Received: 7 September 2020 | Revised: 1 January 2021 | Accepted: 24 January 2021
DOI: 10.1002/mde.3309

RESEARCH ARTICLE

WILEY

Successful retailer strategies in price comparison platforms

Franz Hackl[1] | Michael Hölzl-Leitner[1] | Rudolf Winter-Ebmer[1,2,3,4] | Christine Zulehner[5,6,7]

[1]Department of Economics, Johannes Kepler University Linz, Linz, Austria
[2]IHS, Vienna, Austria
[3]IZA, Bonn, Germany
[4]CEPR, London, UK
[5]Department of Economics, University of Vienna, Vienna, Austria
[6]Telecom ParisTech, Paris, France
[7]WIFO, Vienna, Austria

Correspondence
Franz Hackl, Department of Economics, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria.
Email: franz.hackl@jku.at

Funding information
Oesterreichische Nationalbank, Grant/Award Number: 16255

The choice of an appropriate e-commerce strategy for the listing in price comparison platforms (eBay, Amazon, and price search engines) is crucial for the survival of online stores in B2C e-commerce business. We use a comprehensive dataset from the Austrian price search engine geizhals.at to identify successful e-commerce strategies with regard to these listing decisions. An e-commerce strategy is a set of choices including the listing decision, availability decision, and decisions on a price path and shipping cost. We apply cluster analysis to identify the different strategies that have been used by online retailers. Using various success measures such as revenue, clicks, market share, and the survival of firms, as dependent variables in our regression analyses, we present causal evidence on the effectiveness of different e-commerce strategies.

JEL CLASSIFICATION
L81; L10

1 | INTRODUCTION

As price dispersion is a widespread phenomenon in the B2C e-commerce business,[1] consumers shopping online use internet platforms (such as eBay, Amazon,[2] AliExpress, or price search engines) to compare prices of different retailers. Hence, these platforms are central players in the information procurement of customers. For this reason, the communicated information about the web-shops' offers on these platforms is of central importance for the success of online retailers. This study identifies and investigates the strategic options of online stores on these platforms.

The set of strategic options is strongly determined by the platform design: due to technical restrictions and the need for a clear presentation, these platforms only offer standardized interfaces (variables) for their communication channel of retailers to customers. Analyzing these interfaces, it turns out that there is only a rather limited set of variables retailers can unilaterally decide about. To sell products online, e-commerce retailers can typically determine which products to list (already from the start of the product life cycle or later on), how to price the products over time, whether to make the products available immediately (to put them in stock) or offer longer delivery times (and order from the wholesaler after the order receipt from the customer), and how much to charge for shipping.

We define an e-commerce retailer's strategy as a specific combination of these four variables. Specifically, the retailer decides on the (i) listing of a product, (ii) its price path, (iii) its availability, and (iv) the shipping cost. These four components form the core of an e-tailer's strategy and are communicated to customers via online platforms.[3] Note that these four components are the core information on practically all information portals in e-commerce (Amazon, eBay, and various price search engines). We emphasize that all four components can

[1]For the relevance of price dispersion in e-commerce, see, for instance, Baye et al. (2004) or more recently Gorodnichenko et al. (2018) or Böheim et al. (2019).
[2]Note that Amazon is very successful in pursuing the strategy to become a prominent platform for online retailers under the brand name "Amazon Marketplace."
[3]Unlike in other taxonomies of retailers' strategies (e.g., Tokman et al. 2016 or Homburg et al. 2008), we do not use survey questions but rely on the actual information on the price comparison site.

be directly influenced by the e-tailer, in contrast to other components such as the price rank on the price comparison site, which can only be influenced by the e-tailer indirectly. E-commerce retailers may also apply different strategies for different products. Using *k*-means clustering, we find evidence for the existence of clearly distinguishable strategy clusters. By analyzing the frequency of the different web-shop strategies applied, it is also possible to identify different company types.

As the choice of the correct strategy can be crucial for the survival of online stores in the B2C e-commerce business, we present evidence on the effectiveness of different e-commerce strategies on success variables, as measured by clicks (revenues), market shares, and firm survival. Hence, we investigate the firms' success in their pricing and listing strategies in online platforms, in which all the relevant strategic choices of e-tailers in their search for customer attention are communicated via a dominant online platform.[4]

Note that in markets that are characterized by a strong competition for the attention of customers, clicks on an offer are important indicators for the success of web-shops.[5] Although, we will have our focus on price search engines, our results can be transferred to all other kinds of standardized platforms which juxtapose offers from different retailers.

Reference to a theoretical framework: E-commerce is mainly driven by Bertrand competition. For most of the traded products, however, consumers usually have to incur search costs due to firm heterogeneity. In this case, the theoretical literature on search has shown that some firms are able to increase prices relative to the competition, as discussed by Stahl (1989) and in the survey by Ellison (2016). The purpose of price comparison sites is to make prices highly visible and almost completely eliminate consumers' search costs.[6] To avoid this market situation, firms may react with non-price competition, such as competition on availability policies and shipping costs, and obfuscation (Ellison & Ellison, 2009; Gabaix & Laibson, 2006; Wilson, 2010) by taking actions to make price search more costly.

Note that the assumption that web-shops can freely decide on listing, price, availability, and shipping cost may not hold for all products and/or firms due to exogenous constraints: manufacturer and wholesalers might follow special supply policies (e.g., not to deliver those shops which offer below the manufacturer's recommended retail price). Different kinds of vertical restraints might have influence on the listing and price decisions (e.g., resale price maintenance and exclusive dealing). Although these constraints might be relevant for some products and firms, the competition law of many countries prohibits a systematic restriction of the entrepreneurial activity of retailers by an excessive market power of producers and wholesalers.[7] For this reason, we assume that these constraints are the exception rather than the rule and that our assumption, that retailers can unilaterally decide on their offer, is legitimate.

Data: For our empirical analysis, we use a random sample of about 5000 products offered by 780 retailers that were introduced on the price comparison platform geizhals.at.[8]

Research strategy: As there is practically no systematic scientific analysis on the optimal retailer strategies in online platforms beyond pure pricing strategies—neither theoretically nor empirically—we have chosen the following data-driven research design: (i) by applying k-means clustering based on the broad range of product offers by various retailers we want to find out, whether clearly distinguishable strategies can be identified at all in the e-commerce business. As different strategies might be applied for different products, we want to answer this question at the offer level. We found a rather convincing picture of clearly definable strategies for firms acting on price comparison sites. (ii) By using clicks, market shares, and firm survival as dependent variables in regression analysis, we identify which strategies are more successful than others. With IV regressions in which we use the choice on clustering variables for predecessor goods offered by the respective firms, we demonstrate the causal impact of strategy choices on firm success. (iii) As some firms apply certain strategies at the product level more often than others, it is possible to assign shops to different firm types each focusing on different strategies. (iv) Regression analysis allows us to differentiate between firm types which are successful from other firm types which show a high probability that they will not survive in the e-commerce business. One advantage of this data-driven approach is that it comes up with stylized facts on web-shops behavior which can also serve as a starting point for a more rigorous theoretical approach. It is, for instance, an interesting fact that the most frequent applied strategy is the worst performing managerial practice.

Results: The results of the cluster analysis show that e-tailers apply three different sets of strategies for offering products. We call the major strategy cluster *In-Stock-Offers*, *Permanent-Offers*, and *Long-Shot-Offers*.[9] *In-Stock-Offers* are listed for a short period of time, but the products are made immediately available at that time. They are sold at low prices with low shipping costs and low variability. *Permanent-Offers* are listed for a long time, but the products are not immediately available and are sold at intermediate prices and shipping costs. The price variability is low, but once prices are changed, the changes are large. *Long-Shot-Offers* are not listed for a long time nor are the products immediately available. These offers are characterized by the highest prices and shipping costs. Their prices are changed frequently

[4]We speak from dominant platforms if the viability of retailers depend on the referral of customers via the platform and the platform can enforce identical information on the web-shops homepage and the comparison platform. See, for instance, the MFN-clause of Amazon with publishers in the e-book market as very extreme form of dominance which even resulted in the EU case law "CASE AT.40153 E-book MFNs and related matters (Amazon)."

[5]Clicks (=referral requests to the retailers website) are *the* prerequisite for actual sales. Conversion rates are one of the important key figures in e-commerce and measure the percentage, how many clicks will lead to actual sales.

[6]For example, Baye et al. (2009) find that a firm enjoys a 60% jump in its clicks when it offers the lowest price at a price comparison site. Tang et al. (2010) show that, in general, the introduction of price comparison sites reduced book prices.

[7]In Austria, for instance, the competition authority tracks a series of complaints about discriminating delivery policies in e-commerce.

[8]Johnson et al. (2004) show that consumers do not search much on individual e-commerce sites. A price comparison site may thus cover a substantial amount of e-commerce. geizhals.at is a perfect example for such a market as it is the most important local price search engine. Practically, all Austrian online retailers have to list at this price search engine in order to be able to enter the online business at all.

[9]"Long-Shot" refers to a bet in which the chances of winning are small but the possible gains are large.

but only by small amounts. According to our firm success measures *In-Stock-Offers* are the most successful bids, followed by *Permanent-Offers*. *Permanent-Offers*, however, perform better than *In-Stock-Offers* in terms of the survival of firms. *Long-Shot-Offers* perform worst among all our performance measures. These results can also be confirmed at the firm type level—firm types with a high share of *In-Stock-Offers* perform on average better.

Strengths: (i) Our analysis is based on a large-scale dataset from a dominant price search engine covering essentially the entire national e-commerce market in Austria. (ii) Given the dominance of the price search engine and our full information from the website, the information that it transmits depicts the complete universe of strategic options for the relevant e-tailers. (iii) We can follow firms' strategic behavior over the complete life cycle of products. (iv) Although we show causal results for Austrian e-tailers, the external validity of this study is much larger. As e-tailing is becoming increasingly important in many sectors of the economy, evaluating the strategies of e-tailers in price comparison environments is important as well. Increased competition in Bertrand market structures forces firms to expand their strategies outside of simple price comparisons.

2 | RELATION TO THE LITERATURE

Our analysis provides empirical evidence for e-tailers' strategies in online platforms. There is hardly theoretical or empirical literature which has a broader focus on the effectiveness of different e-commerce strategies in online platforms taking into account the whole universe of strategic options of online retailers. Most of the contributions are dealing with particular aspects of the e-commerce trade. This section addresses the existing literature.

Gorodnichenko and Talavera (2017) or Gorodnichenko et al. (2018) conduct investigations of strategies of firms in pricecomparison sites, which concentrates on pricing itself. Bauer and Jannach (2018) propose a machine learning-based framework for estimating optimal prices in e-commerce. See also Schlosser et al. (2006) on the impact of website design investments on consumers' trusting beliefs and online purchase intentions. However, given the existence of price dispersion in e-commerce Ellison and Fisher-Ellison (2005) conclude that consumers are heterogeneously informed on the markets' price structure. Hence, in consumer decision making, not only the cheapest price but also a number of other factors play a role.

As described in more detail by Ellison (2016) and Ellison and Ellison (2018), firms offering products online have an incentive to obfuscate when consumers bear search costs and price comparison platforms explicitly reduce these search costs. Firms intend to increase consumers' search costs through, for example, add-on prices and, thus, charge prices higher than those under Bertrand competition.[10]

A closely related study to our analysis is that of Ellison and Snyder (2014). They investigate competition among firms participating in an online market and empirically assess the factors that drive firms to change prices. The analysis provides evidence for differences in pricing strategy decisions across firms. The authors embed their results in a framework for simulating counterfactual market settings and use the simulations to examine counterfactuals involving different mixes of firms based on pricing strategies. Whereas Ellison and Snyder (2014) concentrate on firms' pricing strategies in selling a commodity-type memory module, we extend the analysis to more products and further aspects beyond pricing, that is, listing decisions, availability, and shipping costs. Additionally, we investigate which strategies are more successful.

Based on data from price comparison websites, Cao et al. (2003) show that e-tailers can set higher prices and will have higher overall ratings for fulfillment satisfaction if they provide a satisfactory ordering process. On the other hand, reducing the prices do not positively affect satisfaction with the fulfillment process. Hence, the authors conclude that price competition is not a viable long-term strategy for online retailers.

Haynes and Thompson (2014) investigate sellers' entry behavior using data on digital cameras from Nextag.com. They analyze whether sellers employ hit-and-run strategies in line with the theoretical notion of the contestability of markets. Hit-and-run strategies correspond to shorter forays into the market at lower entry prices. The results of their estimations show that sellers with poor reputations and smaller sellers are more likely to favor a hit-and-run strategy than larger sellers with better reputations. They also find that former entrants induce a much larger price response from low reputation incumbents. This finding reflects the more intense competition for price-sensitive consumers who do not care about retailer reputation.

A key aspect of our analysis is listing decisions regarding new products. When, for example, Pauwels et al. (2004) investigate the effects of new products and sales promotions on firm value in the automobile industry, they rely on financial performance indicators such as revenue, firm income, and stock market performance. We do not use these measures of success, as most of our retailers are not listed on the stock market, and thus, data on financial performance is not available. Instead, we measure the effectiveness of firms' strategies using revenue, clicks, market share, and survival.

Frischmann et al. (2012) investigate the use of shipping costs as a strategic variable in e-commerce and distinguish between sellers charging no shipping costs and those charging relatively high shipping costs. These strategies are meant to target different consumer segments, particularly those with biased perceptions of price awareness.

Dinerstein et al. (2018) argue that the design of the platform has implications on sellers and buyers behavior. They argue that a direct comparison of seller listings for a given product reduced prices by 5% to 15%. Although their analysis is focused on situations where products vary only in price an quality, they conclude that "similar forces would be at play for other product attributes that can be changed in the short run."

[10]See McDonald and Wren (2018) for a discussion of an online search obfuscation effort by firms using multiple brands.

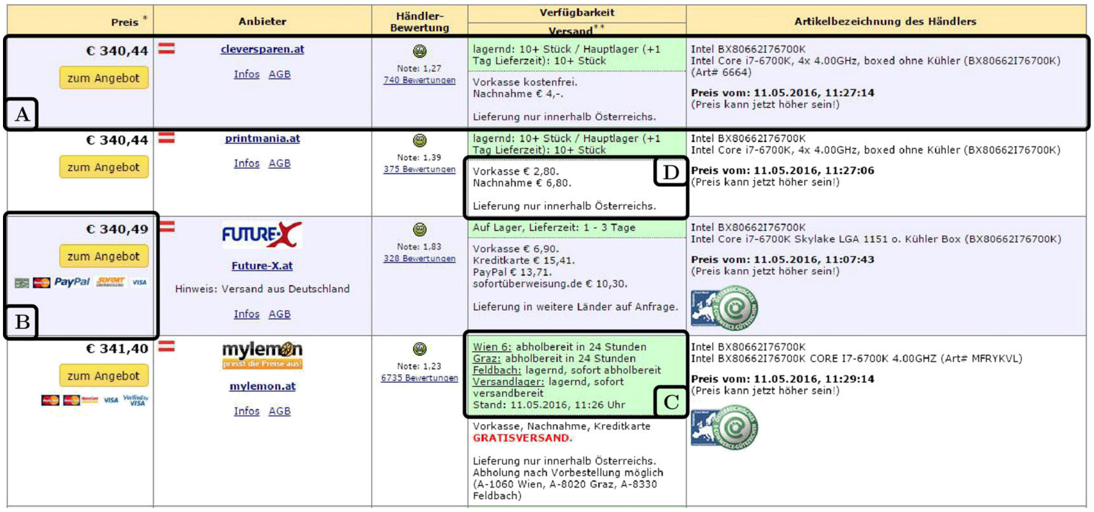

FIGURE 1 Snapshot of the geizhals.at website. Note that the strategic choices, which can be determined solely by the offering firms, reduce to four aspects only: (A) to list the product at all, (B) the price level, (C) whether products are immediately available at the shop (e.g., to have them in stock), and (D) the amount of shipping cost. Variables based on these four aspects will be used in the clustering procedure

3 | EMPIRICAL APPROACH

For our analysis, we use comprehensive data from Austria's largest price comparison portal, geizhals.at, covering the following product groups: IT hardware, software, games, video and photo devices and TV, phones, audio/hi-fi systems, films, household appliances, sporting goods, and drugstore items. According to information provided on geizhals.at, about 1000 retailers utilize the price comparison portal to offer 1,392,241 products for delivery in Austria (excluding Amazon Marketplace). According to the business model of geizhals.at, each retailer must pay a small fee each time an interested customer clicks on a link on the price search engine's webpage to access the e-tailer's webpage (=referral request). It is important to note that geizhals.at is the dominant price search engine in Austria. If an online shop wishes to enter the e-commerce business in Austria in one of the above-mentioned product groups, it is practically impossible to avoid listing its offers on the geizhals.at website. Thus, it is reasonable to assume that our data cover essentially the entire online Austrian market for most of these product groups. We use a sample of these retailers' offers for our analysis of e-commerce strategies in Austria.

3.1 | Identification of e-commerce strategies

E-commerce retailers may apply different strategies depending on the product offered. We define an e-commerce strategy as the set of choices that each retailer has to make for each product, and we will define the "offer level" as the observational unit of the strategy of retailer j for product i. An e-commerce strategy can consist of all decision parameters a firm can use during the life cycle of a product, provided that the strategy is also communicated to the consumer via the price comparison site.

Figure 1 shows a snapshot of an arbitrary hardware product offered by the price search engine. Analyzing this information shows that the set of strategic choices boils down to four essential categories: *(A) the listing decision* (whether the product should be added to the retailer's assortment at all); *(B) the price decision* (the pricing of the product over its life cycle and the target price rank on price comparison portals); *(C) the availability decision* (whether the product is held in storage even before orders arrive or ordered from a wholesaler after a customer places an order); and *(D) the shipping cost decision* (the pricing of shipping and whether this pricing implies a possible obfuscation strategy). These four categories are the foundation for our strategic variables. To characterize e-commerce strategies, we focus only on strategic variables that can be directly influenced by the retailer and directly communicated to the customer via the price comparison portal geizhals.at. No other category of strategic variables can be influenced directly by the retailer and varies across products. In that sense, we cover the entire universe of strategic decisions that a retailer must make in offering a product on geizhals.at.[11]

In order to identify e-commerce strategies, we use a k-means clustering algorithm based on the four strategic categories of listing, availability, pricing, and shipping cost decisions. The k-means clustering algorithm results in a set of meaningful and clearly distinguishable strategy groups.

Unfortunately, we do not observe the costs of different strategies or the firms' profits, in which the cost of different strategies should manifest directly. Hence, we use second-best measures which are supposed to be highly correlated with unobservable profits and which were previously used in e-commerce to approximate the success of a product or firm (Dulleck et al. 2011; Hackl et al. 2014; Smith &

[11]The retailer rating, which is also given on the geizhals.at website, is determined by the customers' evaluation and, thus, can be a long-term strategic element for firms. We will control for firm ratings when we consider success at the firm level.

Brynjolfsson, 2001). (i) *Number of clicks* is an indicator of customers' attention or the demand created by the offer. (ii) *Number of last-click-throughs (LCT)* (Bai & Luo, 2011; Park, 2017; Smith & Brynjolfsson, 2001) is typically seen as a better indicator of an actual sale because it identifies the last firm that a searching customer clicked on during a search on geizhals.at.[12] (iii) *Revenues by Clicks* are calculated as the offered price times the number of clicks. (iv) Finally, *Revenues by LCTs* give the offered price times the number of LCTs.

Note that in price comparison platforms, in which online shops compete for the attention of consumers, the number of clicks (=referral requests to the retailers web-shop) is one of the central key parameters with which success is measured. With clicks and LCTs (as proxies for actual sales), we cover the central inputs for the conversion rate reflecting the relationship between the amount of web-shop visitors and actual sales—an important and frequently used measure for the success in e-commerce. Our strategies are defined at the product level. Hence, firms might apply different strategies for different products. Clicks, LCTs, and our proxies for sales are measures for success at the offer level at which the actual managerial choice between different strategies is done.

As the firm behavior is quite heterogeneous in the usage of different strategies, we can identify specific types of firms (*firm pools*) which differ in their strategy usage. We find evidence for the existence of both clear-cut and mixed retailer strategies. Analyzing at the firm level has the advantage that we cannot only use the above-mentioned click-related success measures but also the survival of e-commerce firms as a yardstick for the firms' performance. This is particularly important for those who are skeptical whether click-related variables are good measures for firm success: it is sometimes argued that firm survival is the ultimate measure of firm performance. This is especially true for the hotly contested e-commerce market. At least in our sample, after 2 years, about 30% of all companies were no longer listed on the price comparison platform.

Although we cannot observe cost of strategies and profits directly, we have a series of very good proxies which should be highly correlated with the success (profits) in order to be able to make reliable statements about the choice of strategy.

3.2 | Data

We use data for new products in geizhals.at to understand firm strategies over the full life cycle of products. We restrict our data to a random sample of about 5% of all products introduced in 2010.[13] The following criteria have been applied in the composition of the dataset: (i) although geizhals.at is available in other countries as well (e.g., Germany, the United Kingdom, and Poland), we only consider the Austrian market. The website geizhals.at only has a dominant position in e-commerce in Austria. Moreover, the default view of the website shows only the Austrian market. This restriction leads to a representative sample of Austria's e-commerce. (ii) We use an inflow sample, only taking into account products that were introduced during 2010. The usage of an inflow sample prevents biased results in favor of long-running products. We use a full year of inflow to prevent biases caused by seasonal effects. (iii) The year 2010 guarantees a sample of new products for which we can observe e-commerce strategies over the *entire* product life cycle. (iv) Products must have been first introduced in Austria. We do not want to bias our findings on e-commerce strategies by considering products already introduced in other geographical markets. (v) Products in the sample must have a minimum of 50 clicks (for Austrian retailers) and a minimum product life cycle of 100 days. (vi) Each product must be offered by at least two Austrian retailers. (vii) Furthermore, we eliminate outliers at the offer level: offers exceeding five times the median price of a product, offers exceeding five times the median shipping costs of a product, and offers with shipping costs above 1000 euros are excluded. In doing so, we eliminate clear input typos. In Table 1 a description of the used variables can be found. Table B.11 in the (Web-Appendix) shows a description of further variables not included in this list.

After applying these restrictions, we obtain 149,862 observations at the offer level, covering 4888 products offered by 780 retailers. Thus, each product is offered by 30 retailers on average. The first section of Table 2 provides descriptive statistics for the variables that are used for the *k*-means clustering. Moreover, Table 2 includes success variables, which are used to evaluate the absolute and relative success of different e-commerce strategies at the product level.

The start of a product's life cycle is easy to define, but the end may be less clear because firms may still offer the product even though demand (clicks) has already disappeared. Thus, we define the end of the life cycle as the point when the 97th percentile of clicks on the product has been reached. For products with very high demand,[14] we set a maximum of 500 clicks as the cutoff to determine the end of the product life cycle.

4 | DESCRIPTION OF E-COMMERCE STRATEGIES

4.1 | Clustering method

To identify different strategies at the offer level, we use a clustering approach. The *k*-means clustering method partitions a dataset into *k* partitions such that the sum of squared deviations from the cluster means (*J*) is minimal (Lloyd, 1982):

[12] As we can distinguish different customers at https://www.geizhals.at using a cookie identifier, we can determine each customers' search episode(s) as a sequence of clicks (=referral requests) from a specific cookie to different e-tailers. A single consumer can have multiple search episodes. We define the LCT as the last click within each search episode, and we assume that it is more probable that the customer made a purchase at this last shop than at any other shop. LCTs are better proxies for actual sales, but they are not perfect (e.g., a cookie identifier may correspond to more than one person, a cookie identifier may be blocked, or a consumer may not make a purchase at the last referral request).

[13] Geizhals.at introduced 101,906 products throughout 2010.

[14] We define a product as being "in very high demand" if the number of clicks in the last three percentiles of its life cycle exceeds is greater than 500.

TABLE 1 Description of variables

Clustering variables	
Availability Percentage	Number of days product is in stock of retailer (relative to the number of days product is offered by the retailer).
Beginning of Offer	Time when retailer offered a product for the first time. Measured in days after the first occurrence of the product on geizhals.at (in days from start of PLC).
End of Offer	Time when retailer removes product from the offered assortment. Measured in days before the product disappears from geizhals.at because no retailer is offering the product anymore (in days till end of PLC).
Listing Percentage	Time product offered by the retailer relative to the duration of the whole product life cycle.
Daily Price Changes	Number of total price changes relative to the number of days the product is offered by the retailer. Price changes are observed at a daily base, so the maximum number of daily price changes is 1. A change of the listing decision for a product (offering or not offering the product) by a retailer is treated like a price change, too.
Planned Price Rank	Average listing rank after a price change.
Coefficient of Variation of Absolute Price	Coefficient of variation of the absolute price of the offer.
Absolute Shipping Costs	Average shipping costs for the offer using payment before shipping.
Success variables	
Click Share	Number of clicks on the retailer's offer relative to the total number of clicks on the product (in %).
Number of Clicks	Number of clicks on the retailer's offer.
Number of LCT	Number of Last-Click-Through clicks on the retailer's offer.
Revenue by Clicks	Revenue in terms of clicks. Number of clicks times the average price offered by the retailer.
Revenue by LCT	Revenue in terms of Last-Click-Through. Number of LCT-clicks times the average price offered by the retailer.
Firm characteristics	
Pick-Up Possibility	Retailer offers the possibility to pick-up products in a store.
Product Mix (HHI)	Indicator for the concentration of the product range of a retailer based on spread of offers among different product categories. High value means concentrated assortment whereas low value indicates a wide range of product types offered.
Firm Rating	Rating of the retailer by users of geizhals.at. 1 means very good, whereas 5 means not very poor performance of the retailer.
Total Clicks on Firm	Total number of clicks on retailer during the year 2010.
No. of Products Offered	Total number of products offered by the retailer during the year 2010.
Average Relative Price	Average relative price (compared to the average product price) of all offers by the retailer.
Product characteristics	
Median Absolute Price	Median price of all offers of the product.
PLC Duration	Full duration of the product life cycle of a product in days.
No. of Offering Firms	Average number of retailers, offering the product.
Price Density	Density of prices for one product. Calculated as (maximum price − minimum price)/number of offering retailers.
Total Clicks on Product	Total number of clicks on the product during the whole product life cycle.

$$J = \sum_{i=1}^{k} \sum_{x_j \in S_i} ||x_j - \mu_i||^2,$$

using data points x_j with means μ_i of clusters S_i. This Euclidean distance operation assigns each data point to the next cluster mean. We use normalized data points between 0 and 1.

Data points are variables which describe an e-commerce strategy and which, therefore, all can be attributed to listing, availability, price, and shipping cost decisions. The selection process of variables takes into account the following considerations.

(i) We use variables that can be determined by the offering retailer itself and, thus, are not driven by rivals' actions. (ii) We avoid variables with high multicollinearity. (iii) We prioritize variables that are immediately observable by customers.

This resulted in the following list of variables: *Listing Percentage* is the percentage of the product life cycle that the product was on offer. *Beginning of Offer* and *End of Offer* are also used to characterize the listing decision within the clustering procedure. *Average Planned Price Rank* serves as an indicator of the firms' target price rank.[15] *Number of Daily Price Changes* is an indicator of a retailer's price activity, and *Coefficient of Variation of Absolute Price* is an indicator of the extent to which prices

[15] We refrain from using the actual observed price rank because this indicator is determined by market behavior.

TABLE 2 Descriptives (means) for the strategy clusters

	All	In-Stock-Offers	Permanent-Offers	Long-Shot-Offers
Clustering variables				
Availability Percentage	21.8	**86.9**	4.6	*2.1*
End of Offer (in days till end of PLC)	324.0	316.6	*113.0*	**453.0**
Listing Percentage	33.3	33.1	**65.3**	*14.4*
Beginning of Offer (in days from start of PLC)	222.0	245.2	*80.2*	**295.8**
Daily Price Changes	0.153	0.139	*0.138*	**0.168**
Planned Price Rank	11.810	*11.070*	11.600	**12.270**
Coef. of Variation of Absolute Price	0.085	0.080	**0.121**	*0.066*
Absolute Shipping Costs	7.745	*7.496*	7.768	**7.845**
Success variables				
Click Share (in %)	3.180	**7.056**	4.097	*0.857*
Number of Clicks	17.240	**45.530**	18.660	*3.423*
Number of LCT	1.247	**3.153**	1.429	*0.264*
Revenue	6060	**12,781**	8269	*1662*
Observations	149,862	33,479	43,414	*72,969*
In percent	100.0	22.3	29.0	48.7

Note: The observational unit is the firm-product level. Highest (lowest) values are marked bold (italics).

have changed. These three variables are used to represent the pricing decision. The availability of the product is captured by the percentage of listing days that a product is in stock (*Availability Percentage*). Finally, the shipping cost decision is covered by *Absolute Shipping Costs*.[16]

The *k*-means clustering algorithm requires an ex ante definition of the number of clustered groups *k*. We use the following statistical measures to determine the optimal number of groups *k*. (i) The kink in the within-sum-of-squares is a measure of the within-group variation and declines for each additional group added. (ii) The proportional-reduction-of-error shows how the within-group variation is reduced by using *k* groups instead of $k - 1$ groups. (iii) The Calinski–Harabasz pseudo-*F* is another measure of the quality of clustering.

Figure 2 shows that the results for all three measures uniquely indicate that $k = 3$ is optimal. As a result, we obtain three e-commerce strategy groups at the offer level by applying *k*-means clustering with $k = 3$.

4.2 | Clustering outcomes

Using k-means clustering, we obtain three clusters,[17] which we call *In-Stock-Offers*, *Permanent-Offers*, and *Long-Shot-Offers*. We deduce the descriptions of these groups from the major clustering variables.[18]

The *In-Stock-Offers* cluster comprises around 22% of all offers. These offers are available for 87% of the listing time. Although they are only offered for about one third of the entire product life cycle, once they are listed, they remain in stock. This high availability is in stark contrast to that in the other clusters, which show availability of less than 5%. Moreover, the prices and shipping costs are lowest in this cluster, and the variability of prices is low as well. It may be that these firms order products in larger quantities and offer them steadily and cheaply from their shelf.

We call the second cluster *Permanent-Offers*; this cluster comprises around 29% of all offers. The main determinant of these offers is long listing behavior; a product is listed most of the time, but it is not kept in stock. Moreover, this cluster has intermediate prices and shipping costs. Prices are not changed often, but when they are changed, the amounts of the prices changes are large. These offers may be seen as firms wanting to list a product without intending to keep it in stock or seeing a necessity for frequent price changes.

Finally, we call the third cluster *Long-Shot-Offers*. Almost 50% of all offers belong to this group. These offers are characterized by the highest prices and shipping costs. The products are generally neither held in stock nor listed for a very long time. Prices are changed very frequently but only by small amounts. Rent-skimming behavior (Varian, 1980) might explain these offers. E-tailers assume that their client base comprises informed and uninformed customers. Informed customers have low search costs and buy from the cheapest website. Offers in the Long-Shot-Offers cluster, however, are addressed towards uninformed customers with higher search cost, who buy both via the referral request at the geizhals.at website and

[16]The robustness checks in Section 6 show that the clustering results do not change if different compositions of the clustering variables are used.
[17]Our results are extremely robust with regard to the randomly chosen seed value which defines the initial cluster assignment of each observation in the iterative *k*-means clustering procedure. The final cluster affiliation does not change with the variation of these seed values.
[18]For descriptive statistics related to the three resulting groups, see Table 2.

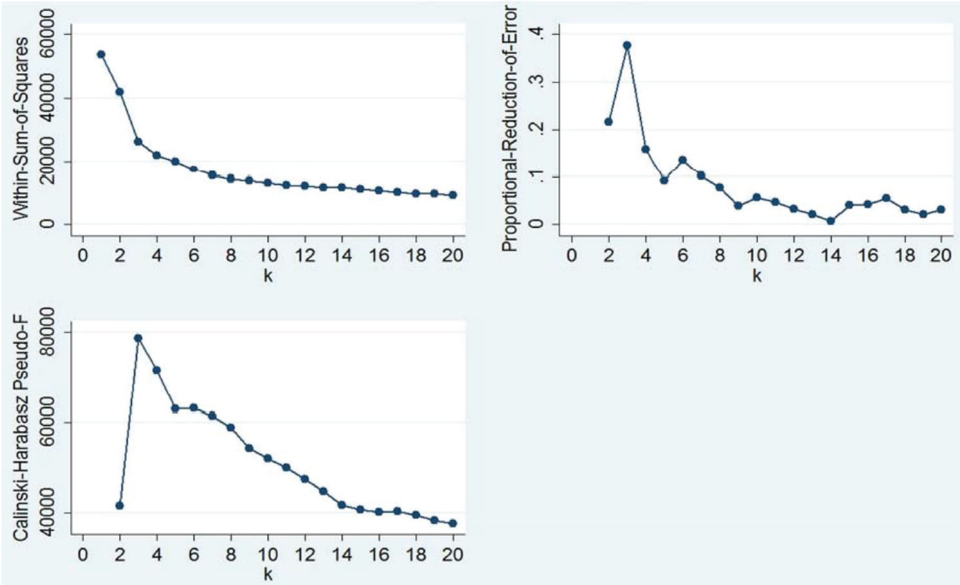

FIGURE 2 Quality measures for the clustering. Different quality indicators for the *k*-means clustering procedure are depicted. The variable *k* on the abscissa refers to the potential number of clusters. Note the kink in the within-sum-of-squares and the maximum in the Calinski–Harabasz pseudo-*F* as well as the proportional-reduction-of-errors for the amount of three clusters

directly from the firms' websites without contacting a price search engine.[19]

Figure 3 gives a schematic presentation of the main components of these three strategies. Additional information, particularly information on outcomes and market-determined characteristics of these clusters, is given in Table 3.

The standardized discrimination function loadings show that the listing decision and availability make the largest contribution to the offer clusters.

5 | SUCCESS OF E-COMMERCE STRATEGIES

We next consider the profitability of the e-commerce strategies identified by our cluster analysis. We analyze the success of these strategies in two steps. First, we concentrate on the offer level and proxy success using demand and revenue. Second, we aggregate our data at the firm level and measure success using firm survival.

5.1 | Offer level

Our first analysis checks which of the In-Stock-Offers, Permanent-Offers, and Long-Shot-Offers clusters are more successful at the offer level. Unfortunately, we cannot directly measure the profitability of a strategy, as the costs of specific strategies and actual purchases are not directly measurable. Instead, we use (see also the Section 3 on Identification) (i) *Number of clicks*, (ii) *Number of LCT*, (iii) *Revenues by Clicks*, and finally, (iv) *Revenues by LCTs*. Although we have no direct measure for profits, it is important to know how to attract demand and generate revenue. Hence, from the perspective of a web-shop, click-related success variables are extremely important key parameters. Additionally, e-tailer and product fixed effects help us account for time-invariant unobserved factors that influence cost and demand in our regressions.

Table 4 shows the results of ordinary least squares and fixed effects regressions for each of the success variables. For the e-commerce strategy clusters, we use dummy variables equal to 1 if the offer belongs to the respective cluster and 0 otherwise. The *In-Stock Offers* cluster acts as the base group for all regressions. Column (1) shows the results without any specific controls. Column (2) uses e-tailer fixed effects to control for unobserved heterogeneity among the offering retailers. Finally, Column (3) adds product fixed effects to control for product-specific heterogeneity. The last specification with e-tailer and product fixed effects is the most appropriate specification, because we are interested in the success of different strategies for the same e-tailer and product, accounting for time-invariant cost and demand heterogeneity.

The results in Table 4 show the strategy ranking in terms of demand and revenue. With respect to demand (i.e., number of clicks and LCT), we find that the *In-Stock-Offers* cluster is always the most successful, followed by *Permanent-Offers* and *Long-Shot-Offers*, which is the least successful cluster. When considering revenues, we no longer find statistical differences between the *Permanent-Offers* and *In-Stock-Offers* cluster when we control for unobserved firm and product heterogeneity. This pattern can be explained by the fact that the *Permanent-Offers* cluster predominantly consists of more expensive products (the mean absolute price is 393 euros for *Permanent-Offers*, whereas that for *In-Stock-Offers* is 342 euros). Thus, Column (2) implies a positive, revenue-increasing effect of *Permanent-Offers*. In any case, the *Long-Shot-Offers* cluster performs the worst.

The quantitative effect of using a different strategy is non-negligible. Looking at our preferred specification with e-tailer and

[19]Legal contracts between e-tailers and geizhals.at commit retailers to list identical prices in the price search engine and on their websites.

FIGURE 3 Schematic representation of cluster descriptives. The figure shows schematic representations of descriptives from Table 2 for the clusters In-Stock-Offers, Permanent-Offers, and Long-Shot-Offers. The illustration of the variables is depicted in proportion to their true means

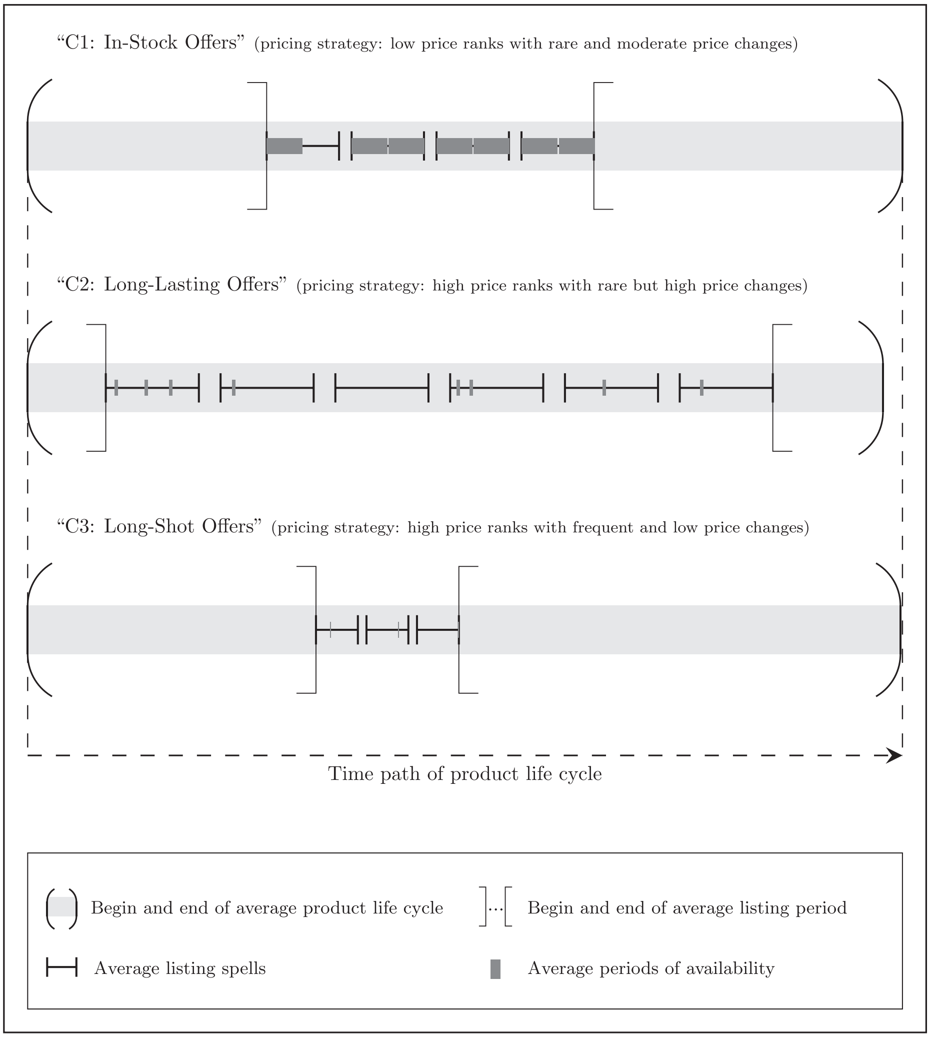

TABLE 3 Further descriptives (means) for the e-commerce strategy clusters

	All	In-Stock-Offers	Permanent-Offers	Long-Shot-Offers
Further descriptives				
Availability at First Offering Day	0.166	**0.635**	*0.030*	0.032
No. of Availability Changes	10.220	12.900	**16.240**	*5.417*
No. of Days Offered	290.000	306.700	**541.600**	*132.600*
No. of Listing Changes	8.512	7.853	**12.600**	*6.380*
Bestprice Percentage	0.082	**0.129**	0.070	*0.068*
Losses Until Reaction	3.565	3.461	*2.473*	**4.263**
No. of Price Changes	38.240	34.900	**70.060**	*16.660*
Rank at First Offering Day	10.590	10.340	*9.021*	**11.650**
Average Relative Price	1.011	*0.988*	1.005	**1.026**
Relative Price at First Day	1.017	*1.005*	*1.005*	**1.029**
Relative Price at Last Day	1.015	*0.984*	1.013	**1.031**
Average Relative Price Rank	0.563	*0.515*	0.549	**0.592**
Top-10 Percentage	0.534	**0.561**	0.532	*0.522*
Coef. of Variation of rel. Price	0.057	0.062	**0.066**	*0.048*
Var. Coef. of rel. Rank	0.285	0.326	**0.356**	*0.224*
Observations	149,862	33,479	43,414	72,969
In percent	100.00	22.34	28.97	48.69

Note: Values in the table represent means of the respective variables. The observational unit is the firm-product level. Highest values are marked bold (italics).

TABLE 4 Success of different clusters at the offer level

	(1)	(2)	(3)
Number of Clicks			
Permanent-Offers	−26.87***	−14.84***	−13.35***
	(0.716)	(1.008)	(1.018)
Long-Shot-Offers	−42.11***	−24.72***	−23.35***
	(0.650)	(0.939)	(0.943)
Constant	45.53***	33.58***	102.0***
	(0.538)	(0.748)	(4.755)
R^2	0.027	0.144	0.234
Number of Last-Click-Throughs			
Permanent-Offers	−1.724***	−0.823***	−0.787***
	(0.0568)	(0.0809)	(0.0818)
Long-Shot-Offers	−2.889***	−1.623***	−1.511***
	(0.0515)	(0.0753)	(0.0758)
Constant	3.153***	2.275***	8.924***
	(0.0426)	(0.0600)	(0.382)
R^2	0.021	0.117	0.208
Revenues by Clicks			
Permanent-Offers	−4511***	1548***	362.9
	(308.9)	(440.2)	(428.1)
Long-Shot-Offers	−11,119***	−3192***	−4756***
	(280.4)	(410.1)	(396.8)
Constant	12,781***	7166***	31,221***
	(232.1)	(326.7)	(2000)
R^2	0.011	0.108	0.260
Revenues by Last-Click-Throughs			
Permanent-Offers	−340.9***	100.4**	−6.995
	(27.08)	(39.03)	(37.99)
Long-Shot-Offers	−881.5***	−297.9***	−385.5***
	(24.57)	(36.36)	(35.21)
Constant	1022***	610.3***	3300***
	(20.34)	(28.97)	(177.5)
R^2	0.010	0.085	0.240
Observations	149,862	149,862	149,862
E-tailer Fixed Effects		X	X
Number of Retailers		780	780
Product Fixed Effects			X
Number of Products			4888

Note: In all regressions, In-Stock-Offers represent the base scenario. Standard errors are in parentheses.
**$p < 0.05$. ***$p < 0.01$.

product fixed effects (Column (3)), switching from an *In-Stock-Offer* to a *Long-Shot-Offer* reduces the number of clicks by 51%, and the amount of LCTs by 48%; revenues drop by 37%.[20] Choosing a different e-tailing strategy has far-reaching consequences on customer attention to products, and, to the extent that this attention is also converted into actual purchase, the consequences are even larger.

As robustness check, we show in Section 6.1 that IV regressions controlling for potential endogeneity of the cluster types confirm the causal interpretation of our results. Moreover, we demonstrate in Section 6.4 the robustness of our results, if we control for the quality of the cluster assignments by using silhouette coefficients of the observations as weights in the regressions.

5.2 | Firm level

Success of e-commerce strategies: In the second step of our analysis, we consider firms. We aggregate the data at the firm level and construct firms' shares of *In-Stock-Offers*, *Permanent-Offers*, and *Long-Shot-Offers*. These variables are related to firms' survival in 2012, which we again interpret as a measure of profitability. We use the dummy variable *Still Alive in 2012* as an indicator for success. Both in Austria and globally, e-commerce is characterized by a high number of market entries and exits. Of the 780 retailers in our dataset that we observe from 2010 on, only 535 are still active in 2012 (i.e., the dummy variable *Still Alive in 2012* is equal to one). Thus, 245 Austrian e-commerce retailers went out of business over this time period. This indicator is important as it allows for conclusions about the profitability of these retailers. Whereas the other success variables relate to revenues or induced demand, the indicator of firm survival allows more direct inference regarding the profits of firms. We augment these regressions with additional explanatory variables, such as pick-up possibilities, product mixes, firm ratings, and the number of products offered by firms.

Our estimation results are given in Columns (1) and (2) of Table 5. We find that firms with high shares of *Permanent-Offers* are more likely to stay alive than firms with high shares of *In-Stock-Offers*. Clearly, this strategy ranking differs from our earlier results on product-specific offers related to demand or attention. The ranking may have changed for two reasons. First, *Permanent-Offers* are more often used for expensive products with possibly higher mark-ups. Second, firms in the cluster *Permanent-Offers* rarely hold inventory but rather sell their products directly via a wholesale firm. Creating attention and demand for low-priced products is not sufficient for firm survival in this cluster. Our results for revenues in Table 4 also show no difference in revenues between *Permanent-Offers* and *In-Stock-Offers*. Finally, higher inventory costs may make *In-Stock-Offers* less profitable.

As before, firms with high shares of *Long-Shot-Offers* perform worst. As firm survival is related to business coming via referral request from the price comparison site geizhals.at, as well as demand from customers, who do not use a price comparison website, this survival analysis is also informative with respect to the rent-skimming strategy mentioned above: as survival at geizhals.at is correlated to actual survival of the firm, rent skimming by addressing customers going directly to the high-priced web-shop of the firm

[20]Percentage values are based on the average number of clicks, LCTs and revenues of the base group, which is *In-Stock-Offers*. See Table 2 for the corresponding values.

TABLE 5 Success on firm level: using strategy shares and firm types

	(1)	(2)	(3)	(4)
	Dependant variable: Still Alive in 2012			
Share Permanent-Offers	0.295***	0.240***		
	(0.0520)	(0.0533)		
Share Long-Shot-Offers	−0.229***	−0.239***		
	(0.0478)	(0.0464)		
F2: Specialized-Suppliers			0.265***	0.202***
			(0.0569)	(0.0611)
F3: Long-Shot Firms			−0.185***	−0.201***
			(0.0502)	(0.0484)
F4: Power-Sellers			0.148***	0.134**
			(0.0554)	(0.0570)
F5: Short-Term Suppliers			−0.123**	−0.105*
			(0.0555)	(0.0539)
F6: Large-Department-Stores			0.0273	0.0112
			(0.0503)	(0.0494)
F7: Mixed-Strategy-Type			0.203***	0.221***
			(0.0533)	(0.0554)
Pick-Up Possibility	−0.105***	−0.0851***	−0.103***	−0.0837***
	(0.0323)	(0.0322)	(0.0325)	(0.0321)
Product Mix (HHI/100,000)	0.476	0.972	0.767	1.321
	(0.782)	(0.796)	(0.790)	(0.806)
Firm Rating	−0.0486*	−0.0321	−0.0435*	−0.0260
	(0.0259)	(0.0264)	(0.0260)	(0.0266)
No. of Products Offered by Firm/100,000	0.138***	0.234***	0.146**	0.244***
	(0.0513)	(0.0691)	(0.0601)	(0.0802)
Constant	0.882***	0.915***	0.853***	0.880***
	(0.0556)	(0.108)	(0.0587)	(0.111)
Product Category Fixed Effects		X		X
R^2	0.134	0.190	0.131	0.193
Observations	774	774	774	774

Note: Dependant variable: Still Alive in 2012. Estimation method: Linear probability model. “Share In-Stock-Offers” and firm type “F1: In-Stock-Firms” represent the base group. A dummy for imputed firm ratings is included. The product fixed effects refer to the product categories used by geizhals.at to which the respective product range of a company predominantly belongs. Standard errors are in parentheses.
*$p < 0.1$.
**$p < 0.05$.
***$p < 0.01$.

(without comparing prices at geizhals.at first) does not seem to pay off—these firms go out of business earlier.

Our control variables perform according to expectations. Larger firms (with more products) live longer, as do those with better consumer ratings. The distribution of the product mix is not important for survival, whereas firms with no pick-up possibility (i.e., firms with no brick-and-mortar stores) live longer.

Success of firm pools: Thus far, we have characterized firms based on the percentages of strategies chosen. However, firms may choose specific e-commerce strategies for specific products. A more nuanced picture emerges if we take these strategic elements into account when characterizing firm types.

Thus, we implemented the following algorithm to assign retailers to firm-strategy pools. (a) Assign a retailer to the pool F1, F2, or F3 if more than 70% of offers fall in the respective cluster (e.g., *F1: In-Stock-Firms* make more than 70% of their offers as *In-Stock-Offers*). (b) If two strategies combine to make up more than 70% of offers, we assign firms to the strategy pools F4, F5, and F6, accordingly: F4 is *In-*

TABLE 6 Description of firm types

Name + Definition: >70%	No. of e-tailers	Description and interpretation
F1: In-Stock-Firms In-Stock-Offers	230	High percentage of immediately available offers; most products in stock; low offered price with small dispersion and low price rank; low shipping cost; number of products offered is low; specialized on few product categories; products are long-living goods; high number few product categories; products are long-living goods; high number of clicks
F2: Specialized-Suppliers Permanent-Offers	53	Offer products only in a few product categories; offered over a long period of the product life cycle; do not put many of the offered products into storage; few price changes; if they adjust prices the magnitude of the change is quite high; products offered are only offered by a few other retailers; offer products with highest absolute price level; high relative price; low number of clicks; good rating
F3: Long-Shot-Firms Long-Shot-Offers	224	Products offered are not in stock; offers are only listed for a very short time of the product life cycle; prices are often changed; relative price level is high; observed shipping costs are beyond the average; no pick-up possibility; offer many products in many product categories; low number of clicks
F4: Power-Sellers In-Stock + Permanent-Offers	59	Offer high expensive products; relative low median price; low shipping cost; high number of clicks; if listed, it is offered for more than half of the product life cycle; high availability; few price changes; offer a small number of products; assortment is not concentrated on certain product categories; good retailer rating; a majority of firms operate brick-and-mortar facilities
F5: Short-Term-Suppliers In—Stock + Long-Shot-Offers	87	Offer products only for a short time of the product life cycle; availability of the products is high; many price changes; variation of price is low; high shipping costs; planned price rank is below the average; wide product portfolio; rather badly rated by customers; products with a short product life cycle; high number of clicks
F6: Large-Department-Stores Permanent + Long-Shot-Offers	117	Highest number of offers; high number of clicks; wide product portfolio; combined with brick-and-mortar facilities; high average price; high shipping costs; do not aim at best-price rankings; low availability; products are listed almost half of the product life cycle; number of price changes is at an average level; if prices are changed, the variation is quite high; products offered are more expensive than the average product; have a shorter product life cycle than the average
F7: Mixed-Strategy-Type remaining	10	Low shipping cost; very good rating; low pick-up possibilities; offer only few products; on markets with few competitors; relatively low price

Stock-Offers and *Permanent-Offers*, F5 is *In-Stock-Offers* and *Long-Shot-Offers*, and F6 is *Permanent-Offers* and *Long-Shot-Offers*.[21] (c) The remaining retailers are assigned to firm-strategy pool F7, which reflects firms with mixed e-commerce strategies.[22]

Looking at the number of retailers assigned to each group, we see that *F1: In-Stock-Firms* (with 230 retailers), *F3: Long-Shot-Firms* (with 224 retailers), and *F6: Large-Department-Stores* (with 117 retailers) are of particular importance. Although the mass of *In-Stock-Offers* is concentrated in the F1 firm pool and that of *Long-Shot-Offers* is concentrated in the F3 pool, we observe the highest number of *Permanent-Offers* in the F6 pool. Pools F1, F3, and F6 account for 73% of all retailers and cover 85% of all offers. Table B.4 in the Web-Appendix gives an overview of the distribution of offers over the firm pools. Table B.12 in the Web-Appendix shows descriptives for the different firm types.

Columns (3) and (4) in Table 5 use these firm pools as explanatory dummy variables.[23] The firm pool *F1: In-Stock-Firms* acts as the base group for all regressions. Starting with the comparison of the large firm pools F1, F3, and F6, we confirm our results at the offer level. We do not observe significant differences between the success of firm pools *F1: In-Stock Firms* (with mainly *In-Stock-Offers*) and *F6: Large-Department-Store* (in which *Permanent-Offers* are predominant). In comparison with the cheap and immediately available products of *F1: In-Stock Firms*, the broad product assortment and loss leader strategies might attract consumers to *F6: Large-Department Stores*. In this case, loss leaders (or complementary) products are not especially cheap but are hard to obtain elsewhere. Customers accept these offers, as they can typically save on shipping costs and only have to deal with one store. In contrast to the results for pools F1 and F6, we do not find any empirical evidence that *F3: Long-Shot-Firms* use a successful e-commerce strategy. The same finding applies to the considerably smaller group of *F5: Short-Term-Suppliers*, which is a mixture of F1 and F3 firms. F5 retailers perform worse than F1 retailers, but better than F3 retailers, which is due to the mixture of the two strategies.

An additional interesting finding is that there are two small firm pools that account for neither the mass of offers nor a large number of retailers but perform better than the successful firm pools F1 and F6. These two firm pools are *F2: Specialized-Suppliers* (53 retailers) and *F4: Power-Sellers* (only 59 out of 780 retailers). Clearly, these

[21]Changing the percentage limit to 60% or 80% does not substantially change the assignment of retailers to firm pools or the corresponding success rates of firms discussed later in the text. The results can be found in Tables B.2 and B.3 in the Web-Appendix https://www.econ.jku.at/t3/papers/MDE2021.pdf.

[22]In Table 6, we provide a short description of the firm types. Subsection A.2 in the Web-Appendix contains a detailed characterization of the the firm pools.

[23]Table B.5 in Web-Appendix shows the success of different firm pools with regard to the number of clicks, revenues, click shares, and the number of LCTs.

two small strategy firm pools perform better than the pools F1, F3, F5, and F6. A detailed inspection of the characteristics of *F4: Power-Sellers* shows that these retailers are similar to F1 retailers. Clearly, these *F4: Power-Sellers* utilize special managerial skills with regard to assortment composition and selective warehousing, which are highly attractive for consumers. The most successful firm pool, however, is *F2: Specialized-Suppliers*. These are shops that identify highly profitable niches of special products that are only occupied by a few other retailers. The final group, *F7: Mixed-Strategy-Type*, also exhibits a high probability of survival. However, as this group consists of only seven firms, we refrain from a characterization of firm strategies.

Controlling for various retailer characteristics, we can show that firms with pick-up possibilities, and therefore higher distribution costs, have lower survival rates. Unlike in the case of pure online trading, a half-hearted switch from a traditional brick-and-mortar store to an e-commerce business might also explain the negative effect of the pick-up variable. Retailers with good firm ratings show higher survival rates in 2012 in regressions without controls for product category fixed effects. Retailers without ratings (whose ratings we had to impute using the average firm rating) are young e-commerce companies that are still trying to build reputations and customer bases. Such firms perform worse than those with at least one rating. Finally, we see a significant survival advantage for larger firms (measured by the number of products offered by a retailer).

6 | ROBUSTNESS

We perform the following robustness checks. (i) We bring causal evidence for the effects of strategy clusters on our success variables. (ii) We check the stability of our results with respect to different product groups, and (iii) we demonstrate that changing our clustering variables does not change the assignment of offers to our clustering categories. (iv) The assessment of different clustering strategies presented so far rests on the relative performances of several success indicators at the product level. However, one might argue, success in absolute terms is the decisive variable, and, thus, we also demonstrate the robustness of our results using absolute measures for success defined at the firm level. (v) Finally, one might speculate whether e-commerce strategies might change over the product life cycle. We demonstrate that only a small share of offers change e-commerce strategy types over the life cycle of the products.

6.1 | Causal evidence at the offer level

For a test, whether the strategy choice of a certain cluster type has causal impact on the success variables, we propose the following instrumentation strategy: for a given firm, we use the cluster variables from a predecessor good offered by the same firm in the same sub-subcategory as instrument in IV regressions. Predecessor products of good i and firm j have been selected in the following way (see also Figure B.2 in the Web-Appendix): (i) predecessors must have their market launch at least 365 days before the market launch of product i. (ii) The end of the predecessors' product life cycle must not lie after the market launch of i (otherwise, the exclusive restriction would be violated). (iii) They must be offered by firm j. (iv) Predecessors must have clicks to calculate a product life cycle with a begin and end time. From potentially 353,494 available candidates for predecessors with a valid market launch date we lose (i) 15,403 products because they have no clicks, (ii) 163,131 products because the ends of their product life cycles lie after the "birth date" of the products in our dataset, (iii) 131,716 products as they were not offered by the respective firm, and (iv) 8887 products due to missing data. Hence, from the original sample size of 149,862, we have only instruments for 34,357 offers.

We use a firm's past strategic decisions as an instrument for a firm's contemporaneous strategic decisions. Our instrumentation strategy takes advantage of the fact that corporate strategic decisions may exhibit temporal persistence. We chose *Daily Price Changes*, *Listing Percentage*, *Availability Percentage*, and *Absolute Shipping Cost* of the predecessor good as instruments as these variables have the highest contribution in the k-means clustering procedure. For a valid instrument, two conditions must hold. First, the instrument has to be relevant and second, it has to satisfy the exclusion restriction.

As the first stage results show, our instruments are relevant. We find that the past strategic decisions are highly correlated with contemporaneous strategic decisions. This is shown by the values of the F-statistics in the first stage regressions. The Cragg–Donald Wald F-statistics are 240.5 without fixed effects and 22.3 controlling additionally with product and e-tailer fixed effect. As these values are larger than 10, we can reject the hypothesis of weak instruments (Staiger & Stock, 1994; Stock et al. 2002). Hence, our first stage regressions show that a company that decides for a specific strategy on predecessor products for certain reasons will also apply this strategy to successor products. We argue that there is no violation of the exclusion restriction because at the time of the strategy decision for the predecessor product, the success of the successor product was unclear.[24] Hence, the instrument should have—above its influence on contemporaneous strategies—no additional impact on the demand for the product. We consider this a realistic assumption as selling strategies for another product, typically 1 year ago, unlikely have an impact on current demand.

Our identification strategy might be invalid, if there are unobserved firm- and product-specific variables which are persistent over a series of products, both correlated with the clustering strategies and influencing demand. Suppose that the quality of service is such a variable. We could assume that service quality for a particular product is correlated with a specific firm strategy. If such missing

[24]We use a frequently employed instrumentation strategy from dynamic panels (Arellano & Bond, 1991) and (Blundell & Bond, 1998). We look at market participants' behavior in other markets at earlier times when the realization of the outcome variable was not known. Our approach is also comparable to the use of prices of the same product in other independent markets in the demand literature (Hausman, 1996; Nevo, 2001) or to the shift-share approach common in the migration literature (see, for instance, Card, 2001).

TABLE 7 Causal evidence: success of instrumented clusters at the offer level

	No Fixed Effects			Fixed Effects		
	OLS	OLS	IV	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Clicks						
Permanent-Offers	−26.87***	−95.13***	−46.94***	−13.35***	−39.80***	−46.32
	(0.716)	(2.570)	(6.848)	(1.018)	(3.977)	(68.14)
Long-Shot-Offers	−42.11***	−110.8***	−75.16***	−23.35***	−49.14***	−147.5**
	(0.650)	(2.635)	(9.683)	(0.943)	(3.968)	(60.62)
Constant	45.53***	115.2***	79.01***	102.0***	99.14***	128.2***
	(0.538)	(2.239)	(5.184)	(4.755)	(16.88)	(47.66)
R^2	0.027	0.051	0.041	0.234	0.267	0.226
Number of Last-Click-Throughs						
Permanent-Offers	−1.724***	−6.846***	−3.654***	−0.787***	−2.553***	−3.876
	(0.0568)	(0.204)	(0.544)	(0.0818)	(0.320)	(5.511)
Long-Shot-Offers	−2.889***	−8.119***	−6.881***	−1.511***	−3.308***	−12.43**
	(0.0515)	(0.209)	(0.769)	(0.0758)	(0.320)	(4.902)
Constant	3.153***	8.508***	6.542***	8.924***	10.42***	12.01***
	(0.0426)	(0.178)	(0.412)	(0.382)	(1.359)	(3.854)
R^2	0.021	0.043	0.035	0.208	0.241	0.192
Revenue by Clicks						
Permanent-Offers	−4511***	−22,642***	−10,958***	362.9	−2929**	−51,977**
	(308.9)	(945.2)	(2518)	(428.1)	(1459)	(24,967)
Long-Shot-Offers	−11,119***	−28,948***	−28,920***	−4756***	−8505***	−68,924***
	(280.4)	(969.0)	(3560)	(396.8)	(1455)	(22,210)
Constant	12,781***	31,013***	25,551***	31,221***	28,042***	38,196**
	(232.1)	(823.3)	(1906)	(2000)	(6191)	(17,463)
R^2	0.011	0.026	0.016	0.260	0.251	0.211
Revenues by Last-Click-Throughs						
Permanent-Offers	−340.9***	−1852***	−801.1***	−6.995	−246.6*	−5452**
	(27.08)	(85.28)	(227.9)	(37.99)	(132.0)	(2291)
Long-Shot-Offers	−881.5***	−2412***	−2797***	−385.5***	−711.8***	−7494***
	(24.57)	(87.43)	(322.2)	(35.21)	(131.7)	(2038)
Constant	1022***	2610***	2268***	3300***	3499***	4442***
	(20.34)	(74.28)	(172.5)	(177.5)	(560.3)	(1603)
R^2	0.010	0.022	0.006	0.240	0.244	0.180
Product Fixed Effects				X	X	X
E-tailer Fixed Effects				X	X	X
Observations	149,862	34,357	34,357	149,862	34,357	34,357
Number of Retailers	780	241	241	780	241	241
Number of Products	4888	2024	2024	4888	2024	2024
F-Stat (Cragg–Donald)			240.498			22.349

Note: In all regressions, In-Stock-Offers represent the base scenario. Columns (1) and (4) should facilitate the comparison and can also be found in Columns (1) and (3) of Table 4. Standard errors are in parentheses.

*$p < 0.1$.

**$p < 0.05$.

***$p < 0.01$.

variables, like service quality, are firm specific, but not specific for each product the firms sell, it will be taken up by our firm fixed effects. A similar argument can be brought forward for unobserved product-specific variables which are controlled for by product-specific fixed effects. In addition, the organization of the price comparison platform makes it unlikely that there are such unobserved factors. If a costumer visits the price comparison platform, she sees the same information that we use in the clustering approach (price path, listing, direct availability, and shipping cost). Clicking on an individual e-tailer's website is usually based on this information, although we cannot exclude that the costumer uses other information as well. However, omitted firm-product-specific (and not only either firm or product specific) variables would question our identification strategy (e.g., if the retailer would have different reputations in the service quality for different product groups).

Our IV regression results can be seen in Table 7. Columns (1) and (4) replicate OLS regressions from Table 4. For comparison reasons, Columns (2) and (5) show OLS results for the reduced sample for which instruments are available. Finally, Columns (3) and (6) depicts the 2SLS regressions. Note, that our IV regressions controlling

TABLE 8 Which e-commerce strategy is used for which product?

	Permanent-Offers	Long-Shot-Offers	Permanent-Offers	Long-Shot-Offers
	(1)	(2)	(3)	(4)
Category hardware	0.232***	0.262***	0.260***	0.400***
	(0.0377)	(0.0330)	(0.0386)	(0.0334)
Category software	0.224***	−0.117**	0.139**	0.116**
	(0.0543)	(0.0506)	(0.0557)	(0.0516)
Category games	−0.545***	0.511***	−0.296***	0.429***
	(0.0629)	(0.0492)	(0.0640)	(0.0500)
Category TV	0.0581	0.139***	0.0607	0.208***
	(0.0392)	(0.0345)	(0.0402)	(0.0348)
Category phone	−0.501***	0.0377	−0.287***	−0.0442
	(0.0496)	(0.0414)	(0.0504)	(0.0419)
Category audio	−0.145***	0.0720**	0.0923**	0.0125
	(0.0351)	(0.0307)	(0.0358)	(0.0312)
Category movies	−0.113*	−0.127**	−0.156**	−0.195***
	(0.0655)	(0.0584)	(0.0665)	(0.0589)
Category household	−0.205***	0.00232	0.206***	−0.236***
	(0.0431)	(0.0376)	(0.0445)	(0.0386)
Category sport	−0.329***	−0.387***	0.0122	−0.729***
	(0.0816)	(0.0726)	(0.0826)	(0.0737)
Category drugstore	−0.784***	−0.316***	−0.354***	−0.450***
	(0.0728)	(0.0589)	(0.0738)	(0.0597)
Category miscellaneous	−0.734***	−0.0935	−0.322	−0.390*
	(0.259)	(0.199)	(0.260)	(0.199)
P Brand Strength (/10,000)			−0.154***	0.00703
			(0.0128)	(0.0116)
P Median Absolute Price (/1000)			0.166***	0.157***
			(0.0125)	(0.0122)
P No. of Offering Firms (/100)			1.024***	−2.939***
			(0.0767)	(0.0711)
P Product Life Cycle Duration (/100)			−0.102***	0.0308***
			(0.00246)	(0.00226)
Constant	0.17***	0.590***	0.811***	0.691***
	(0.0382)	(0.0334)	(0.0470)	(0.0417)
Observations	149,862	149,862	149,862	149,862

Note: Multinomial logit model; "In-Stock-Offers" are the base category. Standard errors are in parentheses.
*$p < 0.1$. **$p < 0.05$. ***$p < 0.01$.

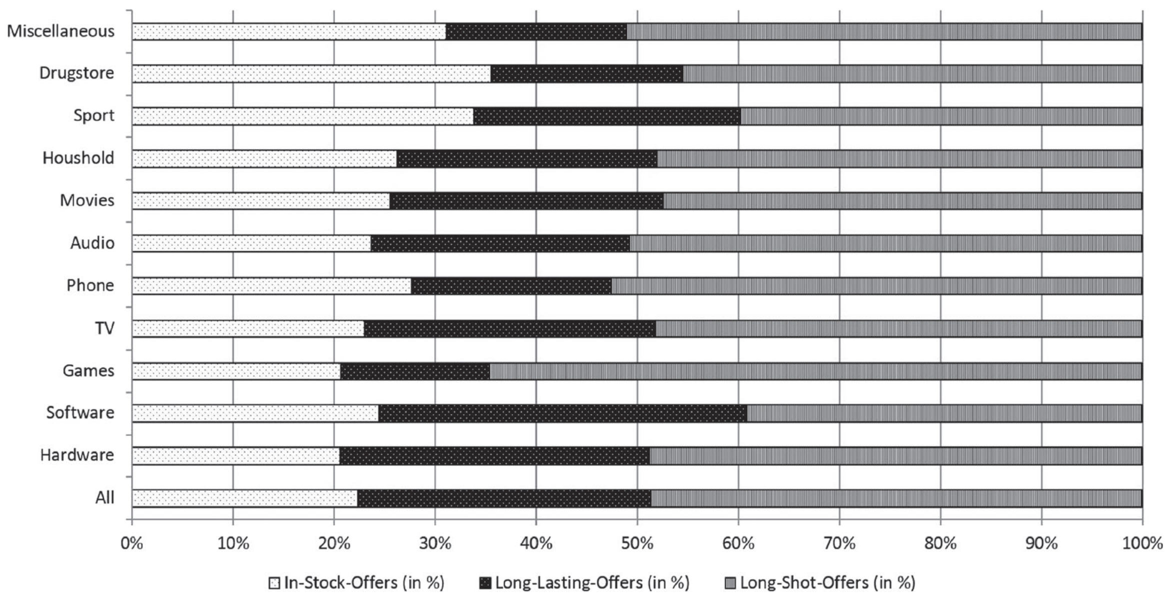

FIGURE 4 Cluster shares across product groups. The figure shows the shares of our strategy clusters (In-Stock-Offers, Permanent-Offers, and Windfall Offers) across different product groups

for potential endogeneity strongly confirm the result of Table 4. Compared with *In-Stock-Offers*, *Permanent-Offers* tend to be less successful (only for the smaller IV sample including all the e-tailer and product fixed effects we have an insignificant coefficient for *Number of Clicks*and *Number of Last-Click-Throughs*). Long-Shot-Offers are again the worst strategy.

Given these causal results, we continue to argue with the OLS coefficients of our full sample for the following reasons: (i) as the IV sample only comprises 23% of the full sample, we prevent a substantial reduction of our sample size. (ii) As the coefficients of our IV regressions are consistently larger than our OLS estimates in the full sample, this corresponds to a conservative approach, which understates our result rather than exaggerate them.

6.2 | Usage of e-commerce strategies across product groups

In a second robustness check, we analyze whether the usage of e-commerce strategies differs across product groups. Particular strategies may be seen as reactions to consumers' search profiles. Consumers may search differently for more durable goods, such as TVs, than for more short-lived products, such as games.

Table 8 shows the results of a multinominal logit model with the choice of e-commerce strategy as the dependent variable and product categories as explanatory variables. In addition to product group fixed effects, we also include explanatory variables, such as the median absolute price and the number of firms that offer the product. Table 8 shows the results with the base group of *In-Stock-Offers*. We note that the product group effects are significantly different from 0 and reflect the percentages in Figure 4. Additionally, we find that higher prices increase the probability of using more *Permanent-Offers*. Furthermore, if there are more firms in the market, we observe more *Permanent-Offers*.

We observe that e-commerce strategies are used differently in specific industries. Next, we evaluate whether they have different success rates in different groups. We calculate success measures comparable to those in Column (3) of Table 4 for each of our product groups. For a better comparison, in Table 9, we show relative changes in the success outcomes when switching from *In-Stock-Offers* to another strategy. We find notable group-specific differences, especially for information goods like software or movies, for which the statistical difference between *Permanent-Offers* and *Long-Shot-Offers* nearly vanishes. Moreover, *Permanent-Offers* is the most successful strategy for selling phones.[25]Although we see some group-specific differences, our main results on the success of different e-commerce strategies hold for most of the product categories. *In-Stock-Offers* are more successful than *Permanent-Offers*, whereas *Long-Shot-Offers* perform worst. The corresponding coefficients for Table 9 can be found in Table B.6 in the Web-Appendix.

6.3 | Clustering using variables determined by competition

Thus far, all our clustering variables can be unanimously determined by the retailer and do not reflect consumer reactions. In this subsection, we present the results of a robustness check, in which the

[25]It should be mentioned that cell phones are very often bundled with a contract from a mobile phone providers.

TABLE 9 Success of different clusters across product groups

	All	Hardware	Software	Games	TV	Phone	Audio	Movies	Household	Sport	Drugstore
Number of Clicks											
Permanent-Offers	−29.32%	−50.72%	−99.14%	−59.07%	−19.24%	28.96%	−23.88%	−29.08%	−13.74%	−22.90%[a]	10.33%
Long-Shot-Offers	−51.28%	−67.43%	−103.45%	−58.67%	−32.12%	−13.56%[a]	−30.61%	−29.26%	−51.91%	−47.57%	−19.15%[a]
Number of Last-Click-Throughs											
Permanent-Offers	−24.96%	−44.19%	−116.48%	−56.62%	−19.76%	40.10%	−17.90%	−27.13%	−12.02%	−18.79%	−2.89%
Long-Shot-Offers	−47.92%	−64.54%	−112.09%	−62.51%	−30.78%	−7.96%[a]	−22.35%	−20.78%[a]	−44.79%	−44.91%	−25.44%[a]
Revenues by Clicks											
Permanent-Offers	2.84%[a]	4.53%[a]	428.09%	−76.99%	3.58%[a]	61.27%	−18.96%	−13.52%[a]	4.30%[a]	−45.61%[a]	−2.35%[a]
Long-Shot-Offers	−37.21%	−41.18%	188.24%	−70.10%	−21.12%	3.26%[a]	−31.94%	−10.90%[a]	−61.19%	−115.06%	−23.53%[a]
Revenues by Last-Click-Throughs											
Permanent-Offers	−0.68%[a]	−0.33%[a]	46.89%[a]	−72.36%	−5.31%[a]	61.50%	−4.04%[a]	−8.38%[a]	4.73%[a]	17.94%[a]	−13.45%[a]
Long-Shot-Offers	−37.72%	−46.87%	−17.58%[a]	−71.90%	−24.41%	3.76%[a]	−13.09%[a]	0.17%[a]	−52.78%	−96.45%	−28.82%[a]
Means of Base Group (In-Stock-Offers) for Each Product Group											
Number of Clicks	45.53	29.2	16.23	85.87	76.18	74.68	65.23	83.84	59.8	66.45	54.1
Number of Last-Click-Throughs	3.153	2.005	1.183	5.044	6.078	6.027	4.323	7.335	3.146	3.124	2.697
Revenues	12,781	6851	1769	16,265	32,475	12,343	14,917	25,332	17,354	14,756	5809
Revenues by Last-Click-Throughs	1022	593.6	122	929.1	2821	1090	1042	2402	954.8	766	329.2

Note: Percentages represent the change in the success variable resulting from a switch from an *In-Stock-Offer* to another strategy at the product level. As reference value, we use the mean of our success variables for *In-Stock-offers* neglecting the product and e-tailer fixed effects. These values are shown in the undermost panel. The negligence of product and e-tailer fixed effects in the reference value is the reason for percentage values below 100%.

[a]The coefficients are not significant.

TABLE 10 Comparison of base clustering with competition influenced clustering

		Clustering base version			
		In-Stock-Offers	Permanent-Offers	Long-Shot-Offers	Total
Clustering with	In-Stock-Offers	22.26%	0.08%	0.06%	22.40%
competition	Permanent-Offers	0.03%	28.76%	0.30%	29.08%
variables	Long-Shot-Offers	0.05%	0.14%	48.33%	48.51%
	Total	22.34%	28.97%	48.69%	100.00%

Note: Columns depict the original assignment to clusters in the base version. Rows indicate the assignment of clusters if additional competition variables are considered in the clustering procedure. Note that a variation of clustering variables does not change the assignment of offers to clusters.

clustering procedure includes additional variables that are typically considered to be important, but are determined by the actions of rivals. These variables are *Bestprice Percentage*, *Losses until Reaction*, and *Coefficient of Variation of Relative Rank*. *Bestprice Percentage* is the percentage of time that a given offer by a retailer was the best price among all retailers; *Losses until Reaction*measures the time between dropping by at least one rank in the price ranking and changing the price of an offer for a given product. Whereas *Bestprice Percentage* is a proxy for the aspired price rank, the other two variables are proxies for the effort to maintain this rank. Note that in all three cases, the effort of a retailer can be thwarted by a competitor setting its prices accordingly.

Table 10 shows that using this new clustering procedure does not imply any significant changes in the allocation of offers to the clusters. The columns in Table 10 show the original assignment of offers to clusters in the base version, and the rows depict the offer allocation using our clustering procedure with the extended set of variables. Of the original 22.34% of offers grouped in *In-Stock-Offers*, 22.26% remain in this category. Only 0.03% and 0.05% of offers change cluster categories. The extremely low off-diagonal values confirm this result for the offers in the other two clusters. Hence, both the descriptive statistics and the results of our success analysis do not change if we add additional competition variables.[26]

6.4 | Quality of cluster assignments using silhouette coefficients

In testing whether the cluster of In-Stock-Offers, Permanent-Offers, and Long-Shot-Offers are more successful in Table 4, we treated all observations equally, irrespective of whether the assignment to a cluster was very clear or ambiguous. The quality of the cluster assignment of an observation can be measured with the silhouette coefficient which quantifies, how similar the observation is to the items of its own cluster compared with the observations in all other clusters (see, for instance, Halpin, (2016); Rousseeuw, (1987)). Note, however, that the calculation of the silhouette coefficient is computationally very demanding,[27] so that we split our original dataset of 149,862 observations into four random samples of equal size and replicate our main estimation results from Column (3) in Table 4. Table 11 reports the original regression results together with four sets of unweighted and weighted coefficients.

Although we find the coefficients from the weighted regressions somewhat smaller than the unweighted counterparts, we see all our qualitative results confirmed. Using clicks and LCTs as demand indicators, our results confirm that the *In-Stock-Offers* cluster is always the most successful, followed by *Permanent-Offers* and *Long-Shot-Offers*, which is the least successful cluster. For the revenue variables, we see again that the statistical differences between the *Permanent-Offers* and *In-Stock-Offers* cluster vanish. In any case, the *Long-Shot-Offers* cluster performs the worst.

6.5 | Clustering and the product life cycle

Some studies (Spann et al. 2015) suggest that firms may use different strategies in different phases of the life cycle of a product (PLC) and that such price dynamics may matter substantially in the sales process.

As the PLCs of our products are quite different, with a mean of 895 days, a minimum duration of 101 days, and a maximum of 1475 days, we construct a relative PLC with three phases based on the average number of offering firms, as follows: the growth phase covers 20% of the PLC, the maturity phase extends from the 20th percentile to the 60th percentile, and the declining phase lasts from the 60th percentile until the end of the PLC. This definition of phases is designed according to the development of the number of firms in a market that follows a distinctive inverted U-shaped pattern. Figure B.1 in the Web-Appendix shows the empirical distribution of offering retailers and clicks based on our data for each percentile of the PLC.

Separately, for each of these three phases of the PLC, we can observe our strategy variables that were used in the clustering

[26]Descriptive statistics (means) of the clusters generated using the extended set of variables can be found in Table B.7 of the Web-Appendix. Furthermore, in Table B.8 of the Web-Appendix, we show estimations results for success using clusters based on the extended set of variables. Note that the means of the respective clusters and our success rate regressions essentially coincide.

[27]Standard procedures to calculate the individual silhouette coefficients require the calculation of distances between all observations in the dataset which alone requires a matrix size of 22.5 billion entries.

TABLE 11 Clusters at the offer level—silhouette coefficient as weight

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Full sample	Sample 1		Sample 2		Sample 3		Sample 4	
	Unweighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
Number of Clicks									
Permanent-Offers	−13.35***	−12.82***	−8.255***	−18.63***	−16.16***	−9.004***	−4.214**	−14.23***	−10.06***
	(1.018)	(1.973)	(1.924)	(2.615)	(2.685)	(1.923)	(1.871)	(2.165)	(1.995)
Long-Shot-Offers	−23.35***	−23.27***	−19.85***	−25.69***	−23.96***	−18.83***	−15.74***	−23.34***	−20.38***
	(0.943)	(1.823)	(1.801)	(2.432)	(2.528)	(1.782)	(1.752)	(2.002)	(1.870)
R^2	0.234	0.333	0.334	0.291	0.280	0.322	0.317	0.293	0.303
Number of Last-Click-Throughs									
Permanent-Offers	−0.787***	−0.617***	−0.240	−1.257***	−1.038***	−0.447***	−0.110	−0.954***	−0.633***
	(0.0818)	(0.158)	(0.157)	(0.211)	(0.212)	(0.154)	(0.153)	(0.177)	(0.165)
Long-Shot-Offers	−1.511***	−1.384***	−1.072***	−1.746***	−1.589***	−1.145***	−0.918***	−1.636***	−1.379***
	(0.0758)	(0.146)	(0.147)	(0.196)	(0.200)	(0.143)	(0.143)	(0.164)	(0.155)
R^2	0.208	0.309	0.308	0.251	0.245	0.287	0.284	0.278	0.292
Revenues by Clicks									
Permanent-Offers	362.9	−497.6	1395	456.1	2012**	3204***	4663***	−1069	395.0***
	(428.1)	(839.3)	(809.6)	(985.7)	(1002)	(890.5)	(876.5)	(919.9)	(849.4)
Long-Shot-Offers	−4756***	−5557***	−4219***	−4523***	−3613***	−2116***	−1449*	−5024***	−4033***
	(396.8)	(775.6)	(757.8)	(916.9)	(943.9)	(825.2)	(820.6)	(850.4)	(796.2)
R^2	0.260	0.356	0.393	0.341	0.363	0.354	0.353	0.321	0.349
Revenues by Last-Click-Throughs									
Permanent-Offers	−6.995	−73.69	57.40	−8.259	−113.1	202.1*	290.5***	145.4*	13.74
	(37.99)	(73.86)	(74.65)	(80.72)	(79.68)	(83.63)	(79.86)	(85.09)	(80.25)
Long-Shot-Offers	−385.5***	−448.0***	−336.3***	−375.8***	−301.1***	−170.0**	−127.7*	−445.6***	−343.9***
	(35.21)	(68.24)	(69.87)	(75.09)	(75.02)	(77.50)	(74.77)	(78.66)	(75.22)
R^2	0.240	0.326	0.343	0.334	0.353	0.320	0.312	0.308	0.332
Observations	149,862	37,466	37,466	37,466	37,466	37,466	37,466	37,464	37,464

Note: Column (1) is a replication of Column (3) from Table 4. E-tailer Fixed Effects and Product Fixed Effects are included. In all regressions, In-Stock-Offers represent the base scenario. Standard errors are in parentheses.
*$p < 0.1$.
**$p < 0.05$.
***$p < 0.01$.

process depicted in Table 2. With the exception of two variables, we use exactly the same variables for a *k*-means clustering procedure calculated separately for each of the three phases.[28] Interestingly, comparing the descriptive statistics of the resulting clusters between the phases does not indicate noteworthy changes.[29] The different clusters in the respective phases exhibit more or less identical descriptive features as the cluster groups for the entire PLC in Table 2.[30]

Based on the descriptive statistics, we find little evidence that firms switch their e-commerce strategies over the PLC. This result

[28]Including *End of Offer* and *Beginning of Offer* would not make sense in different phases of the PLC.

[29]Table B.9 shows the descriptive results for the respective clustering analysis in each of the three phases of the PLC.

[30]The descriptive statistics of the clusters remain their relative positions in the maturity and decline phases over all clusters and variables. We observe only one reasonable shift in relative positions in the growth phase; in contrast to our results in Table 2, *Permanent-Offers* indicate the lowest *Planned Price Rank*. This is not surprising, as *Permanent-Offers* enter the market much earlier in the PLC, when only few retailers are present in the market. Due to the low number of retailers, we observe consequently lower aspired price ranks for this cluster in the growth phase of the PLC.

is also depicted in Table 12, which shows the distribution of the original cluster assignment from Table 2 over the clusters of the respective phases of the PLC (e.g., of the original *In-Stock-Offers*, 84.27% remain in this cluster in the growth phase). Only 8.8% of the offers move to the *Permanent-Offers* cluster, and 6.9% switch to the *Long-Shot-Offers* cluster). Analyzing Table 12, we observe that the assignment of offers to their respective clusters largely does not change. The bold figures show values above 50% for each phase of the PLC and indicate that most of the offers remain in the same cluster.

The exceptions are that 37.95% of offers in the original *Long-Shot-Offers* cluster move to the *Permanent-Offers* cluster in the growth phase, and 35.24% of offers in the original *Permanent-Offers* cluster are assigned to the *Long-Shot-Offers* cluster in the declining phase of the PLC. At least for these two relatively small groups, we find confirmation that retailers switch their e-commerce strategies throughout the PLC. Therefore, it is interesting to examine the characteristics and market outcomes of these two product groups in comparison to the nonswitching offers.

Columns (1) and (2) of Table 13 compare offers that were assigned as *Long-Shot-Offers* over the whole PLC. Some of them (Column (1)) were identified as *Permanent-Offers* in the growth phase. Columns (3) and (4) refer to *Permanent-Offers* that are or are not identified as *Long-Shot-Offers* in the declining phase, respectively. We find better outcomes for those offers assigned to the *Permanent-Offers* cluster as compared with those assigned to the *Long-Shot-Offers*, even if the strategy is carried out in only one phase of the PLC, as in Column (1). On the other hand, offers moving from the *Permanent-*

TABLE 12 Comparison of base clustering with phases of PLC clustering

Original assignment	Growth phase	Maturity phase	Declining phase	Phase assignment
In-Stock-Offers	**84.27%**	**92.46%**	**93.32%**	In-Stock-Offers
	8.80%	2.65%	2.30%	Permanent-Offers
	6.93%	4.89%	4.38%	Rent-Skimming-Offers
Permanent-Offers	3.96%	3.27%	4.24%	In-Stock-Offers
	71.46%	**84.41%**	**60.53%**	Permanent-Offers
	24.58%	12.32%	35.24%	Long-Shot-Offers
Long-Shot-Offers	2.12%	1.16%	0.67%	In-Stock-Offers
	37.95%	12.45%	12.84%	Permanent-Offers
	59.93%	**86.38%**	**86.48%**	Long-Shot-Offers

Note: The table shows how different offers can be assigned to different e-commerce strategies (clusters) over different phases of the product life cycle. Note that the assignment over the product life cycle remains by and large relatively stable.

TABLE 13 Comparison of switching and nonswitching offers in the growth and declining phase of the PLC

	Growth phase		Declining phase	
Original assignment:	**Long-Shot-Offer**	**Long-Shot-Offer**	**Permanent-Offer**	**Permanent-Offer**
Phase assignment:	**Permanent-Offer**	**Long-Shot-Offer**	**Permanent-Offer**	**Long-Shot-Offer**
Clustering variables				
C Availability Percentage	**0.0218**	0.0137	**0.0219**	0.00690
C Listing Percentage	**0.796**	0.222	**0.822**	0.276
C Daily Price Changes	0.145	**0.248**	0.126	**0.188**
C Planned Price Rank	11.92	**12.34**	10.95	**11.80**
C Coef. of Variation of Absolute Price	**0.0550**	0.0298	**0.106**	0.0547
C Absolute Shipping Costs	**7.606**	7.465	**7.503**	6.972
Success variables				
S Click Share	**1.923**	0.668	**4.374**	2.562
S Number of Clicks	**5.964**	2.954	**17.93**	12.53
S Number of LCT	**0.574**	0.272	**1.194**	1.083
S Revenue	**2853**	1465	**7489**	5015

Note: The table highlights those offers/products which switch their e-commerce strategy over time. The values indicate means of various descriptives for different groups of offers. The first two columns compare offers which were identified originally as Long-Shot-Offer in the growth phase but switch to Permanent-Offer over the remaining PLC with those offers which have been assigned stably to the Long-Shot-Offer over the complete product life cycle. Similarly, the last two columns compare offers stably assigned to Permanent-Offers over the complete PLC with those offers which switch from Permanent-Offers to Long-Shot-Offers in the declining phase of the PLC. Higher values are marked bold.

Offer cluster to the *Long-Shot-Offer* cluster in the declining phase of the PLC perform worse than offers remaining in the *Permanent-Offer* cluster even at the end of the PLC. Thus, it seems that some unobservable cost factors related to *Permanent-Offers* force retailers to switch strategies for some of their products to the supposedly cheaper *Long-Shot-Offer* strategy during the PLC.[31]

In the context of robustness checks, however, it is important to note that both groups of strategy switchers are relatively small. For most of the offers, we do not observe a change of strategies over the PLC, and, for this large majority of offers, our results based on using one cluster procedure for the whole PLC hold.

7 | DISCUSSION AND MANAGERIAL CONCLUSIONS

Following the advent of online price comparison platforms (e.g., various price search engines, Amazon, and eBay), price competition has increased enormously for B2C e-commerce firms. As prices are highly visible and entry into such markets is relatively easy, a Bertrand paradox can easily arise in which prices fall to marginal costs even in markets with a limited number of firms. In this situation, firms might resort to non-price competition and obfuscation (Ellison & Ellison, 2009) in their efforts of being listed in online platforms. Firms have a large number of strategy options in such "unfriendly" environments, including listing and stocking decisions, price development over time, and auxiliary options for shipping costs.

Using data from an Austrian price comparison site, we statistically identify three distinct strategies that firms use for specific products (*In-Stock-Offers*, *Permanent-Offers*, and *Long-Shot-Offers*) and causally identify their impact on firm success. Whereas the first two strategies are reasonably successful in terms of attention, clicks, and revenues, the third one is not. In addition to looking at strategies for individual products, we can also characterize firms by their combinations of products and strategies. Here, we investigate the survival of these e-commerce firms in the market.

From these results, we can draw the following managerial conclusions for the behavior of online shops in price comparison platforms:

- One successful e-commerce strategy is ordering a large quantity, selling from the shelf relative cheaply, and removing the listing once the stock is sold (*In-Stock-Offer*).
- An alternative strategy is to list the product most of the time without holding it in stock (*Permanent-Offer*).
- Mixtures of these strategies (i.e., neither listing a product for a long time nor holding the product in stock) do not seem to be very successful.
- Looking at the firm level, a couple of specific strategies might pay off. *Power-Sellers* refers to firms including specifically successful products in their portfolios (i.e., high price and high demand products). *Specialized-Suppliers* refers to firms that concentrate on a few product categories with less severe competition.
- As expected, firms with better consumer-assessed quality ratings and those with generally larger product portfolios survive longer; the opposite is true for firms that incur higher costs by having a separate brick-and-mortar store.
- These results hold true for most product groups.

From a broader point of view, our results can also be interpreted with regard to obfuscation strategies. If consumers differ with respect to their search costs, firms may use mixed strategies for a product and randomize prices. A price comparison platform takes away this advantage. Thus, firms have an incentive to obfuscate using add-on pricing, such as shipping costs, and availability. This is, however, not what we observe empirically: we find the lowest relative average price and absolute shipping cost and the highest availability rates for *In-Stock-Offers*. In contrast to that *Long-Shot-Offers* have the highest relative product prices combined with high shipping cost and lowest rates of availability. The cluster of *Permanent-Offers* position itself between the other two (although availability rates are also extremely low). Hence, we do not find a distinct and clear-cut pattern of obfuscation. The clear ranking of obfuscation variables rather suggests a strategy in which firms specializing in *Long-Shot-Offers* try to skim off rents from uninformed customers in a rather clumsy and—as our empirical results about firm survival confirm—unsuccessful way. On the other hand, firms with *In-Stock-Offers* cater to consumers with lower search costs and charge lower prices and low shipping costs.

Although the almost perfectly competitive market[32] for B2C e-commerce firms in a price search engine environment seems to make marketing endeavors obsolete, firms' carefully chosen strategies can make a difference.

ACKNOWLEDGMENTS

We would like to thank Florian Englmaier, Alex Stomper, and workshop participants in Berlin, Frankfurt, and Munich. Financial support by the Oesterreichische Nationalbank (Anniversary Fund, project number: 16255) is gratefully acknowledged. A Web-Appendix to this article can be found at http://www.econ.jku.at/t3/papers/MDE2021.pdf.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this article.

ORCID

Rudolf Winter-Ebmer https://orcid.org/0000-0001-8157-6631

[31]Our research design does not allow us to identify whether these strategy changes are part of a predefined firm plan at the beginning of the product life cycle or whether the changes are the outcome of a learning process over the product life cycle.

[32]See Hackl et al. (2014) for the effect of the number of firms on mark-ups in e-commerce.

REFERENCES

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, *58*(2), 277–297.

Bai, X., & Luo, M. (2011). How much is trust worth? Evidence from the international online textbook market. *Journal of Internet Commerce*, *10*(4), 245–260.

Bauer, J., & Jannach, D. (2018). Optimal pricing in e-commerce based on sparse and noisy data. *Decision Support Systems*, *106*, 53–63.

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

Baye, M. R., Morgan, J., & Scholten, P. (2004). Price dispersion in the small and in the large: Evidence from an internet price comparison site. *The Journal of Industrial Economics*, *52*(4), 463–496.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143.

Böheim, R., Hackl, F., & Hölzl-Leitner, M. (2019). The impact of price adjustment costs on price dispersion in E-commerce. (*7510*): CESifo Group Munich.

Cao, Y., Gruca, T. S., & Klemz, B. R. (2003). Internet pricing, price satisfaction, and customer satisfaction. *International Journal of Electronic Commerce*, *8*(2), 31–50.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, *19*(1), 22–64.

Dinerstein, M., Einav, L., Levin, J., & Sundaresan, N. (2018). Consumer price search and platform design in internet commerce. *American Economic Review*, *108*(7), 1820–1859.

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

Ellison, G., & Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica*, *77*(2), 427–452.

Ellison, G., & Ellison, S. F. (2018). Search and obfuscation in a technologically changing retail environment: Some thoughts on implications and policy. *Innovation Policy and the Economy*, *18*(1), 1–25.

Ellison, G., & Fisher-Ellison, S. (2005). Lessons about markets from the internet. *Journal of Economic Perspectives*, *19*(2), 139–158.

Ellison, S., & Snyder, C. (2014). An empirical study of pricing strategies in an online market with high-frequency price information. MIT Department of Economics Working Paper No. 11-13.

Ellison, S. F. (2016). Price search and obfuscation: An overview of the theory and empirics, *Handbook on the economics of retailing and distribution*: Edward Elgar Publishing, pp. 287–305.

Frischmann, T., Hinz, O., & Skiera, B. (2012). Retailers' use of shipping cost strategies: Free shipping or partitioned prices?. *International Journal of Electronic Commerce*, *16*(3), 65–88.

Gabaix, X., & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, *121*(2), 505–540.

Gorodnichenko, Y., Sheremirov, V., & Talavera, O. (2018). Price setting in online markets: Does IT click?. *Journal of the European Economic Association*, *16*(6), 1764–1811.

Gorodnichenko, Y., & Talavera, O. (2017). Price setting in online markets: Basic facts, international comparisons, and cross-border integration. *American Economic Review*, *107*(1), 249–282.

Hackl, F., Kummer, M. E., & Winter-Ebmer, R. (2014). 99 cent: Price points in e-commerce. *Information Economics and Policy*, *26*, 12–27.

Hackl, F., Kummer, M. E., Winter-Ebmer, R., & Zulehner, C. (2014). Market structure and market performance in e-commerce. *European Economic Review*, *68*(0), 199–218.

Halpin, B. (2016). SILHOUETTE: Stata module to calculate and graph silhouette width for cluster analysis. Statistical Software Components, Boston College Department of Economics.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition, *The economics of new goods*: University of Chicago Press, pp. 207–248.

Haynes, M., & Thompson, S. (2014). Hit and run or sit and wait? Contestability revisited in a price-comparison site-mediated market. *International Journal of the Economics of Business*, *21*(2), 165–190.

Homburg, C., Jensen, O., & Krohmer, H. (2008). Configurations of marketing and sales: A taxonomy. *Journal of Marketing*, *72*, 133–154.

Johnson, E. J., Moe, W. W., Fader, P. S., Bellman, S., & Lohse, G. L. (2004). On the depth and dynamics of online search behavior. *Management Science*, *50*(3), 299–308.

Lloyd, S. P. (1982). Least squares quantization in PCM. *Transactions on Information Theory*, *28*, 129–137.

McDonald, S., & Wren, C. (2018). Multibrand pricing as a strategy for consumer search obfuscation in online markets. *Journal of Economics & Management Strategy*, *27*(2), 171–187.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, *69*(2), 307–342.

Park, C. H. (2017). Online purchase paths and conversion dynamics across multiple websites. *Journal of Retailing*, *93*(3), 253–265.

Pauwels, K., Silva-Risso, J., Srinivasan, S., & Hanssens, D. M. (2004). New products, sales promotions, and firm value: The case of the automobile industry. *Journal of Marketing*, *68*(4), 142–156.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.

Schlosser, A. E., White, T. B., & Lloyd, S. M. (2006). Converting web site visitors into buyers: How web site investment increases consumer trusting beliefs and online purchase intentions. *Journal of Marketing Research*, *70*, 133–148.

Smith, M., & Brynjolfsson, E. (2001). Consumer decision-making at an internet shopbot: Brand still matters. *Journal of Industrial Economics*, *49*(4), 541–558.

Spann, M., Fischer, M., & Tellis, G. J. (2015). Skimming or penetration? strategic dynamic pricing for new products. *Marketing Science*, *34*(2), 235–249.

Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, *79*(4), 700–712.

Staiger, D., & Stock, J. H. (1994). Instrumental variables regression with weak instruments: National Bureau of Economic Research.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, *20*(4), 518–529.

Tang, Z., Smith, M. D., & Montgomery, A. (2010). The impact of shopbot use on prices and price dispersion: Evidence from online book retailing. *International Journal of Industrial Organization*, *28*(6), 579–590.

Tokman, M., Richey, R. G., & Deitz, G. D. (2016). A strategic choice theory taxonomy of retailers' strategic orientations. *Journal of Marketing Theory and Practice*, *24*(2), 186–208.

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

Wilson, C. M. (2010). Ordered search and equilibrium obfuscation. *International Journal of Industrial Organization*, *28*, 496–506.

How to cite this article: Hackl F, Hölzl-Leitner M, Winter-Ebmer R, Zulehner C. Successful retailer strategies in price comparison platforms. *Manage Decis Econ*. 2021;1–22. https://doi.org/10.1002/mde.3309