by Haynes and Thompson (2008b), which exploits data on digital cameras. Their study, which is related to a similar study by Barron et al. (2004), provides useful insights into the evolution of prices and price dispersion as a function of the market structure. Underlining the potential of research on e-commerce data, the study uses data for about 400 models of digital cameras in the US. It also takes first steps towards taking the life cycle of consumer electronics into account, even though mostly by regarding life cycle effects as a nuisance parameter. Moreover, the authors point to the problem of potentially endogenous right hand side variables and emphasize the need of adequately incorporating seller heterogeneities but they cannot do so, because they only observe the aggregate data that's

[1] Examples are 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).

publicly available. The present study proposes to expand on their analysis, by developing an instrument for the number of sellers in a market and to focus on the determinants of the product life cycle. Haynes and Thompson (2008a) use less detailed data to take a first step towards explaining entry and exit behavior in a shopbot, by estimating an error-correction model and show that the entry and exit into a market is correlated with a measure of lagged price-cost-margin and the number of competitors. Also in the marketing literature Moe and Yang (2009) recently analyzed the product life cycle in e-tailing. However, like the literature in IO, their data did not allow them to take the endogeneity of entry and exit into account.

Our paper sheds light onto the question how the market structure affects the functioning of a market and the price level. Most importantly we are able to take a first step to circumventing the usual difficulty instrumenting the number of competitors in a market. Clearly, even if under potentially reversed signs, these questions are also of great interest to consumers and manufacturers who wish to maximize their benefits. We find a highly significant and strong effect of the number of firms on markups. Ten additional competitors in the market reduce median markups by 0.22 percentage points and the minimum markup by 0.54 percentage points. Accounting for the potential endogeneity of markups and the number of firms in the market, we see a substantially higher negative effect: ten additional retailers reduce the markup of the median firm by 0.88 percentage points and the markup of the cheapest firm by 2.6 percentage points.

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

2 Theoretical Predictions and Relationship of Interest

2.1 Theoretical Predictions

The point of departure for the present study is a series of two papers by Barron et al. (2004) and Haynes and Thompson (2008b). Both confront the predictions of competing model with data that relate the market structure to price level and price dispersion but both have to take the number of competitors in a market as exogenously given, which they themselves point out is possibly not warranted. In what follows we shall briefly summarize their discussions of the differing predictions of the competing models to be tested.

Grossly speaking we distinguish three groups of models which allow for price dispersion and hence a violation of the law of one price: Firstly, search theoretic models (Varian (1980), Rosenthal (1980)), which successfully allow price dispersion by introducing heterogeneity in the search costs of consumers. Secondly, models of monopolistic competition (e.g.: Perloff and Salop (1985)) can account for price dispersion, when extended by introducing asymmetries across firms, such as heterogeneous producer cost or heterogeneous producer demand (cf. Barron et al. (2004)). Thirdly, Carlson and McAfee (1983) present a search theoretic model which accommodates two sources of heterogeneities by assuming a non-degenerate distribution of producers' marginal cost and heterogeneous visiting cost of the consumers. Combining these two types of heterogeneity results in somewhat different predictions about the behavior of price and price-dispersion, given an increase in the number of competitors. Finally, also a simple structure-conduct-performance model (Bain (1951)) can be tested in this context (although with somewhat vaguer predictions), as was pointed out by Haynes and Thompson (2008b).

While all these models differ significantly in their setup, they all have something to say about the impact of market structure on prices and price dispersion. Hence evidence about this relationship is important to test them and to tell which of them is most suitable to think about a market at hand. For the present purposes it suffices to skip a detailed discussion and merely provide a very brief overview over the different predictions off the models.[2]

The first group of search theoretic models, that builds on different consumer types, that are equipped with different search costs (e.g. Varian (1980)), predicts that an increased number of sellers results a larger price dispersion and, somewhat against intuition, a higher average price. The second group of models with differentiated sellers and either production cost or buyer cost asymmetries would expect that a larger number of sellers is associated with a lower average price and smaller price dispersion. Thirdly, the model by Carlson and McAfee (1983) predicts that average prices would go down while price dispersion is expected to rise. According to a structure-conduct-performance model where the incumbents face the threat of entry, prices should decrease or stay equal when more firms enter the market, depending on the strength of the entry-threat. The model is somewhat silent about price dispersion.

3 Data and Empirical Strategy

Price search engine: In our analysis we use data from the largest Austrian price comparison site www.geizhals.at[3]. This platform is Austria's unchallenged market for price comparison for the retailing of consumer electronics. For the study in this paper we use *daily* data on 76 digital single lens reflex cameras from a major hardware manufacturer[4] which come on the market during the period from January 2007 till December 2008[5]. We define a camera's birth by it's appearance on geizhals.at. The cameras were on offer at up to 212 sellers from Austria and Germany.

Available Data: For time t (measured in days) we observe for each product i and retailer j the $price_{ijt}$, the $shipping\ cost_{ijt}$ posted at the website[6] and the $availability_{ijt}$ of the product[7]. Additionally, we observe the customers' referral request ($clicks_{ijt}$) from the geizhals.at website to the retailers' e-commerce website as proxy for the consumers' demand. Customers have the possibility to evaluate the *(service)quality* of the firms on a 5-point scale the average of which is listed together with the price information on geizhals.at. *Whole sale prices* for each product i at time t are obtained by the Austrian representant of the international manufacturer. We do not claim, that these whole sale prices correspond perfectly with the retailers' marginal

[2]For a more detailed discussion of the models the interested reader is referred to the two papers by Barron et al. (2004) and Haynes and Thompson (2008b) our work builds on or to the original papers. Their discussion is clever, concise and insightful, but repeating it here would not add any further insights.

[3]Based on this dataset Dulleck et al. (2011) analyze the search and purchasing behavior of buyers in which the reliability of the retailer gets more important the closer it comes to actual buying decisions.

[4]The hardware manufacturer is a multinational coporation specialized in the manufacture of electronic equipment in several areas. The manufacturer asked to keep his name anonymously. If somebody wants to check the validity of our results we can of course offer more detailed information on the manufacturer.

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

[6]Shipping cost is the only variable which has to be parsed from a text field. We use the information on cash in advance and shipping to Germany, as we lose for this type of shipping cost the least observations. Missing shipping cost are imputed with the mean shipping cost by the other retailers.

[7]If the product is available immediately or at short notice the dummy is 1, if the product is not available, it is 0.

cost. Even though the manufacturer's distribution policy indicates that the retailers should be served by the local representant it can cannot be foreclosed that single retailers procure commodities for instance from the Asian market. Moreover, the local representant might offer special promotions including lower whole sale prices in exceptional cases (e. g. if a retailer commits to promote the manufacturers good in a special way). Finally, it has to be mentioned that besides the wholesale price the retailers in e-commerce might have additional cost for each ordering. Despite of the fact that we cannot guarantee that all e-commerce retailers are balancing their orders according to these whole sale prices they are a very good proxy for the actual marginal cost the retailers' are confronted with[8]. $Price_{ijt}$ and *wholesale* $price_{it}$ are used the calculated the firms' $markup_{ijt}$ and the markets' *price* $dispersion_{it}$.

Organization of data: We reorganized the data in a way so 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 a cross section of 76 product life cycles. This reorganization of data is also important to guarantee that observations are iid. Especially the independence assumption is crucial as listing decisions of e-commerce traders are strategic variables: If we would study product cycles in real time we would have the problem that the listing decisions of digicam X are not independent from the listing decision of the follower model Y. By shifting the product life cycles to identical starting points the iid assumption concerning our data structure is valid. We define the end of a product life cycle if the amount of referral requests diminish to less than 500 remaining *clicks*. Finally, we collapse the data in order to create a panel with products as units of observation and thus obtain a daily unbalanced panel with information on the *products' age*, the *number of firms*, *average markups*, *markup of the price leader*, different measures for *price-dispersion*, and *clicks*.

Descriptives: Table 1 contains summary statistics of the collapsed two-dimensional panel-data. Each observation in the descriptives refer 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 17.8% the average markup for the price leaders falls to 4.6 %. To compare, Ellison and Snyder (2011) report an average markup of 4% for memory modules on Pricewatch.com. We use different measures for the 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. The absolute price gap varies between 0 and 515.9 Euro with a mean of 10.5 Euro. On average a product life cycle amounts to 163 days with a mean of 101 firms which are offering the digicams. A visual inspection of the data (see 2) shows that the estimated markup declines with age, and, more importantly, as the number of firms increases (each observation again corresponds to the data of a single product i at a given day t). However, on average, this pattern is by no means very abrupt, as one might expect in perfectly transparent e-commerce markets. We rather observe a well positive average markup, also with 70 and more firms in the market. In the top left panel, the *median* $markup_{it}$ is scattered against the $number\ of\ firms_{it}$ (in tens) in the corresponding market and the top right panel shows the average. The number of firms ranges from 0 to slightly more than 200 and the median markup have a range from 0% to 37%. It must be noted that we also observe negative markups especially for the minimum

[8]According to the Austrian distributor the Austrian and German wholesale price list are almost identical. Note the manufacturer's incentive to keep cross-border sales between distributors and retailers as low as possible if the manufacturer would pursue substantial price discrimination between countries.

price firms - the average markup of price leaders is 4.6% with a standard deviation of 7.8%. In our dataset we observe for 26.92 % of all best price offers negative markups. This is in line with Ellison and Snyder (2011) who report also a substantial amount of price offers with negative markups for Pricewatch.com. Negative markups might have several possible causes: They might simply point to sell-outs after overstocking, it might be a hint to cases where retailers are not procuring via the official retail channels but exploit price differentials with for instance Asian markets. Finally, loss leader strategies might be responsible for negative markups where a digicam is offered at a price below marginal cost in order to attract new customers or to make profits with complementary goods. In the middle row the median mark-up is plotted against the age of the product. Again the markets' median markups fall on average with the duration of the product life cycle. The spike in the beginning of the product life cycle can be explained with a special price setting behavior of those shops which do not have the digicam available at the short run (not all retailers are delivered by the distributor at the same time). These retailers ask for a very low price in order to postpone the customers shopping decision until the commodity is also available at their own shop. If the product is available the retailer has the incentive to raise the price to the rivals's level. In the lower row, the median markup is plotted against the age of the camera (in months). We typically observe a camera between 7-8 and 15 months. While the line for the averages looks very smooth, the scatter plots on the left of the graphs reveal however, that there is large heterogeneity. Apparently there are three types of digicams: Some appear to be listed by fewer shops (20 and 60, respectively) and then to be taken off the market sooner, whereas another group of cameras seems to be listed by roughly 150 shops on average and then to be taken off the market only after 14 months. The apparent segregation of markets is striking. As expected we observe rather fast market entry within the first two months - after that the amount of firms stagnates. Summarizing the descriptive results it can be stated that markup declines very slowly, given that firms have to compete in prices in this market. Secondly, the life cycle of digital products is short enough to not only allow observing their entire lifecycle, but also observe many thereof, which is the feature our instrumentation strategy shall build on.

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

$$depvar = \alpha_j + \alpha_1 * age + \alpha_2 * age^2 + \beta_1 * \frac{numfirms}{10} + \beta_2 * (\frac{numfirms}{10})^2 + \epsilon_{jt}$$

Dependent variables *depvars* are the minimum markup, price dispersion (measured as the coefficient of variation) and the markup of the mean-price and we regress each of them separately on the number of firms in the market that day. Moreover, we include a quadratic age-trend and thus measure life-cycle effects as a byproduct. However, before we can do so we have to account that it is very easy to list and unlist an item, and that hence the number of sellers can react extremely fast on the market conditions like markups or price dispersion. In all markets - but in particular in an e-tailing shopbot market - it is important to treat market structure as endogenous: due to simple and low-cost market entry and exit, e-tailers can easily adapt to changing circumstances by listing a particular product.

In order to cope with this endogeneity problem we follow an IV-approach and instrument the number of firms. For that purpose we can exploit the long-run availability of markets for brand-name of the full Geizhals.at data. For markets with brand names sellers will typically have an established supply-relationship with a producer or a wholesale importer. In order to instrument the number of firms, we exploit the fact, that we are able to observe the shops behavior in the markets of previously introduced products. We use listing decisions of e-tailers for brand products of our manufacturer in the past as an instrument for current listing decisions. Such past decisions - in particular if they come from different markets (e.g. digicams versus computer products) will influence the listing decisions but will not influence the current market outcome as such.

The *instrumentation strategy* is illustrated with an example in Figure 3: We want to predict whether a shop will list a product on the 10th day after introduction by the shop's general probability of listing a similar item that's been on the market for 10 days. We calculate this probability for shop j and product i by looking at only the group of similar products that were introduced over the six months before product i was introduced. Then we calculate how many of those items, the shop listed on the tenth day after they appeared. Finally taking the share gives us an estimate of shop j's probability to list an item on it's tenth day of existence.

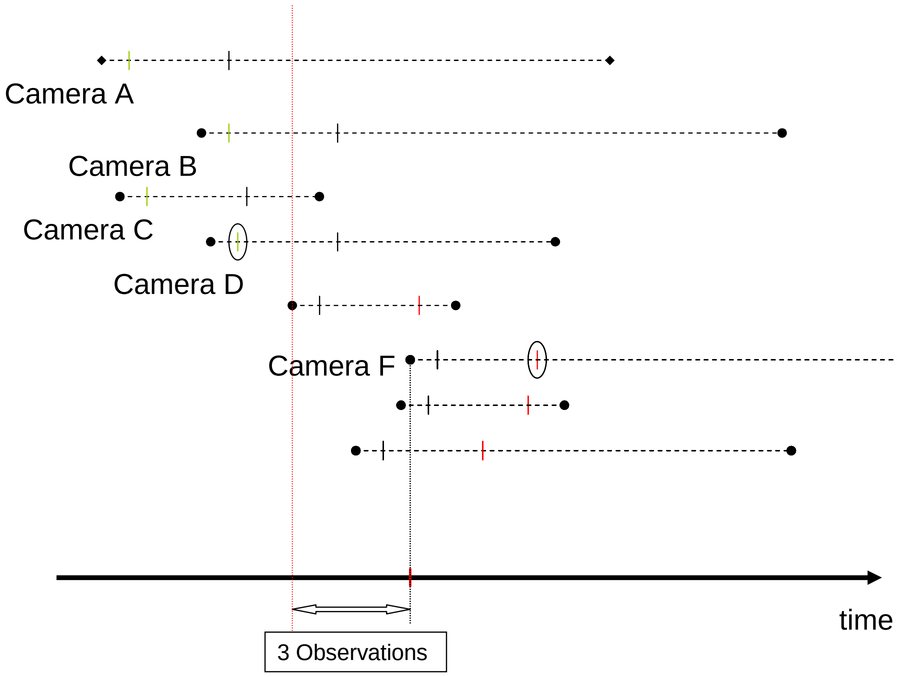
By the same token, we can calculate the share of products listed on the first, second, third, twentieth day. Thus we obtain an estimate of shop j's probability to list an item on it's first, second, tenth, etc. day of existence. Simply by aggregating these probabilities across shops we obtain the predictor of the number of shops on the item. So far this instrumentation strategy does not ensure however, that we always use the same number of products for calculating the shares needed for the instrument, and thus threatens the validity of its standard errors if we do not bootstrap them. Hence, in an effort to robustify our results we calculate an alternative version of the instrument, where we consider only the last 3 items that appeared on the market before item i did.

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 on today's markups and price dispersion. Table 2 presents the first stage regressions and show that the instrument is strong enough to explain the markets actual entry decisions depicted by the number of firms at each point in time of the product life cycles. 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. It is easy to see that the instrumental variables of interest are significantly different from 0 with a probability of more than 99.99%. Moreover they improve the predictive value of the model. The R^2 in the baseline regression without the instrumented number of firms (not shown in table) amounts to 0.3268. Adding our instrument for the number of firms raises the R^2 by 0.0155 in column (1) from 0.330 to 0.363. For the other columns even higher marginal R^2 can be computed. It should also be noted that the F-statistics against the null-hypothesis that the excluded instruments are irrelevant in the first-stage exceed the critical value of 10 substantially. With F-values well above 400 we can proof that our instruments are strong enough to explain the variation in the number of firms. For the following analysis we use columns(2) and (4) to calculate the predicted number of firms for the second stage regressions.

4 Results

Tables 3 and 4 show our basic results for the impact of market structure on markups. This basis specifications are parsimoniuous, they consider only the number of firms on the market

Figure 1: Instrument uses firm's listing behavior in earlier lifecycles



Notes: If we want to predict how many shops will list a product on the qth day after introduction we use the shops' general probability of listing a similar item that has entered the market in the six months before product j. To predict listing behavior for Camera D, we would use information on Cameras A, B, and C. However, to predict how many shops listed camera F on a specific day, we would use the information only from the cameras that saw light later than (and including) Camera D, provided they entered the market before F. Cameras A, B, C and models younger than camera F would be ignored.

- either linearly or in quadratic terms - and the product life cycle. Other product-specific influences are covered by a product fixed-effect. 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 effect of the number of firms on 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 reduce median markups by 0.22 and minimum markup by 0.54 percentage points. The reaction of the cheapest firm is significantly higher as compared to the reaction of the median firm, which might be explained by the high dynamics of online markets, where, in particular, the cheapest price is a focus of considerable attention of both consumers and firms.

If we instrument for the number of firms, we see a substantially higher negative effect: 10 additional retailers reduce the markup of the cheapest firm by 2.6 percentage points and the markup of the median firm by 0.88 percentage points. These figures are large in economic terms considering the standard deviation of the number of firms in our sample - 57 firms. It is not surprising that OLS is underestimating the true 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 reaction of the cheapest firm is considerably higher as compared to 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 as well as the median markup. Numerically, for the cheapest price, the negative influence of the number of firms ceases at 175, for the median markup with 264 firms. As the maximum number of firms in our sample is 203, we can safely assume that for most part of our sample, this negative relationship is a valid description.

Looking at the impact of the product cycle on markups, the picture is not entirely clear: in all 2SLS regressions markups grow over time only to go down at the end of the product life cycle. For the minimum markup the turning point is between 4 and 7 months (Columns 2 and 4) which is around the mean duration of a product life cycle of 5.5 months. For the median markup the turning point is with eight months slightly above the mean duration of the product life cycle.

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 somewhat negative relation between the number of firms and price dispersion, in the 2SLS results in Columns 2 and 4, we see a strong positive relationship. In the linear case, increasing the number of firms by 10 increases the coefficient of variation from 0.1 to 0.11. The situation is quite similar in the quadratic case (Column 4): apart from the first 6 firms, increasing the number of firms always leads to higher price dispersion.

The combined results on markups and price dispersion are only compatible with the model (Carlson and McAfee, 1983), i.e. a search theoretic model which accommodates two sources of heterogeneities by assuming a non-degenerate distribution of producers' marginal cost and heterogeneous visiting cost of the consumers. The other search theoretic models are not in line with our findings of a decreasing median markup, while the models of monopolistic competition predict a decreasing price dispersion, which is not in line with our findings.

4.1 Robustness

In this section we performed several robustness checks. First, we extend the model to check for different effects of market structure on markups over the product cycle. Then, we test the robustness of the basic results by using varying definitions of price dispersion, using other definitions of markups, i.e. including shipping costs into sales prices. Moreover, at the end we take account of the fact that some of the price offers attract less attention of potential buyers; we use click-weighted markups to control for this.

In Table 6 we investigate whether the profit-squeezing effect of a higher number of firms is the same in different phases of the product life cycle. To do this, we estimated the baseline model and added crossterms, interacting the number of firms with age (both linearly and quadratically). For ease of interpretation of the coefficients, in Figures 3 and 4 we also plotted how markups are predicted to depend on the number of firms, separately for different stages of the product life cycle.

In these plots, each line represents a product of certain age and we plotted the curve for products right after their introduction, and after 1, 2, 3, 6 and 9 months in the market respectively. To make the picture clearer, we concentrate for each phase of the life cycle on the typical situation concerning the number of firms.[9] Interestingly, our plots show a very consistent pattern. In Figure 3 we see the pattern for minimum markups. Apart from the 9th month we see a clear pattern: markups decline with more firms, regardless of the life cycle of the product. Moreover, over time, markups typically go down; this trend is quite visible in the first three months, but disappears later on. For the case of median markups in Figure 4, we can see a fairly similar pattern: at all times, more firms in the market means lower median markup. Again, there is an initial reduction of markups over the time of the life cycle, but this trend turns around after three months.

Our first robustness check in Table 7 concerns our definition of price dispersion. We experiment with different definitions: apart from the coefficient of variation we use the absolute price gap between the cheapest price and the second cheapest price, 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 they received. All these variations show a similar pattern: increasing number of firms is first reducing, then increasing price dispersion; in all cases, the turning point is below the average number of firms in the sample.

Further robustness checks concern the measurement of prices. Consumers typically pay the product price plus shipping costs. It is well-known that firms can follow specific price-setting strategies to set visible prices - the product price - very low and non-visible prices, like shipping costs, etc. relatively high (see Ellison and Ellison (2009)). In such a case, the total price including shipping costs should be used to calculate the markup of the firm. Unfortunately, we do not know "actual" shipping costs of the firms, therefore, we calculate an artificial markup: product price plus announced shipping costs minus wholesale price. As there are different shipping costs possible, we concentrate on those, which are mostly observed in the data, which are shipping costs to Germany when paying cash in advance. If firms can vary their announced shipping costs, they should also react to the market structure; i.e. the number of firms in the market. In Table 8 we show that, in fact, our qualitative results are fairly similar: both minimum and median markup decline with the number of firms and price

[9]We plot only in the region between the 38th and 62th percentile of the distribution concerning firm sizes to avoid extrapolation of the polynomials.

dispersion is increasing.

Finally, we investigate whether our results are influenced by the fact we treat all product offers symmetrically in our regressions. In particular, in questions of price dispersion researchers mistrust typically price offers which are way too high (cf. Baye et al. (2004)). This suggests to weigh price offers with the number of clicks they are receiving in order to give the low ranked - and maybe less reliable - price offers less weight. When we do this in Table 9, we see our main results unchanged.

5 Conclusions

In this paper we estimate the effect of market structure on market performance in e-commerce. As endogeneity of market structure and market performance might be an issue, we use the behavior of sellers in high frequency product life cycles to develop an instrumental variable for the number of firms in a market. To analyze the effect of market structure, we use data on 76 different digital cameras and find that an increase in the number of sellers in a market by 10 reduces the mark up of the price-leader by 2.6 percentage points and that of the median firm by 0.9 percentage points. While we also find negative correlations between market structure and performance using OLS regressions, our instrumental variables results - allowing a causal interpretation - are stronger. Moreover, we find a positive effect on the coefficient of variation.

When we differentiate market structure effects over the full life cycle of a product, we find a negative impact over all phases of the life cycle, with somewhat diminishing effects over time. Our results refer to e-tailing in the presence of a price-search engine with very narrow defined products. In such a situation, consumers have a very easy time to collect information about prices and reliability of the sellers. Still, it takes a large number of sellers and a relatively long time till mark-ups of firms dissipate.

The markup of the price-leader diminishes as well over the life cycle of the product. If we evaluate our results at sample means we can compare the competitive effect of more firms to the effect of time: having one more firm in the market reduces the mark up of the price leader by the same amount as three additional weeks in the product life cycle. In other words: a consumer will get the same price reduction if she waited for three more weeks or went to a market with one additional firm.

References

Bain, J.S., "Relation of profit rate to industry concentration: American manufacturing, 1936-1940," *The Quarterly Journal of Economics*, 1951, *65* (3), 293–324.

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., J. Morgan, and P. Scholten, "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.

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

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*, 1–25.

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, *63*, 440–449.

Davis, Peter, "Spatial Competition in Retail Markets: Movie Theaters," *Rand Journal of Economics*, 2006, *forthcoming.*

Dulleck, Uwe, Franz Hackl, Bernhard Weiss, and Rudolf Winter-Ebmer, "Buying Online: An Analysis of Shopbot Visitors," *German Economic Review*, 2011, *12* (4), 395–408.

Dunne, Timothy, Mark J. Roberts, and Larry 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 Ch.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, Paul 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, "Entry and Exit Behavior at a Shopbot: E-sellers as Kirznerian Entrepreneurs," 2008.

_ **and** _ , "Price, price dispersion and number of sellers at a low entry cost shopbot," *International journal of industrial organization*, 2008, *26* (2), 459–472.

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: Journal of the Econometric Society*, 1980, pp. 1575–1579.

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

Toivanen, Otto and Michael Waterson, "Empirical Research on Discrete Choice Game Theory Models of Entry: An Illustration," *European Economic Review*, 2000, *44*, 985–992.

_ **and** _ , "Market Structure and Entry: Where's the Beef?," *Rand Journal of Economics*, 2005, *36*, 680–699.

Varian, H.R., "A model of sales," *The American Economic Review*, 1980, pp. 651–659.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
average price in €	902.8	1353.8	99.1	7864.3	17209
medprice in €	892.9	1342.2	98	7990	17209
minimum price in €	810	1252.2	78	7084.6	17209
number of sellers (divided by 10)	10.6	5.7	0.1	20.3	17209
age in days	162.5	109.7	1	450	17209
wholesale price in €	727.2	1077.7	79	5801.4	17209
indicator: clicks for product i exist at t	0.9	0.3	0	1	17209
aggregate clicks at product i	28.3	39.3	0	646	17209
average clicks per shop offering product i	0.3	0.6	0	20.5	17209
markup of price-leader in %	4.6	7.8	-28.2	35.2	17209
median markup of firms offering product i in %	17.8	2.9	0	35.2	17209
markup of price-leader incl. shipping cost*) in €	7.5	6.8	-22.4	36.1	16827
median markup incl. shipping cost*) in €	19.3	3.6	-3.8	36.1	16827
coefficient of variation of the prices	0.1	0.2	0	5.7	17117
standard deviation of prices in €	65.7	176.7	0	6050.7	17117
coefficient of variation of prices incl. shipping cost*)	0.1	0.2	0	5.6	16502
absolute price gap between best price and second lowest price in €	10.5	25.8	0	515.9	17117

NOTES: The unit of observation is product i at time t (product-time panel). The time-variable are days since market introduction. The dataset covers 76 products. *) Markups and measures for price dispersion including shipping cost refers to the gross price including the fee for shipping cost the customers have to pay.

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

VARIABLES	(1) number of firms/10	(2) number of firm/10	(3) numfirm^2/100	(4) numfirm^2/100
instrumented	0.22***	0.81***	1.13***	12.08***
number of firms / 10	(0.011)	(0.027)	(0.219)	(0.536)
instrumented		-0.04***		-0.71***
(number of firms^2)/100		(0.002)		(0.032)
age (months)	1.62***	1.46***	33.04***	30.11***
	(0.027)	(0.027)	(0.526)	(0.534)
age^2	-0.10***	-0.09***	-2.06***	-1.88***
	(0.002)	(0.002)	(0.036)	(0.036)
constant	4.44***	3.35***	46.12***	25.80***
	(0.072)	(0.085)	(1.410)	(1.662)
observations	17,209	17,209	17,209	17,209
marginal R^2	0.0155	0.0361	0.0012	0.0214
products included	76	76	76	76
F test (all $u_i = 0$)	498.8	501.4	462.8	464.1
Log Likelihood	-42280	-42007	-93490	-93243
DoF (model)	78	79	78	79
rss	137182	132886	5.270e+07	5.120e+07

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; The R^2 of the baseline regression without the instrument (not included in the table) amounts to 0.3268.

Table 3: Minimum markup

VARIABLES	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
number of firms/10	-0.5355***	-2.6552***	-0.9712***	-3.0992***
	(0.011)	(0.130)	(0.030)	(0.128)
(number of firms^2)/100			0.0241***	0.0882***
			(0.002)	(0.009)
age (months)	-1.3595***	2.6895***	-1.3596***	0.4940***
	(0.039)	(0.254)	(0.039)	(0.150)
age^2	0.0396***	-0.2009***	0.0420***	-0.0619***
	(0.003)	(0.015)	(0.003)	(0.010)
Constant	15.9550***	26.8130***	16.9919***	24.7184***
	(0.109)	(0.684)	(0.128)	(0.355)
Observations	17,209	17,209	17,209	17,209
R^2	0.465		0.472	
products included	76	76	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Table 4: Median markup

VARIABLES	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
number of firms/10	-0.2163***	-0.8796***	-0.6792***	-0.9722***
	(0.006)	(0.054)	(0.017)	(0.066)
(number of firms^2)/100			0.0256***	0.0184***
			(0.001)	(0.005)
age (months)	-0.1398***	1.1271***	-0.1399***	0.6691***
	(0.023)	(0.106)	(0.022)	(0.077)
age^2	0.0030*	-0.0723***	0.0055***	-0.0433***
	(0.002)	(0.006)	(0.002)	(0.005)
Constant	20.7644***	24.1619***	21.8659***	23.7249***
	(0.063)	(0.284)	(0.072)	(0.182)
Observations	17,209	17,209	17,209	17,209
R^2	0.132		0.172	
products included	76	76	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Table 5: Coefficient of variation

VARIABLES	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
number of firms/10	-0.0024***	0.0106***	-0.0028*	-0.0024
	(0.001)	(0.003)	(0.001)	(0.006)
(number of firms^2)/100			0.0000	0.0019***
			(0.000)	(0.000)
age (months)	-0.0037**	-0.0286***	-0.0037**	-0.0693***
	(0.002)	(0.007)	(0.002)	(0.006)
age^2	0.0002	0.0017***	0.0002	0.0043***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.1274***	0.0604***	0.1283***	0.0366**
	(0.005)	(0.018)	(0.006)	(0.016)
Observations	17,117	17,117	17,117	17,117
R^2	0.004		0.004	
products included	76	76	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Table 6: Interacting the number of firms and the product life cycle

VARIABLES	(1) minimum markup	(2) median markup	(3) coefficient of variation
number of firms/10	-0.9129	-1.4065***	0.2393***
	(0.772)	(0.310)	(0.031)
(number of firms^2)/100	-0.0589	0.0394*	-0.0140***
	(0.051)	(0.020)	(0.002)
(number of firms/10) x age	1.1203***	0.3669***	-0.0708***
	(0.258)	(0.103)	(0.010)
(number of firms/10) x age^2	-0.0059	0.0207**	-0.0001
	(0.023)	(0.009)	(0.001)
(number of firms^2)/100 x age	-0.0268*	-0.0162***	0.0047***
	(0.015)	(0.006)	(0.001)
(number of firms^2)/100 x age^2	0.0001	-0.0005	-0.0001***
	(0.001)	(0.000)	(0.000)
age (months)	-8.5742***	-0.6309	-0.0003
	(1.070)	(0.429)	(0.042)
age^2	-0.0442	-0.2505***	0.0282***
	(0.150)	(0.060)	(0.006)
Constant	25.1128***	25.0261***	-0.2919***
	(1.167)	(0.468)	(0.054)
Observations	17,209	17,209	17,117
products included	76	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Table 7: Alternative versions of price dispersion

VARIABLES	(1) coef. of variation	(2) abs. price gap	(3) std. dev.	(4) clw coef. of variation
number of firms/10	-0.0024	-3.2428***	-20.0615***	-0.0095*
	(0.006)	(0.577)	(5.939)	(0.005)
(number of firms^2)/100	0.0019***	0.1590***	3.2730***	0.0018***
	(0.000)	(0.040)	(0.407)	(0.000)
age (months)	-0.0693***	-0.1040	-86.4986***	-0.0441***
	(0.006)	(0.595)	(6.124)	(0.005)
age^2	0.0043***	0.0013	5.2337***	0.0028***
	(0.000)	(0.039)	(0.397)	(0.000)
Constant	0.0366**	22.5227***	49.7434***	-0.0004
	(0.016)	(1.494)	(15.377)	(0.018)
Observations	17,117	17,117	17,117	14,948
products included	76	76	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Table 8: Markup and price dispersion including shipping costs

VARIABLES	(1) min markup	(2) med markup	(3) coeff. of variation
number of firms/10	-3.2448***	-1.4456***	0.0021
	(0.151)	(0.109)	(0.008)
(number of firms^2)/100	0.0579***	0.0104	0.0027***
	(0.010)	(0.007)	(0.001)
age (months)	1.5107***	1.7477***	-0.0984***
	(0.200)	(0.145)	(0.010)
age^2	-0.1154***	-0.1100***	0.0062***
	(0.013)	(0.009)	(0.001)
Constant	30.9010***	28.5859***	-0.0674**
	(0.575)	(0.417)	(0.034)
Observations	16,827	16,827	16,502
products included	76	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Table 9: Markup and price dispersion weighted by clicks

VARIABLES	(1) clickweighted minimum markup	(2) clickweighted median markup
number of firms/10	-2.9305***	-2.5049***
	(0.206)	(0.191)
(number of firms^2)/100	0.0598***	0.0873***
	(0.013)	(0.012)
age (months)	0.7782***	-0.1593
	(0.200)	(0.185)
age^2	-0.0808***	-0.0268**
	(0.013)	(0.012)
Constant	27.3837***	25.6425***
	(0.666)	(0.616)
Observations	15,711	15,711
products included	76	76

NOTES: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1;

Figure 2: Median markup plotted against the number of firms and age of product.

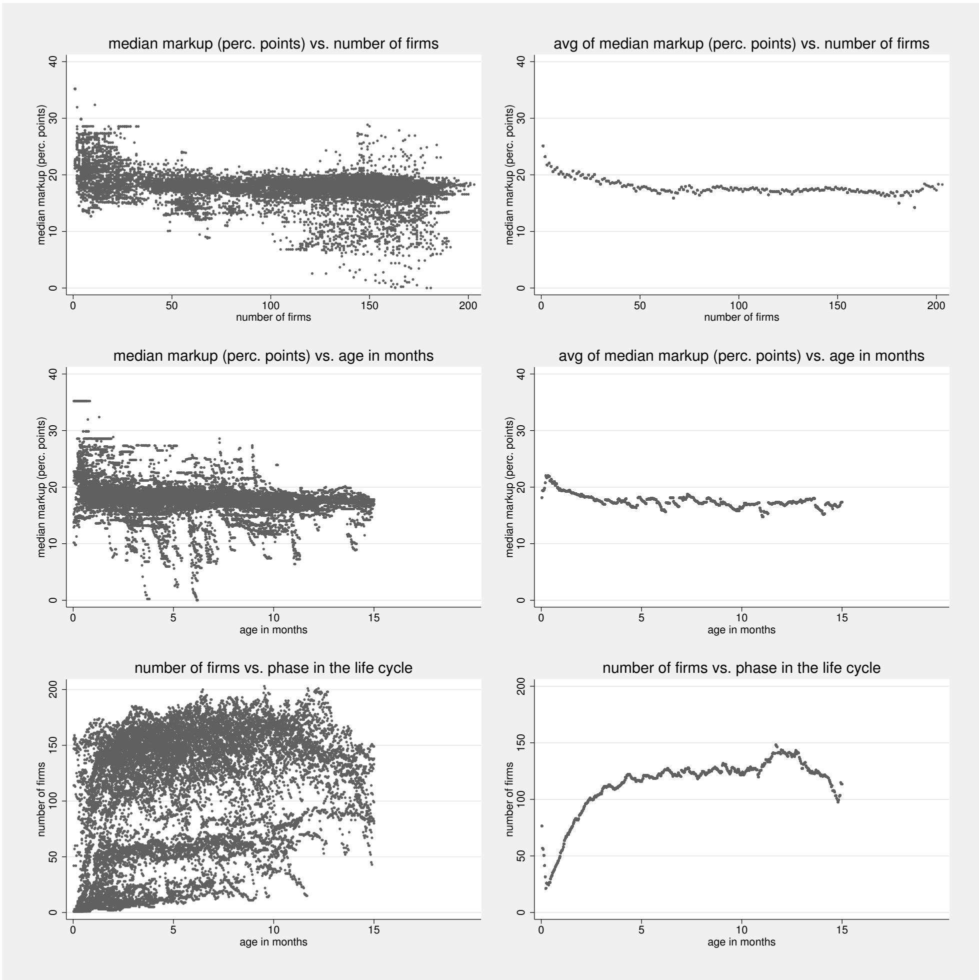


Figure 3: Minimum markup in different phases of the product life cycle

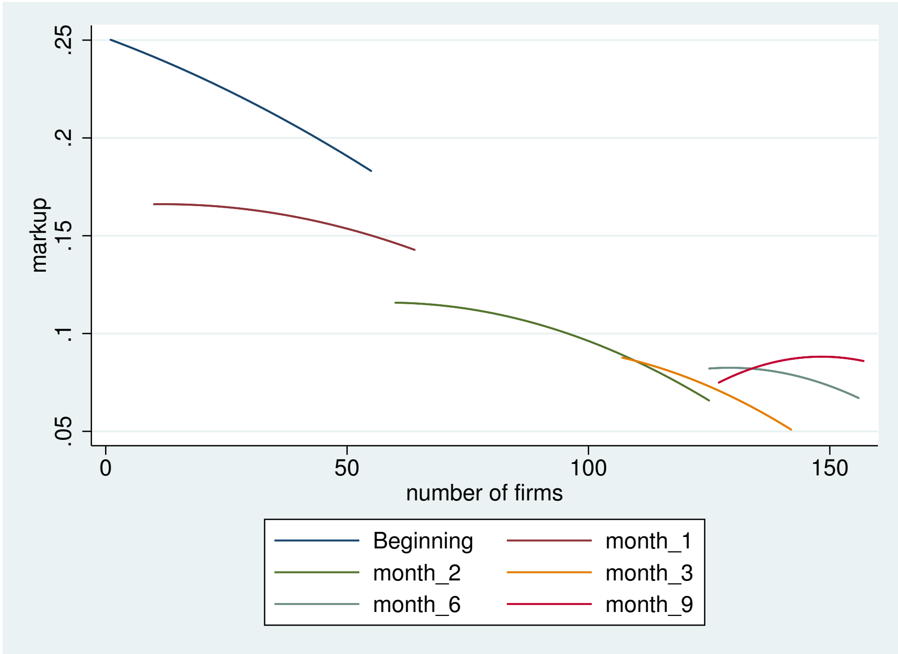


Figure 4: Median markup in different phases of the product life cycle