Investigating Personalized Price Discrimination of Textile-, Electronics- and General Stores in German Online Retail

Tsagana Badmaeva1, Joschka Andreas Hüllmann1

1 WWU Münster, Interorganizational Systems Group, Münster, Germany
{t_badm01, huellmann}@uni-muenster.de

Abstract. Developers of pricing strategies in e-commerce businesses see a wide range of opportunities for deploying online price discrimination techniques given their ability to track consumers’ online identity and behavior. In theory, an increasing use of personal data enables organizations to show every single consumer their own personalized price, which is determined by the consumer’s characteristics, e.g. age, gender, surfing history, or location. This paper aims to explore the existence of online price discrimination activities within the German e-commerce market using a three-method approach. First, inquiring the online retailers via email and investigating their public documents; second, surveying students; and third, using a software crawler to simulate surfing activity. Our results do not provide any evidence of individualized price discrimination, which, we argue, is due to economic and political reasons, not technical reasons.

Keywords: online price discrimination, tracking, privacy, e-commerce, personalization.

1 Introduction

The German e-commerce market increases steadily as people are more comfortable and willing to purchase goods online1. At the same time, the trend for personalization in online retail is growing [1, 2]. We see prospects of e-commerce businesses to collect and evaluate personal data such as age, gender, location or device used in order to offer specific prices for each customer – called personalized price discrimination. However, the public is wary of this development and raises concerns about privacy, tracking and consumer protection. Institutions such as the OECD [3], the German Federal Ministry of Justice [4], consumer protection initiatives [5, 6], newspapers [7, 8], and the citizens question what effects data and algorithms have on consumers. For a debate on price discrimination, it is vital to uncover discriminatory practices in order to make interac-

---

1 https://www.statista.com/statistics/453490/e-commerce-retail-revenue-share-germany/ (last access 2018-10-12).
tions with online retailers transparent. Using personal data for setting prices is a sensitive issue, in particular in Germany [9], which could require political or legislative measures in the future [10].

Previous studies [6–8] explored personalized pricing practices of internationally operating e-commerce retailers, ranging from hotel-booking sites to general online retailers, by using software scrapers. However, the studies do not focus on the German market, while existing German studies prefer a manual approach to collect data and have small sample size [9–11]. We lack comprehensive information on the current status of personalized pricing in German online retail. To close this gap, we suggest triangulating results from manual data collection and automatic scraping.

In this paper, we explore the prevalence of personalized pricing activities performed by textile-, electronics- and general stores in the German online retail market. We selected the eleven biggest online retailers by revenue and investigated the presence of price discrimination practices in their shops by combining surveys, an automated crawler and a discourse analysis. Our study contributes an overview of the tracking ecosystem and reports on the existence of personalized pricing activities in German online retail. It provides survey items and a software crawler to identify such mechanisms. Our research does not detect any price discrimination practices for the selected online retail shops, which, we argue, results from economic and political reasons.

2 Background

2.1 Price Discrimination

Effective pricing strategies are key to increase demand, sales and profit [16]. One pricing strategy is price discrimination (or differential pricing). In general, price discrimination is defined as a variation of price cost ratio across units or groups of buyers [17]. In our context, we define price discrimination as follows: the producer sets varying prices, which includes discounts, for an identical product or service for different consumers based on the consumer's characteristics, the time or location of purchase, the amount of purchase, or other relevant information [14], [16–18]. The goal is that the consumer pays the highest amount that they are willing to pay, i.e. the marginal willingness to pay (WTP) [21]. Price discrimination requires information about consumers and the producer’s ability to estimate the consumer’s marginal WTP as accurately as possible [19]. In an ideal information environment, the producer can set the price to the consumer’s marginal WTP to extract the maximum consumer surplus and maximize his profit. Depending on the available information, three degrees of price discrimination are distinguished [17, 22].

Third degree price discrimination describes different prices for segments of consumers and is the most common form of price discrimination. It requires sociodemographic information about the consumer [17]. Examples are student or senior citizen discounts or regional price variations. Second degree price discrimination is based on the amount or quality of a given product or service, usually including consumers’ self-selection of amount or quality, to mitigate the difficulty of distinguishing consumer types. Examples for amount are volume discounts and examples for quality are airfares,
namely business and economy tariffs. **First degree price discrimination** requires the most information as each person gets a personalized price based on their characteristics [17]. With detailed information about the consumer, the producer can determine the maximum WTP and capture the maximum consumer surplus. Collecting enough information to implement first degree price discrimination is difficult. Nonetheless, large-scale data collection is a first step to enable first degree price discrimination. For example, Shiller [23] demonstrated a profitable first degree price discrimination model based on 5000 trackable consumer characteristics using behavioral and sociodemographic traits. Historically, this type of price discrimination is known from the used cars or insurance market or from the bazaar in one-to-one transactions. Now, it is possible to implement and automate it for all consumers online.

The degrees are not mutually exclusive [19]. Personalized price discrimination is based on the segments of third degree price discrimination, with the ultimate goal to achieve a first degree price discrimination. Alternatively, we can explain personalized price discrimination as an extreme case of third degree price discrimination with segments of size one.

### 2.2 Conditions and Implementation of Price Discrimination

Certain conditions must be met to implement price discrimination. A firm (1) needs some market power, or consumers buy elsewhere, (2) it needs to control the sale of the product or service to set the price and (3) the consumers’ willingness to pay must differ from one another [17]. Sufficient information about consumers to distinguish them into segments must be available [24, 25]. The segments need to be actionable, substantial, accessible, measurable, profitable and stable and are distinguished by behavioral, psychographic, sociodemographic and geographic criteria (for more information see [23–25]). Behavioral criteria include used channels, brand loyalty, buy volume, and previous purchases. Psychographic criteria include lifestyle, social identity and personality, for example, attitude towards risk, expectations of quality of products. Sociodemographic information includes gender, age, children, job, education, income. Geographical criteria are distinguished into macro (country, city) and micro (district, street). Based on data for each customer, segments can be inferred from quantitative statistical analysis such as cluster analysis or machine learning algorithms [26, 27].

The data is sourced internally or acquired externally. Internal data includes customer master data from CRM systems, i.e. data that a consumer enters on the website, for example, name, address, city, etc. Other data is based on the actions that customers perform, i.e. which articles they view, click, purchase, rate and review. Besides, the data includes technical information such as IP address, device, user-agent, or operating system. Prevalent shop software solutions offer the functionality to collect this data, and to use it for consumer segmentation as well as price discrimination. In addition to analyzing internal data, shop solutions offer integrations with external data sources.

---

2 Market segmentation is a concept that goes beyond price discrimination.
External data sources originate either from third party analytics services or trackers\(^3\), which capture users’ online activities on the web \cite{28} and may provide extended insights \cite{29,30}, or from offline datasets, e.g. voter records, or vehicle owner data. There are multiple services available, for example, TruSignal, Adobe Customer Experience Cloud, Google Analytics 360, Bluekai, Lotame, and more. However, the market of these services is highly concentrated around a few big players, who accumulate data, which makes it valuable to acquire, but also sensitive to privacy issues \cite{29–31}.

### 2.3 Consumer Perspective

Growing consumers’ awareness of aggressive tracking and the amount of gathered data lead to a discussion on the privacy and ethics of tracking and personalized prices. This is exacerbated by the fact, that the big players accumulate most of the data \cite{32,33}, and the extent of data collection and sharing is concealed from consumers \cite{30}.

Since online retailers are not transparent about their pricing mechanisms, it is unknown if personalized pricing takes place at all, and what type of data significantly changes the segmentation of consumers and the resulting prices \cite{15}. The growing data collection and the lack of transparency on what consequences particular information has, lead to privacy issues and may unsettle consumers \cite{34}. There is an information asymmetry between the online retailers and the consumers \cite{31}. Consumers can only believe the privacy statements, which are mandated since the EU-GDPR and given by online shops, as there are no (easy) neutral ways to check for personalized price discrimination \cite{35}. For a debate on personalized price discrimination in Germany, it is crucial to provide empirical means to investigate the current extent of personalized price discrimination in online retail. Therefore, we contribute such means to assess the extent of personalized price discrimination, empirically, and report on our results.

### 3 Related Work

Previous studies have found little to none price discrimination in online retail. Despite checking 200 U.S. shops, Mikians et al. \cite{12} found only few price differences caused by a combination of system, sociodemographic and behavioral criteria. Hannak et al. \cite{11} explored U.S. shops and the tourism industry using Amazon Mechanical Turk and simulated user profiles. They found different prices on nine of sixteen e-commerce websites due to changes in system and behavioral criteria. Vissers et al. \cite{36} and Hupperich et al. \cite{13} investigated the tourism industry, both using simulated browser fingerprints and user profiles. Vissers et al. did not find any price discrimination, while Hupperich et al. find differences based on geographic criteria. For both studies, other criteria did not lead to any price changes. Schleusener and Hosell \cite{4} and Kraemer et al. \cite{14} focused on the German market using a manual approach. The former found

---

\(^3\) Due to space limitations, we cut the technical description of third-party trackers and the tracking ecosystem. Instead, an explanation is provided in \cite{38,39}.
price discrimination in the tourism industry, with system criteria being relevant. Interviews with professionals corroborated their results. Kraemer et al. [14] found price discrimination based on geographic and system criteria.

The findings of previous studies are inconclusive, as some results indicate price discrimination in the tourism industry, while other studies find no price discrimination at all. Furthermore, only two studies focus on the German online retail market, leaving it underexplored and requiring further research.

4 Research Design

To investigate the existence of price discrimination in the online retail market, we select the eleven biggest German online e-commerce retailers according to their annual revenue. They are the most relevant shops from a consumer rights perspective, i.e. more revenue means more involved consumers, and we choose textile, electronics and general stores, because they are less explored compared to the tourism industry. First, we analyze the discourse on price discrimination using public information on the selected shops, shop software vendors, tracking providers followed by contacting them. Second, we build a survey based on the found information, which captures personal user information. Test persons filled out the survey and use their personal devices to check prices in each shop, manually. Third, we use an automated software crawler, which simulates user activity and user profiles, to check the prices in each shop.

Table 1. Selected Online Shops (Revenue in Million EUR in 2015)4

<table>
<thead>
<tr>
<th>Shop</th>
<th>Revenue</th>
<th>Sector</th>
<th>Shop</th>
<th>Revenue</th>
<th>Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>amazon.de</td>
<td>7,790.60</td>
<td>General</td>
<td>tchibo.de</td>
<td>450.00</td>
<td>General</td>
</tr>
<tr>
<td>otto.de</td>
<td>2,300.00</td>
<td>General</td>
<td>conrad.de</td>
<td>433.20</td>
<td>General</td>
</tr>
<tr>
<td>zalando.de</td>
<td>1,031.80</td>
<td>Textile</td>
<td>alternate.de</td>
<td>376.70</td>
<td>General</td>
</tr>
<tr>
<td>notebooksbilliger.de</td>
<td>610.90</td>
<td>Electronic</td>
<td>hm.de</td>
<td>344.60</td>
<td>Textile</td>
</tr>
<tr>
<td>cyberport.de</td>
<td>491.30</td>
<td>Electronic</td>
<td>esprit.de</td>
<td>342.00</td>
<td>Textile</td>
</tr>
<tr>
<td>bonprix.de</td>
<td>484.70</td>
<td>Textile</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First, for the discourse analysis, we investigate the most prevalent trackers among the top 20 worldwide [33, 37] and observe the involved actors as well as their activities. We evaluate public documents on their websites in order to understand what, how and why they track the information of consumers in their services. Besides the trackers, we look at the three biggest standard shop software by market share in 2017 (Shopify, WooCommerce, Magento)5. We evaluate their websites for functionality related to tracking, consumer segmentation and price discrimination; and what consumer criteria the software makes use of. We also contact them directly, and assess available third


The market share has changed since we collected the data.
party addons for each shop software. According to the EU-GDPR, e-commerce companies have to disclose how they obtain, process and use personal citizens' information as well as the purposes of doing so. This implies that companies are to disclose whether they perform any kind of price discrimination activities using personal data of customers. Using this opportunity, we assess the privacy notices of each online shop and contact the data protection departments of each online shop. The question that we forwarded to the data protection departments is:

"Due to the EU-GDPR, we would like to inquire, whether and in which way personal data influences the determination of visible prices, including potential discounts or vouchers (also via mail or email), in your online shop." (Translated from the original German statement).

Second, the three checked shop software packages provide the functionality to implement personalized price discrimination, but there is no information if any of the shops use this feature. Only the number of reviews hints that the features may be used, so we use the criteria collected from the discourse analysis to develop a survey. The survey consists of two parts. The first part is a list of 20 hyperlinks to products of the selected shops. The second part asks for behavioral and sociodemographic criteria as well as system attributes as seen in Table 3. We distributed the survey to four different student groups in Germany depicted in Table 2. The students checked the prices of the 20 selected products with their own devices, browser settings, and accounts and filled out the survey accordingly. Some students used proxies and VPNs, but the most common types of connection were University WiFi and mobile data. Due to logistical issues, we were unable to record the type of connection.

Third, due to the logistics of conducting a survey, we develop an automated software to triangulate our findings based on [12, 13]. The goal of the software is to simulate ordinary user activity for seven days, to build five online user profile and accumulate cookies. After seven days, it checks the prices using the five profiles.
Each user profile is generated from two keywords (Table 4) that the crawler uses to engage on search, news and social media sites (Table 5), and is assigned a distinct proxy to simulate geographic location. The software is implemented in Python using Selenium with the Firefox driver and the source code is available on GitHub.

5 Findings

Discourse Results. The tracking ecosystem sees actors with varying primary activities, from data management platforms to marketing solutions, who track consumers’ information. We found that most of them have dedicated services that perform tracking. The tracked data is distinguished into sociodemographic, behavioral and system criteria, sourced internally and externally. The purposes of tracking by the actors overlap. The criteria in Table 6 depicts what data the clients of the actors can acquire, not what data the firms behind the services collect and analyze.

We reached out to the three online shop vendors to inquire whether their software supports personalized price discrimination. Although it is not a core functionality of Shopify and WooCommerce, it is feasible through addons. For Shopify, multiple addons are available, for example, “Customer-Specific Pricing”, “Storakle”; or “Segment
Builder”, while for WooCommerce, we only found “WP Statistics”. Magento has advanced customer segmentation built-in, which can be enhanced by the addon “Prices per Customer”. Magento also allows the integration of external data, for example, the import of offline purchase data.

We found that all of the observed online retailers authored a privacy declaration, in which they depict data use for targeted advertising, statistics and server operations, but not for pricing purposes. After inquiring about the use of price discrimination via email, the majority of the contacted online retailers (7) denied usage of discrimination activities towards customers based on their personal data, while Otto.de, Bonprix.de, hm.de and Amazon.de did not answer the mail at all. Tchibo.de and Esprit.de state that they select consumers for discounts and gifts based on personal data and previous purchases, i.e. they engage in indirect price discrimination based on ex post discounts.

**Survey Results.** Due to space limitations, we do not report all descriptive statistics in detail. We only report them for selected items and only the top three values: n=61, age (M=24.84; SD=4.5), gender* (64% male; 18% female), device type (34% Windows; 31% Android; 13% iOS), and language (62% German; 26% English; 12% other). Most selected hobbies are travelling (46%), technology (41%) and sports (20%) and most visited shops include Amazon.de (43%), Zalando.de (28%) and Notebooksbilliger.de (25%). Top installed apps are YouTube (56%), Facebook (46%) and Instagram (38%). The first round with 26 international and PhD students showed a few insignificant prices differences. For four distinct users, prices differed by 9.00 EUR, 2.00 EUR, 0.05 EUR and 0.09 EUR on Otto.de, Zalando.de and Cyberport.de (last two). For the biggest difference, 9.00 EUR and 2.00 EUR, we compared the survey characteristics of the users with differing prices. Although, both of them were using mobile devices, the model and language of those devices were distinct from the other test persons. The second round of tests, in which 13 bachelor students participated, showed two differences in prices at 0.05 EUR and 0.60 EUR for Tchibo.de and Alternate.de. The last round with 22 bachelor students found only one instance of price differentiation by 0.10 EUR on Amazon.de. Overall, we tested 93 prices of products. Due to insufficient cases of price differentiation – presumably caused by transcription errors – the results do not substantiate the existence of systematic online price discrimination in the observed online retailers.

**Crawler Results.** Utilizing the crawler, we scraped the prices of 865 products with each of the five profiles, totaling 4325 checks in 27 sessions. We crawled all prices within the sessions at the same time, while the different sessions took place throughout one month. The shops Notebooksbilliger.de, Hm.de and Conrad.de are missing in the results due to their bot protection. They recognized our crawler and prevented the page load and scraping of prices. To add another electronics store, we crawled Mediemarkt.de instead. Apart from software errors, when the website layout changed and the scraper failed to extract the correct price, we did not find a single difference in prices in the entire data set for any of the selected products and shops.

---

* Eleven participants did not fill out their gender.
6 Discussion

Our results are consistent with previous studies and show that personalized price discrimination is not widely used in online shops. Nevertheless, as evidenced by the practices of the tourism industry [4, 11, 13], the implementation of personalized price discrimination is generally feasible. The first key condition to implement it, is sufficient information on the consumers to distinguish them into segments [19, 24, 25]. Based on the current data collection means, enough information on the consumers is available to derive such segments and to implement third degree price discrimination [17, 22]. Schleusener and Hosell [4] hypothesize a reason for the lack of personalized price discrimination may be the lack of technical expertise by German online retailers. However, our results show that all standard software packages provide the means to collect internal data and offer integrations with external data providers. Furthermore, if the shops have the expertise to develop a custom shop software, then they have the expertise to implement consumer segmentation and price discrimination. Hence, we argue that the lack of expertise is not the reason for the non-deployment of price discrimination practices.

Instead of technical obstacles, the reasons may be economical. The data providers do not disclose any prices on their websites, but acquiring external data may be costly compared to the value it provides, so acquiring external data may not be profitable for online shops. Yet, this does not explain the lack of price discrimination using internal data. As stated in the second condition for price discrimination [17], the retailer needs market power, or consumers will switch to other shops, if the same products are offered elsewhere. This is exacerbated by readily available online search. Another reason may be that employing personalized price discrimination can hurt the brand [4], because changing prices can lower the perceived price fairness of consumers [16].

Online retailers acquiring and exchanging customer data with third parties (including outside of EU legislation), without knowledge by the consumers may bring about dissatisfaction [10]. German citizens, compared to other countries and cultures, are sensitive to privacy issues, in particular since the EU-GDPR is effective [9]. The threat of legal action by consumer protection initiatives, and potential repercussions due to mistreatment or theft of data, may bar online retailers from implementing personalized price discrimination. It may be too risky to lose the trust of consumers and their brand value over the comparatively small profits, delivered by personalized price discrimination practices.

We conclude that the reasons for the lack of price discrimination in Germany are not technical, rather it must be for economic and political reasons. The digitalization, paired with intransparency of data collection and usage, especially by international players, brings rapid changes and means that price discrimination can start to occur any time. Thus, it is imperative to monitor potential price discrimination continuously. In a bigger project, our software can be improved with better monitoring abilities to provide ongoing assessment of personalized price discrimination.

We also received anecdotal references from colleagues and consultants, which we were not able to corroborate. Contrary to our findings, three retail and e-commerce consultancies claimed that their clients are engaging in online price discrimination. We
assume that either their clients have not rolled out any price discrimination yet, or they are in pre-testing, or they are not part of our selected shops, or our methods did not detect them.

7 Conclusion

In previous studies, German shops were not widely tested using surveys and automated scraping, and the ecosystem with its players was not systematically explored. Hence, our study makes a valuable contribution, as we investigate the prevalence of personalized price discrimination practices in the German e-commerce market through triangulation. By using three different methods, we reach a holistic picture of the ecosystem, its players and their use and exchange of data for the purpose of price discrimination. Contrary to our expectations, the results of our study do not substantiate that online retailers in Germany perform online price discrimination practices. In particular, the discourse analysis shows that the technical possibilities for price discrimination are widely available. Yet, both the survey and the crawler results show no evidence of price discrimination by the tested online shops. As discussed, we argue that this is due to political and economic reasons, not because of technical reasons.

The discourse analysis is limited to the official documentation, references and responses of the players. The survey and crawler are restricted by the sample size. Other limitations include the colocation of the survey participants and the usage of University Wi-Fi, since one of the characteristics that was proven to alter prices within consumers is the geographic location. Some participants made errors while transcribing prices. Several shops detected the crawler as a bot. We used free proxies, which might be blacklisted and negatively affect the price scraping, and the crawler directly accessed the product pages and did not traverse the shop websites. Nevertheless, our results match those observed in earlier studies. Future iterations of this experiment should be conducted with larger sample size and an improved crawler. The reasons for price discrimination can be further explored by conducting expert interviews.

Various institutions such as the OECD [3], the German Federal Ministry of Justice and Consumer Protection [4], consumer protection initiatives [5, 6], newspapers [7, 8], and most importantly, the German citizens and consumers recognized price discrimination as a major concern. Personalized price discrimination is a prime example of the algorithmization and digitalization of society and economy, and what effects data and algorithms may have on its people. Due to the digitalization and rapid changes of the online ecosystem, we encourage a continuous assessment of price discrimination practices in the future.

8 Acknowledgements

We thank Felix Murrenhoff, who programmed the original version of the crawler and collected data.
References