Balancing User Privacy and Personalization*

Malika Korganbekova
Northwestern University
Kellogg School of Management

Cole Zuber
Wayfair LLC

July 7, 2024

CLICK HERE FOR MOST RECENT VERSION

Abstract

Privacy restrictions imposed by browsers such as Safari and Chrome limit the quality of individual-level data used in personalization algorithms. This paper investigates the consequences of these privacy restrictions on consumer, seller and platform outcomes using data from Wayfair, a large US-based online retailer. Large-scale randomized experiments indicate that personalization increases seller and platform revenue and leads to better consumer-product matches with 10% lower post-purchase product returns and 2.3% higher repeat purchase probability. Privacy restrictions can distort these benefits because they limit platforms' ability to personalize. To evaluate privacy restrictions of interest, we (i) re-train the platform's personalization algorithm with lower-quality data and generate counterfactual recommendations, and (ii) next, we simulate consumers' search and purchase behavior under counterfactual recommendations using structural modeling. We find that two main policies imposed by Safari and Chrome disproportionately hurt price responsive consumers and small/niche product sellers. To address this, we propose and evaluate a probabilistic recognition algorithm that associates devices with user accounts without using exact user identity. Our findings demonstrate that this approach mitigates welfare and revenue losses significantly, striking a balance between privacy and personalization.

Keywords: consumer privacy, personalized recommendations, large-scale A/B tests, multi-session consumer search, Probabilistic Identity Recognition, Gaussian Processes

*Corresponding author: malika.korganbekova@kellogg.northwestern.edu. We are grateful to Eric Anderson, Brett Gordon, Robert Porter, and Caio Waisman for their invaluable mentorship and advice. We thank Guy Aridor, Vivek Bhattacharya, Andrey Fradkin, Samuel Goldberg, Rafael Greminger, Igal Hendel, Charles Hodgson, Yufeng Huang, Gaston Illanes, Garrett Johnson, Alibek Korganbekov, Aliya Korganbekova, Lakshman Krishnamurthi, Tai Lam, Alexander MacKay, Ilya Morozov, Andrey Simonov, Daniel Spulber, Artem Timoshenko, Anna Tuchman, Udayan Vaidya, Nils Wernerfelt, Hans Zhu, participants at the MSI Business Analytics conference at the Wharton School, and seminar participants at the Kellogg School of Management and Northwestern Department of Economics for their helpful comments and suggestions. We also thank anonymous reviewers in the MSI A. G. Clayton Dissertation Competition and Shankar-Spiegel Dissertation Competition for their helpful feedback. Special thanks to Wayfair's Data Science teams for their help securing data for this project, especially Alexei Alexandrov, Vinny DeGenova, Colin Gray, Matt Herman, Chia Hua Lee, Yasaman Khazaeni, Jeffrey Mei, and Ben Schroeder. All remaining errors are our own.

1. Introduction

Data collected about consumers by businesses underpin online personalization of product rankings (e.g., Amazon, Wayfair), movies (e.g., Netflix), music (e.g., Spotify), and other offerings. However, the growing availability and potential mishandling of individual-level data raises significant privacy concerns. Regulators are concerned that consumers have limited knowledge of how their data are used and of the inferences firms may draw about them based on this personal information.[1] To address these concerns, regulators and internet browsers have enforced limitations on online consumer-tracking. For instance, Safari may log consumers out of the website and clear their browsing history after seven days of inactivity, and Chrome plans to prevent cross-website user-tracking starting in 2024.[2] Under these restrictions, unless consumers log in, online platforms will no longer be able to recognize them continuously, and therefore will have an incomplete (fragmented) view of their browsing history. When platforms use this fragmented data in their personalization algorithms, it may affect the quality of recommendations made to users. Yet despite the growing number of privacy restrictions, there is little empirical evidence on their necessity and impact on the economic outcomes of consumers, sellers, and platforms. Such evidence is critical both to help design future data regulation and to help firms adapt their strategies to an increasingly privacy-conscious world.

The current paper has three goals. *First*, we use a large-scale field experiment conducted with Wayfair, a large US-based online retailer, to establish that personalized recommendations lead to better consumer-product matches and benefit both sellers and the platform. *Second*, we quantify the extent to which privacy restrictions distort personalization benefits. While the field experiment fully disabled personalization, under actual privacy restrictions platforms retain personalization but use fragmented (distorted) data. To assess the impact of privacy restrictions, first, we re-train platform's personalization algorithm with distorted data and generate counterfactual recommendations. Next, we simulate consumers' search and purchase behavior under counterfactual recommendations using a structural model. We estimate the model exploiting experimental variation in the data. We show that browser-induced privacy restrictions reduce the algorithm's prediction accuracy and result in lower-quality recommendations. The counterfactual simulations indicate that lower-quality recommendations decrease consumer welfare by 30% (from \$25 to \$18), and the adverse effects are more pronounced for price-responsive consumers. Moreover, smaller-revenue sellers and niche-product sellers[3] experience a disproportionate revenue loss of 8.6%, while larger sellers are relatively unaffected. *Third*, to help platforms mitigate the negative consequences of privacy restrictions, we evaluate probabilistic recognition algorithm proposed in Korganbekova and Zuber (2023). The machine learning algorithm probabilistically associates devices with unique user identities by exploiting

[1] Competition & Markets Authority UK Report

[2] See Google delays move away from cookies in Chrome to 2024.

[3] We use Deep Learning tools to identify products that are visually less similar to mass-market products.

detailed behavioral data and IP address information, even when the exact user identity is unknown. We show that the algorithm can recover up to 56% of welfare loss for the consumers and up to 73% of revenue loss for smaller sellers, providing a promising solution to mitigate the impact of privacy restrictions. Next we describe each of these findings in more detail.

To verify whether privacy policies should be a cause for concern, we first quantify the effects of personalization on consumers and sellers. To this end, we ran a large-scale field experiment where we randomly turned off personalization on product ranking pages on Wayfair. The experiment included 9 million consumers and ran for two years, from January 2020 to December 2021. Consumers in the treatment group saw personalized product rankings tailored to their browsing histories, while control group consumers saw non-personalized bestseller rankings.[4] We find that consumers in the treatment (personalized) group were 10% less likely to return a product post-purchase and were 2.3% more likely to repeat purchase a product in the same product category. These results suggest that consumers in the personalized group got better product matches than consumers in the non-personalized group. Moreover, unlike bestseller rankings, which highlight the most popular products, personalized rankings provide smaller sellers greater opportunity for prominence on the platform. Specifically, smaller-revenue and more niche sellers' products are 15% more likely to be shown on top of product ranking pages, and sellers earn up to 87% more revenue from personalized impressions compared to bestseller rankings. Overall, the experimental results suggest that personalization benefits both consumers and small and more niche product sellers.

Next, we quantify the extent to which browser-induced privacy restrictions distort personalization benefits. To do that, we need to evaluate consumer choices and compare seller and platform outcomes in two worlds: *a privacy-unrestricted world* in which platforms retain the ability to track consumers, and a counterfactual *privacy-restricted world* in which the platform continues to personalize but uses incomplete (fragmented) data.[5]

We leverage our access to the platform's personalization algorithm to generate counterfactual personalized rankings that would have been shown under privacy restrictions. First, we distort the individual-level data to mimic the impact of the privacy policy of interest. We focus on two privacy policies implemented by largest browsers: Chrome and Safari. For instance, to evaluate Safari's policy that clears browsing history after seven days of inactivity, we keep only the most recent seven days' worth of browsing data for each consumer. Next, we re-train the personalization algorithm with the new, distorted data input. The re-trained algorithm generates counterfactual product rankings, and we simulate consumers' response to these ranking using our structural model, which we describe next.

[4] Bestseller rankings are generated based on the aggregate historical popularity of the products.

[5] Note that the experiment turned off personalization completely; however, under privacy restrictions, the platform will continue personalizing, using fragmented data.

To evaluate how consumer choices will change under newly generated personalized rankings, we develop and estimate a demand-side model of multi-session consumer search. The model captures the main features of the online purchase funnel: viewing, clicking, and purchasing. Viewing depends on product rankings generated by the platform, and the platform can control which products are more prominent in the search results. Consumers can click only on the products they view. We model viewing and clicking separately because we have pixel-level data that allow us to distinguish between the two.

We describe consumer behavior as follows. Consumers know their preferences over the product characteristics observable from the ranking pages (e.g., prices, ratings, images),[6] but they have uncertainty regarding the unobservable product characteristics that they learn only after clicking on a product (e.g., reviews). While on the ranking page, consumers construct a utility index based on the observable characteristics of the products they have viewed. They have rational beliefs over the indices of the products they have not yet viewed. Further, they choose to click on a product if the maximum product index within the viewed set exceeds the *expected* maximum index within the non-viewed set of products. After clicking on a product, consumers learn the true utility of the product and update their beliefs about all the remaining products' quality. Next, consumers decide among (i) continuing to click, (ii) viewing additional products, (iii) purchasing a clicked product, or (iv) leaving. If they return to the website for subsequent sessions, the platform may personalize product rankings, and consumers follow the same search process under new rankings.

We model consumer's utility as a Gaussian Process over the observable product characteristics, which enables us to incorporate the key components of the model: viewing, clicking, purchasing, and learning. Moreover, the Gaussian Process specification also allows us to accommodate multi-session search: the consumers' posteriors from the previous session become their priors in the next session. We use experimental variation in the product rankings and pixel-level data to estimate consumer preferences, search costs, and learning parameters. We validate the model using data from a natural experiment: a short-term Chrome privacy policy change in 2020.[7]

In the counterfactuals, we fix estimated model parameters and simulate consumer search and purchase patterns under generated counterfactual personalized rankings that would have been shown under privacy restrictions. We focus on the two most prominent policies implemented by Safari and Chrome, as mentioned above.

We find that the policy that clears consumers' browsing history after seven days of inactivity (Safari) leads to a nearly 50% reduction in the prediction accuracy of the personalization algorithm. This results in lower-quality recommendations, which in turn reduce consumer welfare by 19% (from

[6]We project each image into two-dimensional space using Siamese Neural Network and UMAP, and directly use these vectors in the structural model.

[7]The policy was introduced at the beginning of the pandemic and affected users of a specific Chrome version.

$25 to $20) and decrease smaller sellers' revenue by 5.6%. For the Chrome policy that blocks cross-website tracking and affects consumers arriving from advertising channels (26% of the traffic), we find qualitatively similar but larger effects.

The counterfactual analysis yields several important insights. First, privacy policies disproportionately affect consumers who are more price responsive and have high search costs. These consumers either leave the website faster, as they do not see relevant product rankings, or tend to buy lower utility products. Second, the personalization algorithm tends to switch to emphasizing the popular products due to lack of data on smaller sellers. This leads to worse outcomes for the smaller/niche sellers. Third, while the platform loses revenue because of privacy restrictions, the impact on its instantaneous profit is relatively small. The reason is that larger sellers' products are easier and cheaper to ship than those of smaller sellers. Overall, these results highlight that privacy restrictions may hurt more vulnerable consumer groups and smaller sellers, which necessitates careful consideration of alternative policy design.

In the last part of the paper, we explore alternative strategies that platforms may take to mitigate the negative consequences of privacy restrictions. In Korganbekova and Zuber (2023), we propose using IP address information as well as consumers' detailed behavioral data to probabilistically recognize consumers even when the exact user identity is unknown. Structural model described above allows us to evaluate the algorithm. First, we obfuscate consumer identity as if it is not known. For each device, we predict which user the device belongs to. Next, we generate personalized rankings based on the predicted user's browsing history. Finally, we simulate consumer search and purchase process using the newly generated rankings according to the structural model. We show that this intervention can recover up to 56% of consumer welfare losses and up to 73% of small seller revenue. Therefore, even under privacy regulation, platforms can adapt their strategies to continue showing personalized content.

In this paper, to analyze the impact of privacy restrictions, we focus on one retailer: Wayfair. That brings up the question of generalizability, and whether our results are specific to Wayfair's personalization algorithm. We believe our results are generalizable to other online platforms. While different platforms use separate recommendation algorithms, these algorithms are ubiquitous and relatively standard. For example, major platforms often feature their algorithms as part of academic papers and workshops, which leads to adoption of similar algorithms by other platforms (see YouTube Covington, Adams and Sargin (2016); Pinterest Eksombatchai, Jindal, Liu, Liu, Sharma, Sugnet, Ulrich and Leskovec (2018); Wayfair Mei, Zuber and Khazaeni (2022); and Amazon Linden, Smith and York, 2003. It is also unlikely that the platforms can drastically improve their personalization algorithms under privacy restrictions, given that the main issue is not the algorithm but rather the inability of the platforms to recognize consumers.

2. Related Literature

This paper contributes to several strands of Economics and Marketing literature. First, the paper contributes to the privacy literature. A large number of empirical papers study the impact of European General Data Protection Regulation (GDPR) on firm outcomes (e.g., Aridor, Che, Nelson and Salz (2020), Goldberg, Johnson and Shriver (2019), Johnson, Shriver and Goldberg (2023), Zhao, Yildirim and Chintagunta, 2021). While GDPR's consent-based policy asks users to allow tracking, our paper focuses on policies that block online tracking *by default*. Our two-year, long-term experiment and unique rich data allow us to credibly estimate the effects of personalization on consumers, sellers, and the platform. In a related paper, Sun, Yuan, Li, Zhang and Xu (2021) ran a similar experiment with Chinese retailer Alibaba, but their experiment ran only for seven hours, preventing conclusions regarding long-term consequences of personalization. Moreover, we develop a structural model that allows us to mimic and evaluate not only full disablement of personalization, but to simulate a world in which platforms continue to personalize but use fragmented data.

Second, our findings extend the data fragmentation literature. Data fragmentation happens because of privacy restrictions, when instead of observing consumers' full browsing history, the platform is unable to track consumers over time and, therefore, observes only disconnected (fragmented) partial views of their browsing history. Prominent papers in this area include those by Coey and Bailey (2016) and Lin and Misra (2022), who examine the analytic form of the estimation bias caused by data fragmentation. Our paper empirically examines the changes in the predictive performance of personalization algorithms (i.e., training accuracy) and extensively describes the effects of data fragmentation on consumer, seller, and platform outcomes. Wernerfelt, Tuchman, Shapiro and Moakler (2022) study a similar phenomenon in the Facebook advertising context.

Third, we contribute to the personalization and consumer search literature. There is a large empirical search literature that estimates structural search models built on the tractable solution offered by Weitzman (1978) (e.g., Kim, Albuquerque and Bronnenberg (2010), Ursu (2018), Honka and Chintagunta (2017), Compiani, Lewis, Peng and Wang (2021), Morozov (2023), Seiler (2013), Seiler and Pinna, 2017). Our model has several important distinctions compared to extant work. First, motivated by the empirical patterns in the data, we disentangle viewing a product and clicking on it. The pixel-level data allows us to move away from one of the main assumptions in the aforementioned papers, i.e., that of consumers' full awareness. Consumers in our model have limited awareness of the products, and the platform's rankings affect their awareness given that platforms can make some products more prominent than others. Gibbard (2022) and Greminger (2022) were among the first papers to allow for limited consumer awareness. However, to the best of our knowledge, our paper and Choi and Mela (2019) are the first to use data enabling to separate treatment of viewing and clicking actions. Second, we allow for consumer learning in our model. The decision to model consumer learning was driven by the empirical patterns in the data similar to the ones observed in Bronnenberg, Kim and Mela (2016) and Hodgson and Lewis (2022). Incorporating learning into the model requires us to resort to a near-optimal consumer search policy, as opposed to the optimal policy suggested by Weitzman (1978) under stricter assumptions. Given that we model search via Gaussian Processes, we use a near-optimal heuristic for optimal search policy: Upper Confidence Bound algorithm (Auer, 2002). We prove the finite regret bounds of the algorithm in our setting in Appendix B.

The remainder of the paper is organized as follows. Section 3 describes the empirical setting and data. In Section 4, we describe the experiment conducted with the platform and show the experimental results. In Section 5, we introduce the model and explain the estimation details. Section 6 shows main counterfactual results. Section 7 concludes.

3. Empirical Setting

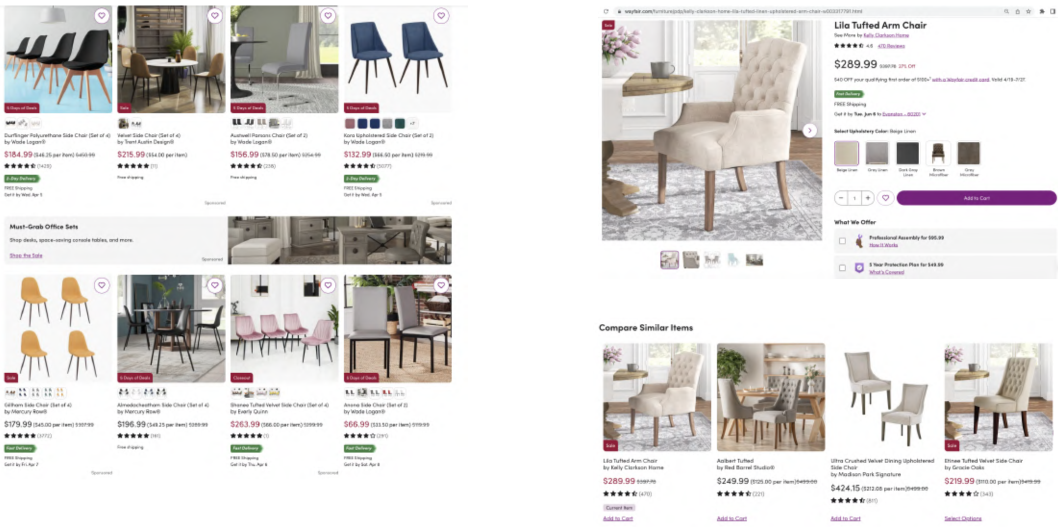
To study the impact of privacy restrictions on platforms, we have collaborated with Wayfair, a large US-based online retailer of furniture and home goods. Wayfair does not own sellers' inventory, instead acting as an intermediary between product sellers and consumers. The platform owns and maintains the website interface and all personalization algorithms. In this section, we provide an overview of the platform's interface and discuss how privacy restrictions may influence consumer recognition within the platform's ecosystem.

Suppose consumers arrive at wayfair.com for the very first time and search for the keyword 'dining chairs'. Consumers are taken to a product category page, where the platform displays an ordered list of dining chairs (hereafter, "ranking page" or "product rankings"). Figure 1a shows an example of a ranking page. These pages typically feature 48-96 products allocated in a matrix form. Consumers can observe product images, prices, ratings, and numbers of ratings on the ranking pages. During the initial session, the platform does not recognize consumers and lacks information about their preferences. Therefore, the platform presents consumers with non-personalized bestseller product rankings. These rankings are based on the overall historical popularity of products and are not customized to the consumers' browsing history.

Clicking on a particular dining chair on the ranking page takes consumers to the product page where they can find detailed product descriptions, e.g., chair dimensions, materials used, assembly requirements, and other specifications (Figure 1b). Many product pages include product reviews and ratings. Consumers can read these reviews for insights into the chair's quality, comfort, durability, and overall customer satisfaction. On the product pages, consumers are presented with additional product recommendations known as *Compare Similar Items* (Figure 1b). These recommendations are not personalized based on the consumers' browsing history but rather represent a set of products

FIGURE 1: Ranking pages and product pages

(A) Product ranking page (B) Product page



that are comparable to the product currently being viewed. Consumers can opt to click on one of the four recommended products within the widget, which will redirect them to the respective product's page.

Subsequently, consumers are presented with three options: (i) to continue searching by clicking or scrolling further, (ii) to make a purchase of the best product they have come across thus far, or (iii) to exit the website. Importantly, consumers retain the flexibility to revisit the website and continue the search process at any given point in time.

In subsequent visits, consumer recognition becomes crucial. Suppose that during the initial session, a consumer exclusively clicks on blue chairs or applies a filter to see only blue-colored chairs. The website sets a cookie - a small text file - on the consumer's browser that keeps track of the specific types of products consumer clicks on or types of filters she applies.[8] These cookies, which are set by the wayfair.com website and are only readable and writable by this website, are called **first-party cookies**. They allow websites to store valuable information about consumers' browsing history on their domains.

Next time the consumer re-visits wayfair.com, the website checks for the presence of any cookies associated with the wayfair.com domain on her browser. By examining the first-party cookie files, the platform identifies that the consumer had exclusively clicked on blue dining chairs in the past. The platform's personalization algorithm incorporates the consumer's preference for blue chairs and presents personalized rankings that prioritize blue chairs at the top of the results.[9,10] This

[8]Platforms can set cookies for many different consumer actions, such as sorting, adding a product to the basket page, and others.

[9]The layout of the ranking page looks exactly the same as the non-personalized pages, but products are ranked such that more relevant products to a consumer gain more prominence on the ranking page results.

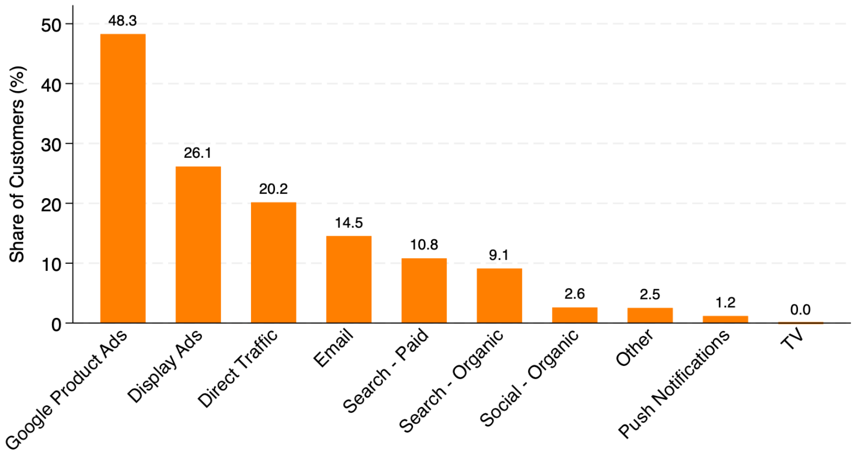
[10]It is worth noting that, at the time of this research, the platform employed single-category personalization, i.e. browsing data from other categories was not utilized in determining the rankings.

personalized approach may reduce search frictions and enhance the consumer's browsing experience on the website. However, certain restrictions disallow first-party cookie tracking.

Privacy restrictions associated with first-party cookies. In 2019, Safari introduced a new version of its browser that includes Apple's Intelligent Tracking Prevention (ITP),[11] which automatically removes first-party cookies after a week of user inactivity. This feature not only logs consumers out but also requires Wayfair to reset first-party cookies.[12] As a result, if the consumer returns more than 7 days after the previous session, Wayfair loses access to the browsing information stored in the expired cookie, as it becomes inaccessible.

Initially, first-party cookie tracking restrictions impacted only those first-party cookies that were used in third-party contexts, i.e., for cross-website user tracking. The restrictions had limited impact on the first-party cookies used within the website itself (see Figure E4 and Table E11). However, since 2022, Safari has been strengthening the regulation and blocks first-party cookie tracking more often, which affects the platform's ability to recognize consumers.[13]

FIGURE 2: Traffic sources



Notes. This figure shows the distribution of traffic by different arrival channels, e.g. Direct Traffic, Google etc. The total doesn't sum up to 100, because one consumer can arrive through multiple channels. 26.1% of consumers arrive from Display advertising, and a large share of consumers arrive either from Google or from Direct Traffic. Unless consumers login, platform relies on third-party cookies for display ad recognition and on first-party cookies for the consumers arriving directly to the website.

Advertising is another critical aspect where platforms' ability to recognize consumers is at risk. The platforms show re-targeted ads on third-party websites (display ads) to encourage consumers to re-visit the website. In our data, approximately 26% of traffic arrives from display advertising, which is a large portion of traffic (Figure 2).

[11]See Apple's Intelligent Tracking Prevention (ITP).

[12]Technically, the cookies that are set to expire in 7 days are the ones that are set from JavaScript. However, majority of cookies are set through JavaScript.

[13]See the Apple's secret Safari cookie crackdown article for more information.

Display advertising works via **third-party cookies** that facilitate cross-website user tracking. To illustrate, let's say a consumer leaves Wayfair and visits weather.com. On weather.com, third-party companies like Wayfair or data aggregators can place their cookies. As a result, when weather.com loads, it reads the Wayfair cookie and displays ads that align with the consumer's browsing history on wayfair.com (Figure 3).[14] Third-party cookies are used both to show advertising and to recognize consumers if they click on an ad and re-visit the website.

FIGURE 3: Wayfair ads on weather.com

Notes. This figure shows an example of the retargeted advertising that was served through third-party cookies. The screenshot was taken by the authors on weather.com.

Privacy restrictions associated with third-party cookies. Safari and Firefox already block third-party cookie tracking,[15] and the largest web browser Chrome[16] plans to block them in 2024. Blocking third-party cookies limits platforms' ability to serve advertising and to recognize consumers. Under these restrictions, all third-party requests reset users' cookie files which means that platforms will not be able to recognize users and pull their browsing history.

Notably, when consumers voluntarily log in during each visit, the platform can recognize them without relying on cookies. However, our data reveal that approximately 37% of consumers choose to log in, leaving the platform to rely on cookies for the recognition of 38% of consumers (Table 1), which corresponds to millions of consumers.[17]

To summarize, platforms use first- and third-party cookies to recognize consumers and to

[14] In reality, the online ad system works in a more complex way. Third-party vendors form coalitions to map a user's identifier from a demand-side platform to a data management platforms. This process is called cookie syncing and it is used by AdTech platforms, demand-side platforms, data-management platforms (DMPs), ad exchanges, supply-side platforms (SSPs) and various other data providers. This means that the user data is exchanged across different platforms, which creates significant privacy concerns. See DMPs and Cookie Syncing for details.

[15] For instance, Safari's Intelligent Tracking Prevention browser setting prevents cookies being read in a third-party context. In the first version of ITP, Apple limited third-party cookie reads to a 24-hour window, but rolled complete block later. Firefox followed Safari and by blocking third-party cookies in Version 50+ of the browser.

[16] Chrome is the leading global browser with 62.85% share as of June 2023. See Browser market shares statistics.

[17] To further examine the breakdown between first-party and third-party cookie reliance, a rough estimation can be made. Among the 26% of display advertising traffic, assuming an average of 38.46% recognition through cookies, approximately 10% of the traffic depends on third-party cookie recognition. Similarly, if 37% of the traffic from Google Product Ads and Direct traffic log in and 38.46% are recognized through cookies, it can be inferred that first-party cookies account for 26.3% of traffic recognition (=48.3% (Google) + 20.2% (Direct)) × 38.46% (cookie-recognized).

Table 1: Recognition rates by device types

	Overall	**Desktop**	**Mobile Site**	**App**
Logged in (%)	37.62	33.95	22.63	84.06
Cookie-recognized (%)	38.46	44.22	48.69	8.43
Not recognized (%)	23.92	21.83	28.68	7.51
Total	100.00	100.00	100.00	100.00

Notes. This table reports login and cookie-recognition rates by different devices. From the data, for each consumer who searched in dining chairs category, we determine whether she was logged in at least once, or was cookie recognized or never recognized. We then calculate corresponding shares overall, on desktop, mobile site and Wayfair app.

personalize user experience. Approximately 38% of traffic is recognized via first- and third-party cookies in the platform we collaborate with. Privacy policies block platforms' access to the cookies and reset unique user identifiers. This threatens platforms' ability to recognize consumers and to personalize their experience.

This paper aims to assess the impact of privacy restrictions on consumers, sellers, and the platform. Additionally, we explore alternative strategies that the platform can adopt to mitigate the consequences of these privacy restrictions. But first, it is crucial to quantify the potential losses stemming from the absence of personalization. Therefore, after describing the data in the next subsection, we quantify the effects of personalization on consumers, sellers and the platform.

3.1. Data

In this subsection, we describe the data that was generously provided by Wayfair. Our main sample consists of 30 million consumers who browsed the platform in 2018-2022. Our data have several important and unique components.

Clickstream data. We have access to Wayfair's full high-frequency pixel-level clickstream data that tracks consumers' actions on the website. The main data span the two-year experimental period from January 2020 to December 2021. We also have access to the historical data from 2018 to 2020, which we use to evaluate consumers' and sellers' historical outcomes. The data are at the device - customer id - URL - action timestamp level. Thus, each row in the data represents a single action, such as click, scroll, tap, hover, or zoom, taken by a consumer on a given device. Consumers' behavior is captured across all platforms (e.g. desktop, tablet, mobile) and devices. We view the existing clickstream data as a *relatively* complete description of consumers' browsing history. Note that every consumer can be associated with multiple devices and multiple browsers (see Table E10). The platform stitches together corresponding devices and browsers and links them to a unique consumer identifier for each consumer to the best of their ability. If consumers login, then it is straightforward for the platform to link browsing history. Otherwise, the platform relies on cookie files to recognize consumers and to build a *relatively* complete view of consumers' browsing history across sessions and devices.[18]

Pixel-level data. A unique aspect of our data is that for each consumer, we observe which products appeared on a consumer's screen at each point in time. Thus, we can restore consumers' scrolling behavior and explore which part of each page consumer *viewed.* The data are crucial to assessing consumers' awareness of products, and it is similar in spirit to eye-tracking data.

Login and traffic source data. At each point in time, we observe whether the consumer was logged in, the referrer URL (website the consumer was on before arriving on a focal webpage), and the channel consumer used to arrive on the website, such as Google ads or Direct Traffic.[19]

Rankings and recommendations data. For a given URL in the clickstream data, we can re-create the layout of the webpages that were shown to the consumer. Most importantly for our setting, we can recover the ordered product rankings and recommendation widgets that were shown to each consumer. For the product rankings, we have determined which products were personalized and which ones were non-personalized on the ranking pages. Moreover, for all product pages, we kept historical scores outputted by a recommendation widget algorithm that allows us to determine the set of products that were shown on each product page as part of additional product recommendations.

Transactions data. For each consumer, we observe the set of products they purchased (if any), corresponding prices of the products, and the indicator for whether the consumer returned the purchased item. We use these data to evaluate consumer choices and to proxy for the quality of product matches. In particular, we use product returns and repeat purchases to evaluate consumers' satisfaction with the purchased products.

Prices and wholesale costs. We observe a daily panel of each product's prices, wholesale costs set by the seller, any discounts/allowances provided by the seller and shipping costs.[20] We use these data to calculate product markups and margins to assess the platform's profitability.[21]

Seller-level data. For each seller on the platform, we observe the number of products they carry, the revenue earned for each product, and all the historical data on the wholesale costs and revenue. We use these data for seller-heterogeneity analysis.

Product characteristics. We have data on product characteristics that are observable to the consumer from the ranking pages, i.e., prices, product rating, and number of ratings. Moreover, we also use data on products' style, material, chair width, height, etc. in our descriptive analysis.

Image data. We have access to images of products in select categories, namely, dining chairs, sofas, and ottomans.[22] We use these images to train a Deep Learning algorithm, namely Siamese Neural

[18]The platform backfills customer identifiers whenever possible. It means that they retroactively add the unique customer identifiers in case consumer logs in, for instance.

[19]These data were used to plot Figure 2.

[20]Wayfair uses a dynamic pricing algorithm and takes wholesale cost and any discounts set by the sellers as an input in the pricing formula. See Wayfair pricing.

[21]For data sensitivity reasons, we only show relative changes in the profits whenever applicable and never reveal actual markups.

[22]We utilize multiple product categories to collect sufficiently large training data and enable the model to learn

Network (Bell and Bala, 2015), which allows us to represent every image as a 512×1 vector (image embedding). We then calculate cosine similarities between the image vectors to determine niche products and run a heterogeneity analysis based on that. We tag products as niche if the cosine similarity of the product's image to all the remaining images in the same category is low, i.e., the product differs from other products. We also use these image embeddings in our structural model. Details of the Deep Learning model training are given in Appendix G.

To reduce data dimensionality, we focus on one product category: dining chairs.[23,24] We chose the dining chairs as a category of interest for several reasons. First, it is a large category at Wayfair that contains more than 30,000 products and is a big-ticket category where the median and mean product prices are $349.99 and $431.99, respectively (Table 2). There is also significant horizontal differentiation in the category because chairs differ in color, style, upholstery material etc., which creates heterogeneity in pricing (Table 3). Therefore, consumers must engage in extensive search: median consumer (among both purchasing and non-purchasing consumers), arrives for 2 sessions and spends 15 minutes on the website. Moreover, it takes some time for the consumers to decide which

TABLE 2: Session summary statistics

	Observations	**Min**	**Mean**	**Median**	**Max**	**St.Dev.**
# of sessions	635,267	1.00	3.53	2.00	29.00	4.72
Session Duration (minutes)	635,267	0.04	35.82	15.21	273.00	50.81
Interarrival time (days)	274,745	0.00	14.09	8.00	59.00	15.39
# of products	635,267	0.00	6.32	2.00	63.00	10.40
Price ($)	35,873	4.46	431.99	349.99	2999.99	291.27

Notes. This table reports summary statistics of consumer searches in dining chair category.

product to purchase, which is why the median interarrival days is 8 days. Thus, this is a category where privacy restrictions may play an important role because at the next consumer visit, the cookie files may already be deleted.[25] Finally, this is the category where consumers tend to repeat purchase items. For instance, consumers may buy 1-2 chairs and subsequently purchase additional chairs. The repeat purchase behavior is important in assessing the customer satisfaction with the product.

Overall, our data are unique because we observe consumers' search behavior at a very granular level, including their pixel-level actions. Moreover, we know whether consumers logged in or were cookie recognized, as well as the channel they used to arrive to the website, e.g., advertising. The fact that we observe detailed consumer behavior and the sellers' and platform's price and cost data allows us to credibly evaluate the impact of privacy restrictions on all main parties on the platform:

diverse image aspects, including shape differences and other characteristics, across various product categories.

[23] We explicitly state in the rest of the paper if we switch to analyzing all of the data in a particular analysis for statistical power purposes.

[24] We deliberately selected a category that experienced minimal Covid-related impact, while consciously avoiding categories heavily affected by the pandemic such as office furniture or kitchen appliances.

[25] Recall that the Safari policy will reset first-party cookies in 7 days.

Table 3: Heterogeneity in dining chairs

Upholstery Material	Mean Price ($)	Median Price ($)	St.Dev. Price ($)	% of Products
Genuine Leather	739.10	579.00	565.93	10.18
Fabric	547.77	419.99	412.51	43.92
Faux Leather	472.98	369.99	354.47	26.63
Velvet	450.88	334.99	362.24	19.06
Metal	228.12	243.74	73.75	0.03
Wood	226.89	225.99	75.92	0.07
Plastic / Acrylic	161.88	180.49	69.58	0.08
Wicker / Rattan	147.60	141.80	53.33	0.04

Notes. This table reports the distribution of prices in dining chairs by upholstery material type. The table is sorted in descending order by first column (mean price).

the platform itself, sellers, and consumers.

4. Experimental Results

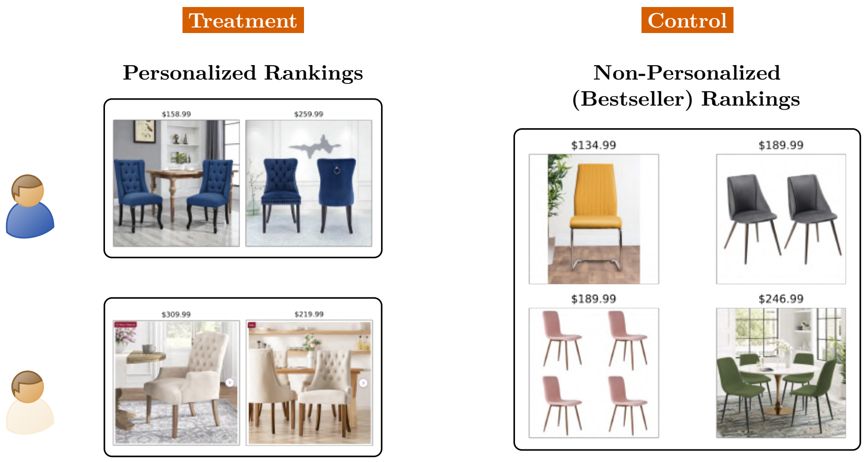
To quantify the effects of personalization on consumers, product sellers, and the platform, we ran a large-scale field experiment, where we randomly turned off personalization on product ranking pages on Wayfair. This corresponds to the full disablement of personalization on the platform for a random sample of consumers. The experiment included 9 million consumers and ran for two years from January 2020 to December 2021. Consumers in the treatment group saw personalized product rankings tailored to their browsing histories, while control group consumers saw non-personalized bestseller rankings. Bestseller rankings are generated based on the aggregate historical popularity of the products, i.e., even if the platform had consumers' browsing history, the data were not used.

One could argue that personalization should benefit platforms, otherwise they would not run the algorithms. However, the effects of personalization on the consumers and product sellers are unclear. On the one hand, platforms have the incentive to provide better consumer-product matches to nurture long-term consumer loyalty. On the other hand, regulators are concerned that platforms may prioritize their own commercial interests over better consumer product-matches. For instance, the platform could show higher margin items to the consumers they view as less price elastic.[26] In the next subsections, we empirically investigate which mechanism prevails.

To illustrate the experimental variation, suppose there is a consumer who clicked on blue dining chairs during the first session and a consumer who clicked on the white chair (left-hand side of Figure 4). When consumers re-visit the website, if they were randomized into the **treatment** (personalized) group, the platform would serve them personalized rankings that are generated by the personalization algorithm. The inputs to the algorithm are the consumers' clicks, add-to-carts, and purchases within a product category. The algorithm itself is a Deep Learning-based algorithm that learns the similarities

[26] See Algorithms: How they can reduce competition and harm consumers 2021 (CMA) for more details.

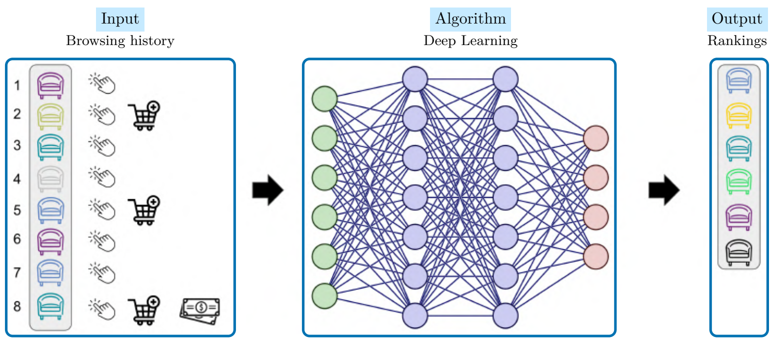
FIGURE 4: Experiment Design



between the co-clicked (co-purchased) products and outputs the set of personalized rankings (see Figure 5).

If the consumers were randomized into the **control** group, the platform serves non-personalized bestseller rankings regardless of the type of chairs they saw previously (right-hand side of Figure 4).

FIGURE 5: Personalization Algorithm



Note that during their very first session on the platform, consumers in the treatment and control group will see the same set of recommendations. The reason is that the platform does not yet have data to use for personalization.[27] Thus, the results in this section should be treated as Intent-To-Treat (ITT) rather than standard Average Treatment Effects (ATE).

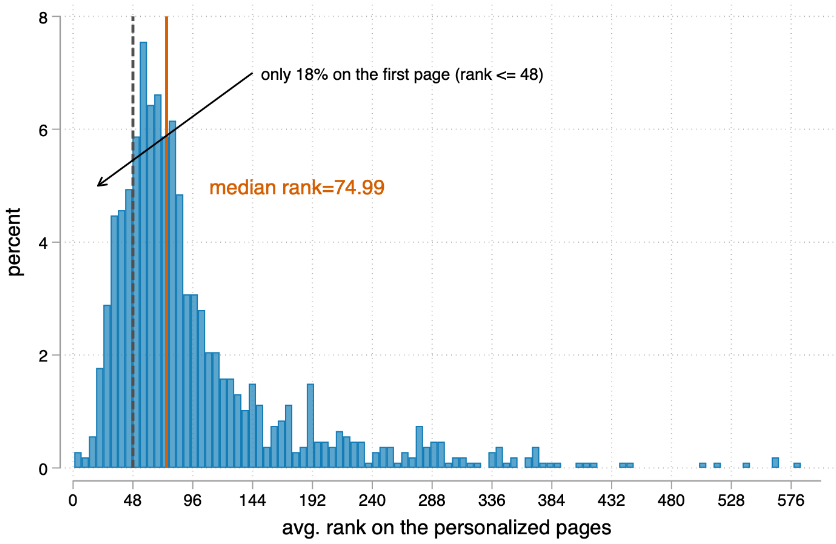
We did randomization checks based on both consumer groups' demographic data, as well as their historical search and purchase behavior. Table C5 shows that there are no significant differences between consumers in treatment and control groups, which validates correct randomization in the experiment.

[27]There is also a rule on the platform that prohibits changing rankings within the 30 minute interval during a session, not to confuse consumers. Therefore, the rankings do not change within a session.

4.1. Experimental changes in product rankings

First, we illustrate the impact of the experiment on product rankings. In the personalized (treatment) condition, products that were predicted to be more relevant to the consumer were placed higher on the ranking pages. To validate the experiment, we take the top 50 most popular products in the dining chair category overall and show the distribution of their ranks in the experimental (personalized) condition. In the bestseller non-personalized condition, these products would have been shown at the top of the product ranking results. However, Figure 10 shows that in the personalized condition the median rank of the top 50 product is 74. Given that there are 48 products on each ranking page, this means that in the personalized condition 50% of the time a top product is shown on the second page. This makes sense because the platform typically fills the first 72 positions (1.5 pages of the ranking results) with the personalized results. Note that it is completely possible for the top products to be personalized in case there are consumers who click on these or similar products.[28]

Figure 6: Rankings of Top 50 bestselling products in personalized condition



Notes. This figure shows the ranks of the Top 50 products when shown on personalized pages. Despite being most popular products, the median product is shown on position 74, on the second page of the ranking results. The reason is that other products that are more relevant to consumers' individual tastes take higher positions in the ranking results.

Meanwhile, smaller sellers jumped in the ranking results by a median of 4 pages or 241 positions, as illustrated in Figure 7. Thus, the experiment changed the rankings of the products as expected, which further validates the experiment.[29]

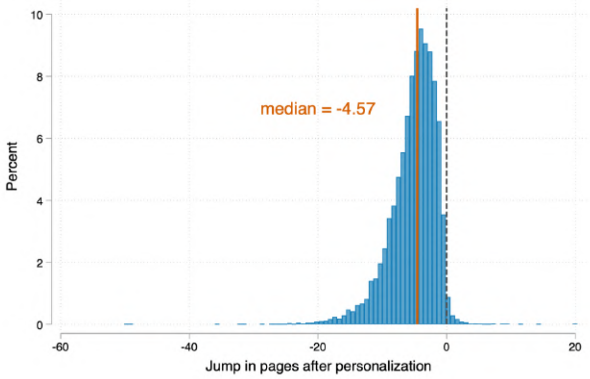
To summarize, the experiment changed product rankings quite substantially, where bestseller

[28]Appendix H describes the personalization algorithm in more detail.

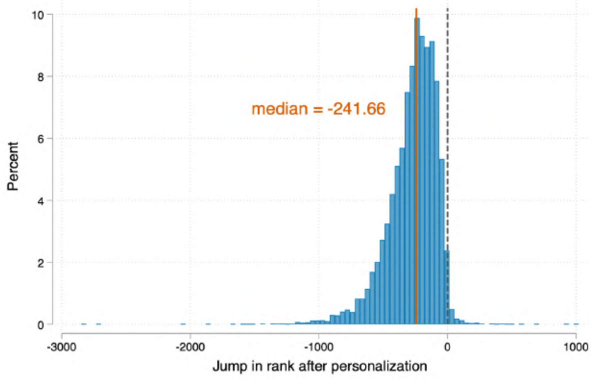
[29]Moreover, Figure D3 shows that the treatment intensity, i.e. the share of the personalized products on the ranking pages increases as the platform collects more and more data about a consumer. Note that while the platform tries to populate the first 72 positions with personalized products, it may not always be feasible due to factors such as product unavailability or the algorithm finding fewer than 72 relevant products.

FIGURE 7: Changes in ranking results in the treatment condition

(A) Δ ranking pages



(B) Δ product ranks



Notes. This figure shows the jump in the pages and overall ranks of less popular products as a result of personalization. Figure (A) shows that median product jumped four pages higher, and Figure (B) shows that the overall rank decreased by 241 positions.

products were shown on the second page of the ranking results or even further. Meanwhile, less popular products gained the opportunity to be more prominent on the website. Next, we explore how these experimental changes affected consumers, sellers, and the platform.

4.2. EFFECT OF PERSONALIZATION ON CONSUMERS AND THE PLATFORM

In this subsection, we estimate the effects of the experimental changes in the rankings on consumer and the platform outcomes in the dining chairs category. Since there is experimental variation in the treatment assignment, we estimate the treatment effect by regressing the outcome variable for each consumer (y_i) on the randomized treatment assignment ($treatment_i$) as follows

$$y_i = \alpha + \beta treatment_i + \varepsilon_i \tag{1}$$

where y_i is the outcome variable of interest, such as clicks or revenue, $treatment_i$ is the treatment indicator equal to 1 in case consumer i is randomized into the personalized ranking group, and 0 otherwise.

Table 4 shows that personalization does not affect the probability of clicking, but does increase the probability of adding to cart by 1.1%, the probability of basket page visit by 1.4% and the purchase (conversion) probability by 1.4% (Columns 1-4 in Table 4). Consumers in the personalized group bring more revenue (+2.1%) and more profit (+1.5%), which is partially driven by a larger number of purchase instances (+2.4%) (Columns 5-7). Table D6 in the Appendix shows the effects for the full experimental data, i.e., all 9 million consumers; and Table D8 shows the experimental results among consumers who had a browsing history before the experiment. The results are consistent across different samples.

Personalization is clearly beneficial to the platform through higher revenue and profit. On

TABLE 4: Effect of personalization on consumer and platform outcomes

	Logistic				*OLS*		
	(1) Clicks	(2) Add-to-cart	(3) Basket page	(4) Converted	(5) Log(Revenue)	(6) Purchases	(7) Log(Profit)
Personalized	0.002	0.011**	0.014***	0.014**	0.021***	0.024***	0.015**
	(0.012)	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.006)
Intercept	2.988***	0.246***	0.148***	-0.870***	1.947***	1.095***	–
	(0.008)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)
Observations	635,267	635,267	635,267	635,267	635,267	635,267	635,267

Notes. This table reports the output from the estimation of equation 1. Data is at the consumer-level. Columns (1)-(4) report the logistic specification and Columns (5)-(7) report the OLS specification results. Robust standard errors in parentheses. The intercept in profit Column (7) is hidden for data sensitivity reasons. Statistical significance: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

the consumer side, purchase metrics exhibit a positive trend, however it is not yet clear whether consumers benefit from personalization. One concern that regulators have is that personalization may drive consumers to purchase higher margin items. To test this hypothesis, we analyze consumers' purchase outcomes. We estimate Equation 2 on the consumer-purchased product-level data. Table 5 shows that purchasing consumers in the personalized group buy 0.5% higher priced items, and they buy more items (+0.9%) than the purchasing consumers in the non-personalized group. This leads to the platform earning 1.5% higher revenue from the purchasing consumers. However, contrary to the regulators' concerns, we do not find significant differences in the platform profits from purchasing consumers.[30] These results suggest that while consumers in the personalized group purchase slightly more expensive products and drive revenue, personalization algorithm per se does not lead consumers to higher margin items.

$$y_{ij} = \alpha + \beta treatment_i + \varepsilon_{ij} \tag{2}$$

TABLE 5: Effect of personalization on consumer outcomes: purchase outcomes

	Purchase Outcomes			
	(1) log(price)	(2) log(quantity)	(3) log(revenue)	(4) log(profit)
Personalized group	0.005***	0.009***	0.015***	0.004
	(0.001)	(0.001)	(0.003)	(0.003)
Intercept	6.022***	2.075***	7.343***	–
	(0.001)	(0.001)	(0.002)	(0.002)
Observations	2,022,708	2,022,708	2,022,708	2,022,708

Notes. This table reports the output from the estimation of equation 2. Data is at the consumer-purchased product level. The intercept in Column (4) is hidden for data sensitivity reasons. Statistical significance: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Next, we explore the quality of the product matches between personalized and non-personalized groups. To proxy for the match quality, we use detailed data on *post-purchase product returns and repeat visits*. Note that consumers tend to repeat visit and to repeat purchase products in the dining

[30]The profit intercept is hidden to adhere to the legal agreement with the platform and prevent the disclosure of their margins.

chairs category. We observe in the data that consumers purchase a set of 1-2 chairs and then repeat purchase after some time. Our hypothesis is that they test the chairs before purchasing the full set. We estimate the logit specification of Equation 1 to identify the probability of repeat purchases. The outcome variables y_i are the indicator variables for whether the consumer repeat purchased 7, 30, 90, 150, 365, or 500 days after the first purchase. Columns (1)-(6) of Table 6 show the estimation results. We find that the repeat purchase probability is similar in the personalized and non-personalized groups 7 days after the first purchase. However, 30 to 500 days after the first purchase the repeat purchase probability increases by 2.2-3.9%. Table D7 shows the same regression results for the full set of consumers across all product categories in the experiment and confirms that the results are similar. The fact that consumers in the personalized group are more likely to repeat purchase a product within the same product category suggests that they are satisfied with the initial product match.

TABLE 6: Effect of personalization on repeat visits and product returns (dining chair consumers)

	Repeat purchases						*Returns*
	(1) 7 days	(2) 30 days	(3) 90 days	(4) 150 days	(5) 365 days	(6) 500 days	(7) product returns
Personalized	0.007	0.022*	0.025***	0.027**	0.023*	0.039***	0.003
	(0.015)	(0.012)	(0.008)	(0.011)	(0.012)	(0.015)	(0.006)
Personalized× Personalized product							-0.103***
							(0.022)
Intercept	-1.663***	-0.726***	0.218***	0.214***	0.726***	0.971***	-2.727***
	(0.010)	(0.008)	(0.006)	(0.008)	(0.008)	(0.011)	(0.004)
Observations	136,643	136,623	268,058	136,585	136,568	90,480	1,898,251

Notes. This table shows the effects of personalization on repeat purchases and product return rates. Columns (1) - (6) are estimated using logit version of 1. Data are at the consumer level. Column (7) is the estimation of Equation 3. Data are at the consumer-purchased product level. Consumers in the personalized group might buy the item that was part of the organic rankings and wasn't personalized to them and we control for that by interacting the treatment dummy with the indicator for whether the product was personalized. For statistical power, we've included all consumers who were shopping in dining chairs category and their visits to the same marketing category, i.e. dining chairs, chairs, Each column represents the set of people who purchased a dining chair and we check the probability they will purchased in 7, 30, 90 etc. days. Robust standard errors in parentheses. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

To further investigate the quality of product matches, we show that consumers in the personalized group are 10% less likely to return the product post-purchase (Column 7 of Table 6). The specification we use is the logit form of the following equation:

$$y_{ij} = \alpha + \beta_1 treatment_i + \beta_2 treatment_i \times personalized\ product_j + \varepsilon_{ij} \tag{3}$$

where y_{ij} is the dummy variable indicating whether consumer i returned purchased product j; $treatment_i$ is the indicator for the treatment assignment of consumer i where $treatment_i = 1$ if the consumer is in the personalized group, and 0 otherwise; $personalized\ product_j$ is the indicator for whether consumer bought a product that was *personalized* to her. Recall that typically the first 1-2 pages of product rankings feature personalized products, and the rest of the products are non-personalized to consumers' browsing history. Therefore, it is possible that consumers in the personalized group buy a product that was not personalized to them. We find that consumers who were in the personalized group *and* bought a product that was personalized to them were 10% less likely to return the purchased product (Column 7 of Table 6). The results are robust both in the dining chairs category and in the full sample of experimental consumers (see Table D7).

Overall, the results above suggest that consumers benefit from personalization through better product matches, which we measure via higher repeat purchase probability and lower probability of returning a product.

So far, we focused on consumers' purchase outcomes, but the entire search process is important. Therefore, next we explore whether consumers in the personalized group (independent of purchase status) experience lower search costs than the consumers in the non-personalized group. We measure search costs by exploring consumers' filtering behavior and the time they spent searching. Estimating Equation 1 with the search-related outcome variables, we find that consumers in the personalized group are 1.9% less likely to filter a ranking page, and conditional on filtering they apply marginally fewer filters. Moreover, we find that consumers in the personalized group spend 3.6% fewer days searching for a product compared to the consumers in the non-personalized group. These results suggest that consumers in the personalized condition incur less search costs on the platform.

TABLE 7: Effect of personalization on search costs

	Logistic	*OLS*		
	(1) Filtered (0-1)	(2) Log(# of filters applied)	(3) Position of the filter	(4) Log(days till purchase)
Personalized	-0.019***	-0.007*	-0.021**	-0.036*
	(0.007)	(0.004)	(0.010)	(0.020)
Intercept	-1.716***	1.435***	2.768***	2.409***
	(0.005)	(0.002)	(0.007)	(0.014)
Observations	635,267	140,574	179,628	43,110

Notes. This table shows the output from the estimation of equation 1. Robust standard errors in parentheses. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

The takeaway from this subsection is that both consumers and the platform benefit from personalization. Moreover, the increase in the platform's profit is driven by the repeat purchase behavior of consumers rather than diversion towards high margin products. These results suggest that the platform and consumers' incentives are aligned in the sense that better consumer-product matches lead to better outcomes for both consumers and the platform.

4.3. Effects of personalization on product sellers

Next, we explore the effect of personalization on product sellers. To understand how personalization affects different sellers' products, we proceed in several steps.

Small-revenue and less experienced sellers. First, we use pre-experiment data from January 2018 to January 2020 to calculate each product's historical popularity, i.e., revenue earned and quantity sold. We also ranked products by their historical revenue within a category of products, where rank 1 means that the product is a best-seller, rank 2 means the product is the second best, etc. Next, we construct a consumer-product-seller level dataset to investigate whether consumers in the personalized group buy more or less popular products. Table 8 shows the results of estimation of Equation 2. We find that consumers in the personalized group purchase products that have 6% less historical scaled revenue (-\$251 with the intercept of \$4,011), have 5% less quantity sold (2.4 with the intercept of 49.4), and are 94 positions ranked lower (less popular) compared to the consumers in the non-personalized condition. These results suggest that personalization leads consumers to purchase less popular products.

Table 8: Effect of personalization on sellers

	Historical *product*		
	(1) Revenue	(2) Quantity	(3) Relative rank
Personalized	-251.784***	-2.406***	94.239***
	(34.734)	(0.506)	(26.417)
Intercept	4011.621***	49.388***	1.3e+04***
	(25.001)	(0.364)	(19.014)
Class FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes
Observations	429,417	429,417	429,417

Notes. This table shows the output from the estimation of equation 2. Data is at the consumer-purchased product level. Robust standard errors in parentheses. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Next, we focus on the below-median historical revenue sellers and investigate the importance of personalization for them. Table 9 shows that smaller sellers get 20% higher revenue (+\$79 with the intercept of \$395) in the personalized condition, and they are 15% more likely to be shown on the first 2 pages of the product ranking results.

Table 9: Importance of personalization for below median-revenue sellers

	OLS	*Logit*
	(1) Below median-revenue seller Revenue	(2) First two pages
Personalized	79.201***	0.147***
	(8.232)	(0.011)
Intercept	395.966***	6.221***
	(5.469)	(0.008)
Observations	2,415,416	16,612,314

Notes. This table shows the output from the estimation of equation 1. In Column (1) data is at the seller level, in column (2) the data is at the seller-product level. Robust standard errors in parentheses. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

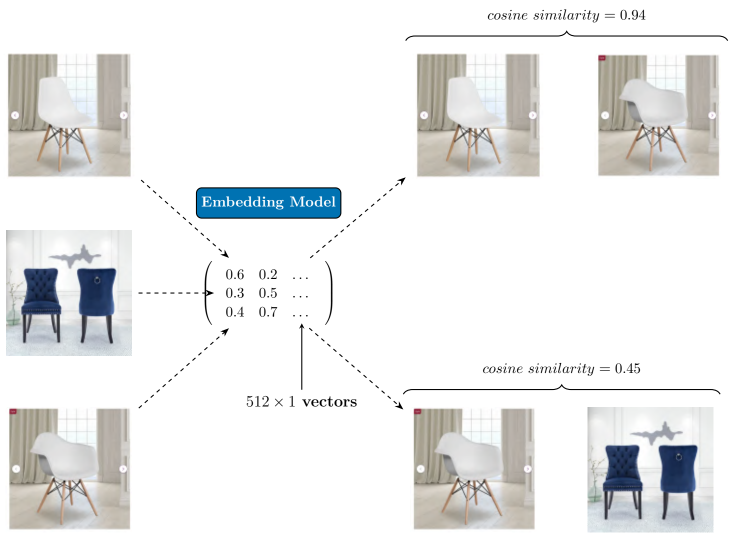
Thus, personalization leads consumers to purchase smaller sellers' products, and the latter get significant part of their revenue from personalized impressions, especially since they are placed more prominently on the website under personalization.

Next, we focused on the impact of personalization based on sellers' experience on the platform. We find that less experienced sellers benefit more from personalization, whereby their revenue increases by 3.2-4.1% in the personalized condition compared to non-personalized (bestseller) condition. Thus, personalization benefits both smaller revenue and less experienced sellers by attracting consumer demand to them. It is important to recognize that in personalized condition the type of seller is not crucial as long as the product is relevant to the consumer, which is why both large and small sellers are treated similarly.

Niche sellers. Traditional view on personalization is based on Chris Andersen's long-tail literature, which suggests that personalization should benefit long-tail products. To empirically test that, we train the image recognition Deep Learning model to, first, efficiently identify niche products versus mass products. Next, we estimate the impact of personalization on product with different nicheness-level.

The embedding model takes as an input more than one million product images and outputs a 512×1 vector representation of an image (embedding), so that similar images are close in this vector space and dissimilar images are farther away. Figure 8 illustrates the idea. A well-trained model should output vectors such that the cosine similarity between similar white chairs is high, e.g., 0.94, and the cosine similarity between dissimilar white and blue chairs is low (e.g., 0.45). Appendix G provides more details on the algorithm training.

FIGURE 8: Illustration of the image embedding process

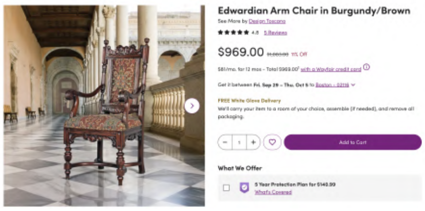


Notes. This figure illustrates how the image embedding algorithm works. Image embedding is a compact representation that captures the essential information about the image. Each image is converted into a numerical vector of 512×1. This procedure allows us to calculate pairwise cosine similarities between vectors.

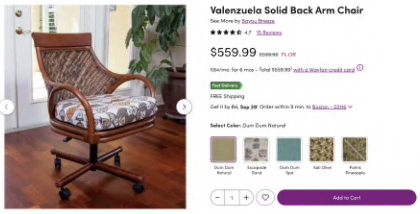
We use the embedding vectors to calculate pairwise similarities between a product and all the remaining products within a category. High cosine similarity corresponds to mass-market products, that are very similar among each other. Low cosine similarity products are the niche products that are dissimilar from the rest. Figure below shows a randomly sampled examples of the results. Panel A shows a product that has mean cosine similarity of 0.07 to the other dining chairs. This means that the chair is a very niche one. Panel B shows a product that has a cosine similarity of 0.2, which means that it is a mid-niche product. Panel C exhibits a mass-market product that has a cosine similarity of 0.8, that is, it is very similar to all the other chairs in the dining char category.

FIGURE 9: Examples of various-niche products

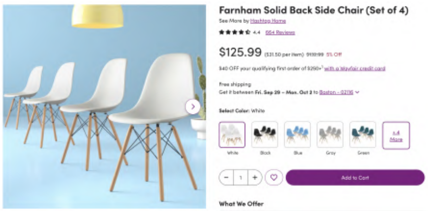
(A) Cosine similarity = 0.07



(B) Cosine similarity = 0.2



(C) Cosine similarity = 0.8



To understand the impact of personalization on niche vs mass-market product, we split all products into five quintiles and constructed a product-year-week panel to understand how a product's revenue changes after being personalized across different quintiles.

$$log(revenue)_{jt} = \alpha + \beta_1 personalized_{jt} + \beta_2 cosine\ quintile_j + \beta_3 personalized_{jt} \times cosine\ quintile_j + \varepsilon_{jt} \tag{4}$$

The results of the regression 4 (Table 10) suggest that contrary to the traditional view, the type of products that benefit the most from being personalized are the mid-niche sellers (quintile 3, see Figure 10).

The rationale is that consumers who prefer very niche (quintile 1) products can find them on their own because they use specific keywords. Meanwhile, mid-niche products are the ones that the platform can help consumers find. Thus, we find that personalization benefits smaller historical revenue sellers and mid-niche product sellers, because they are more likely to gain prominence on the website and are more relevant to the consumers who view them.

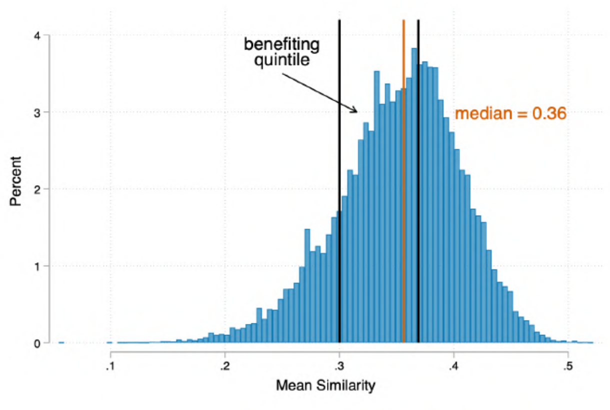
To summarize this section, experimental results suggest that platform, consumers and smaller sellers benefit from personalization. These results serve as reduced-form measures of welfare gains. In the next section, we develop a structural model to evaluate how privacy restrictions affect the benefits of personalization.

Table 10: Heterogeneous effect of personalization on sellers by nicheness

	(1) Log(revenue)
Personalized	0.148***
	(0.007)
Cosine Q2	0.058***
	(0.008)
Cosine Q3	0.062***
	(0.008)
Cosine Q4	0.030***
	(0.008)
Cosine Q5	0.078***
	(0.008)
Personalized× Cosine Q2	0.002
	(0.009)
Personalized× Cosine Q3	0.022**
	(0.009)
Personalized× Cosine Q4	0.014
	(0.009)
Personalized× Cosine Q5	0.008
	(0.009)
Intercept	0.362***
	(0.005)
Observations	1,835,424

Notes. This table shows the output from the estimation of equation 1. In Column (1) data is at the seller-product. Robust standard errors in parentheses. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure 10: Distribution of mean cosine similarities



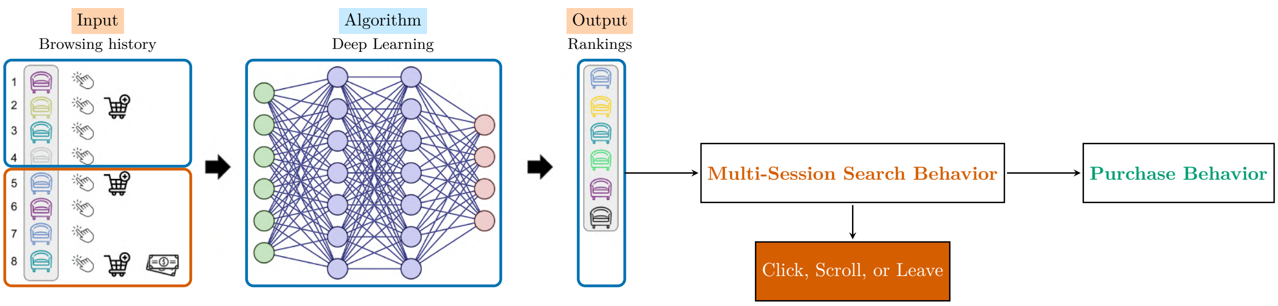
Notes. This figure shows the distribution of mean cosine similarities for each product in dining chairs category. We take fifty thousand products that historically existed in the category and calculate pairwise cosine similarities between them. We then calculate mean similarity by each product and plot the resulting distribution. The black lines show that mid-niche products benefit more from personalization. The result is based on the regression Table 10 .

5. Model

In the previous section, we established that personalization algorithms can benefit consumers, smaller sellers, and the platform. Next, we want to quantify the impact of privacy restrictions on personalization benefits, and understand any heterogeneous effects on different consumers and sellers. To achieve that, first, we construct a counterfactual world in which the platform starts using lower quality (fragmented) data as a result of privacy restrictions. We then re-train Wayfair's personalization algorithm using lower quality data. For example, Chrome 2024 policy will block the ability of the platform to recognize consumers who arrived through display advertising channel. We take Wayfair's data and act as if consumers who arrived from the display advertising channel were not recognized by the platform, which is why the platform could not connect consumer sessions and thought that the consumer arriving from display advertising is a completely new consumer. We input fragmented data to the existing algorithm, which outputs counterfactual recommendations that *would have been* generated had the Chrome restriction been in place. The outline is illustrated in Figure 11.

Next, we need to simulate how consumers will search and purchase on the website under the counterfactual recommendations. We develop a multi-session consumer search model that will allow us to estimate the underlying consumer preferences and search costs. We can then fix these parameters and simulate the changes in consumer choices under the new set of product rankings. Finally, simulations will help us get to our main outcomes of interest: changes in consumer choices, consumer welfare, seller revenue, and platform revenue and profit.

Figure 11: Overview of the counterfactual analysis



One might wonder why we focus on consumer behavior and do not model sellers' and platforms' decisions. We do not have to model platforms' actions because we observe the personalization algorithm they use and re-train their algorithm directly. It is hard for the platforms to change the algorithms in a fast manner, so we operate under the assumption that the platforms' algorithm does not change in the counterfactuals. In the last part of the paper, we relax that assumption.

Next, we do not model sellers' response to the privacy restrictions and to the changes in the personalization algorithms. The reason is that sellers observe aggregate performance of their products on the platform and do not see whether traffic is driven by personalization. Thus, they will not be able to distinguish between the mechanisms that could drive traffic changes. Moreover, they have limited knowledge about the details of the algorithms. Therefore, we only focus on consumer behavior in the model. Next, we show the empirical facts that motivated the model and formalize the model.

5.1. Empirical Facts to motivate the model

This section reports empirical patterns in the data that motivate ensuing model. Consumers' search behavior can be described by a sequence of decisions and we organize the discussion to capture the progression of consumer search and to highlight the decisions that are important for the counterfactual analysis.

Viewing. When consumers reach a product category page, they do not see the entire ranking results. They usually view only top part of the results and have to scroll down to view additional products (Figures 12 and 13). Formally, this means consumers have limited awareness of products and have to incur additional costs to view the remaining products.[31] We observe that there are significant differences in viewing behavior between consumers who see personalized results versus non-personalized results. Table 11 shows that consumers in personalized group viewed four products less within a page and viewed 1.5 pages less than consumers in the non-personalized group. Overall, personalized group consumers viewed 83 products less and purchased products placed higher in the ranking results. This highlights that personalized rankings change the incentives of consumers to scroll and view additional products. When we change the rankings in the counterfactuals we expect that consumers' viewing incentives will change, and it is important to capture that in the model.

Figure 12: Top view of the ranking page

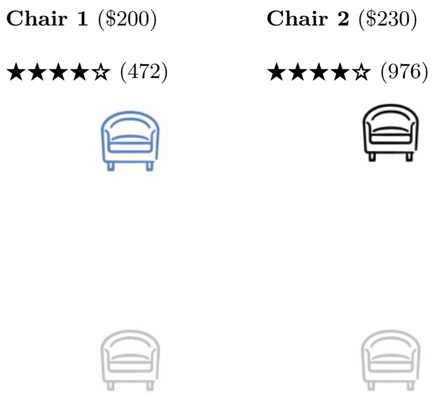
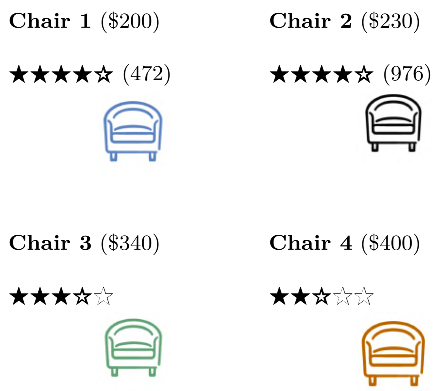


Figure 13: View after a scroll



Learning. We observe that both in the personalized and non-personalized groups consumers exhibit

[31] Traditional search models usually assume full awareness, which means that consumers observe the entire set of products that are available on the website. Two recent papers that restrain from this assumption are Greminger (2022) and Gibbard (2022).

TABLE 11: Viewing patterns

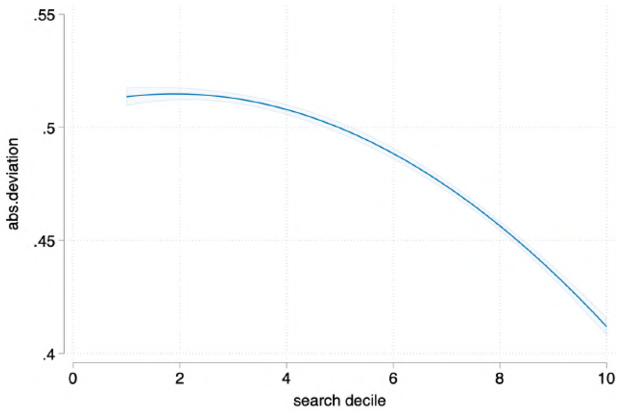
	Search patterns			Purchase
	(1) In-view (within a page)	(2) Search page	(3) In-view (full rank)	(4) Rank (Purchased product)
Personalized	-4.163***	-1.587***	-83.752***	-25.133***
	(0.102)	(0.047)	(2.200)	(4.489)
Intercept	18.679***	2.830***	109.314***	49.232***
	(0.045)	(0.046)	(2.164)	(1.861)
Observations	2,536,098	2,536,098	2,536,098	5,659
Clusters	635,267	635,267	635,267	4,699

Notes. This table reports the output from the estimation of equation 2. Data is at the consumer-session level and tracks the number of products and pages that were in view for a consumer. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

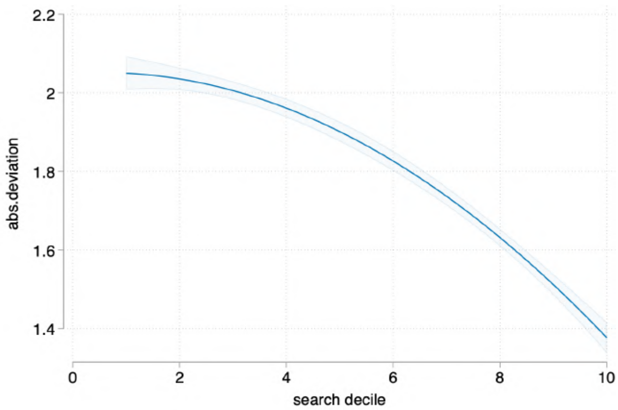
learning behavior similar to the one described in Bronnenberg, Kim and Mela (2016). After clicking on a product and observing its utility, consumers seem to update their beliefs about the remaining products that are similar to the clicked ones. We observe in the data that consumers gradually converge in the attribute space to the product they eventually purchase (Figure 14). We also observe in the data that consumers tend to stop searching for products that are similar to the ones that they previously did not like (Table E19). This phenomenon is called *spatial learning* in the literature (Hodgson and Lewis, 2022). We think that learning is important to account for because if consumers learn fast they can search in a more efficient way even under distorted personalized rankings. If we do not account for learning, we would overestimate the negative effects of privacy restrictions.

FIGURE 14: Convergence patterns during consumer search

(A) log(price)



(B) chair width



Notes. This figure shows that consumers gradually converge towards a chosen product during search. The x-axis shows consumer's search decile (progression), and y-axis shows the absolute deviation of the searched product attribute from the chosen (purchased) product's attribute.

Multi-session search. As was mentioned in Section 3, dining chairs are big-ticket products that require some consideration before consumers purchase. Since rankings change at each re-visit, it is crucial that we account for the multi-session aspect of search.

Recommendation widgets. Recommendation widgets are additional product recommendations that are featured on the product pages. Technically, we could ignore them because they are not personalized and do not use individual-level data. However, consumers extensively click on the products in the widgets. Thus, we need to consider the widgets to get a complete view of consumers' search paths. Moreover, the products in the widgets are similar to the one consumers clicked on, which is why the widgets could be viewed as personalized to consumers' current browsing history. We ran a randomized experiment with the platform where we randomly removed recommendation widgets from the product pages. We use this experiment together with the above described experiment to estimate model parameters.

Refinement actions. Table 7 showed that consumers in the personalized group tend to filter less, which could indicate that their search costs are lower. However, we do not model refinement actions because we do not see any experimental evidence to support that filtering affects click, add-to-cart or purchase behavior. Table D9 shows that consumers who filter are more active on the website, but there are no differences in the outcomes between personalized and non-personalized group consumers who filter. Moreover, we do not see significant differences in the types of filters applied in each group (Figure E7). Thus, we decided not to model refinement actions since they do not seem to change the outcomes under personalized rankings.[32]

Inter-session actions. We decided not to model consumers' inter-session behavior for several reasons. First, we do not have data on how consumers search outside Wayfair.[33] Second, we do not see any differences in the distribution of traffic sources that consumers in the personalized versus non-personalized groups use to return to the platform. We checked both the channel types (e.g., direct traffic, email) and the referral URLs, and did not see any differences (see Figure E5 and Figure E6). Thus, our hypothesis is that personalized rankings affect consumers' behavior within a website visit but do not affect the way they search on other platforms, websites.

Thus, the main components in our model are viewing, clicking, learning, and purchasing patterns of consumers. We take the spatial learning model proposed in Hodgson and Lewis (2022) as a baseline and extend their model by allowing for consumers' limited awareness and multi-session aspect of search, which is crucial for our counterfactual analysis. Next, we formalize the model.

5.2. Set-up

Consumer i arrives to the website at time $t = 1$ and searches for a product on the ranking pages. There are J products on the website. Consumer has limited awareness, which means she observes only part of the products that are featured on the ranking pages. For instance, due to screen size limitations, consumer may view only several products, and she has to scroll down to view additional products. We call the set of products that consumer has viewed at time t her *awareness*

[32] Moreover, it is hard to solve the model with refinement actions. For an example paper, see Chen and Yao (2017).

[33] We attempted to match our data with Comscore, however, due to the incompleteness of Comscore, did not get a good overlap.

set A_t. For each product in her awareness set, consumer observes a vector of product characteristics: $X_j = [\text{price}_j, \text{rating}_j, \#\text{ratings}_j, \text{image}_j],\ j \in A_t$. These characteristics are observable directly on the ranking pages. We represent images in two-dimensional space using the image embeddings trained using Siamese Neural Network (Appendix G) and UMAP.

Consumer's information set at the beginning of search is as follows. She knows (i) her preferences towards observable product characteristics, (ii) her awareness set A_t, (iii) the observable characteristics X_j of viewed products, (iv) the value of the outside option, and (v) she has rational beliefs over the distribution of the observable characteristics of the products outside her awareness set. Given this information set, consumer chooses between leaving, viewing and clicking.

Leaving the website is possible at any point in case consumer has sufficiently low beliefs over the product payoffs. If consumer decides to stay she chooses between viewing and clicking.

Viewing additional products is costly. Consumer incurs a scrolling cost of c_s to expand her awareness set: $A_{t+1} = A_t \cup R_t$, where A_t is the initial awareness set, R_t is the set of products viewed after scrolling, and A_{t+1} is the resulting awareness set.[34]

Clicking on a product reveals additional product characteristics, such as product reviews, but is costly too. Consumer incurs a clicking cost of c_j and can only click on a product she viewed (is aware of). Clicking on product j reveals the true utility of the product. We follow the specification proposed by Hodgson and Lewis (2022) and model the utility of clicking on product j as:

$$u_{ij} = m_i(X_j) + \xi_j + \varepsilon_{ij} \tag{5}$$

where $m_i(X_j) : X_j \to \mathbb{R}$ is the function that maps observable product characteristics X_j to the payoffs, ξ_j is the unobserved product quality common to all consumers and drawn iid from $N(0, \sigma_j^2)$, and ε_{ij} is the idiosyncratic taste shock drawn iid across consumers and products from $N(0, \sigma_\varepsilon^2)$.

There are several important components of this specification that are worth mentioning. First, consumer forms prior beliefs over product payoffs on the ranking pages (before clicking). In particular, given products' observable characteristics, consumer i forms a prior belief over the mean payoff $\mu_i(X_j)$ and the prior uncertainty $\kappa_i(X_j, X_j)$ of product $j \in A_t$. This part is captured by $m_i(X_j)$ in the utility specification 5. In particular, we assume that $m_i(X)$[35] is a function sampled from Gaussian Process with mean $\mu_i(X)$[36] and covariance $\kappa(X, X')$.[37,38] Second, to reveal other a-priori unknown product characteristics, consumer clicks on a product and reveals ξ_j part of the utility. Additionally, there is an idiosyncratic taste shock to the utility ε_{ij}, which is revealed after clicking on the product page. Together all these components constitute the true utility u_{ij}.

[34] After scrolling consumer observes the same set of characteristics for each newly viewed product, i.e. price, etc.

[35] $m_i(X)$ is the vector $J \times 1$, where each element corresponds to each product j, i.e., $m_i(X_j) \in m_i(X)$.

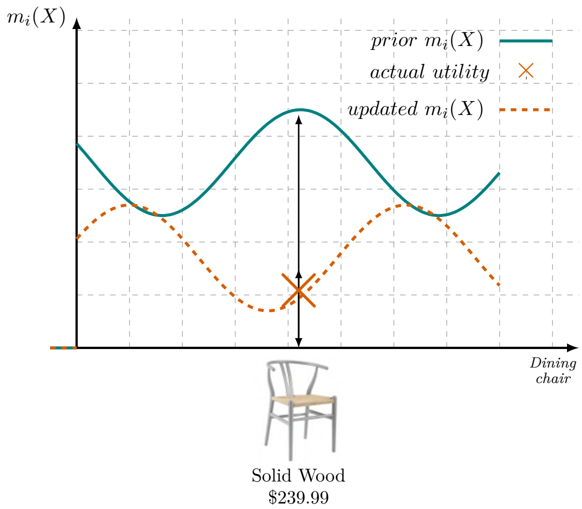
[36] $\mu_i(X)$ is the vector $J \times 1$, where each element corresponds to each product j, i.e., $\mu_i(X_j) \in \mu_i(X)$.

[37] $\kappa(X, X')$ is the $J \times J$ variance-covariance matrix of products' payoffs.

[38] We describe this specification in more detail in the next subsection.

Clicking on a product helps the consumer evaluate the correctness of her prior beliefs. Suppose consumer had high prior belief $\mu_i(X_j)$ regarding product j, but the actual utility of the product appeared to be low. This is illustrated in Figure 15. The teal line is the payoff function $m_i(X)$ sampled from Gaussian Process with consumer's prior mean and variance-covariance beliefs. Suppose consumer thought that the payoff of the solid wood chair is high, so she clicked on it. The true utility u_{ij} (red cross) turned out to be low. For instance, reading reviews revealed that the product is not good. As a result, consumer updates her posterior beliefs given the new information.

FIGURE 15: Illustration of consumer learning process



Thus, clicking (i) reveals additional product characteristics, (ii) reveals product utility, and (iii) allows consumer to update her beliefs. Moreover, on the product pages, consumer views additional product recommendations that are similar to the clicked product. This automatically expands consumers' awareness set at *no cost*. Thus, after clicking product j, consumer's awareness set becomes $A_{t+1} = A_t \cup R_{jt}$ where R_{jt} is the set of products that are recommended on product j at time t.[39],[40]

After clicking a product, consumer can either purchase the clicked product, go back to the ranking page and click on one of the products she is aware of, view additional products, or leave the website. Next, we describe how consumer decides which action to take.

5.3. Consumer's decision problem

Technically, one would have to write and solve a full dynamic Bellman equation to get consumer's optimal search path. However, with thousands of consumers and thousands of products it is not

[39] For simplicity, we assume that while consumers have rational expectations over the remainder of the ranking pages, consumers do not form beliefs over the recommendation widget. The rationale is that the platform shows most similar products which means the index of the products in the recommendation widget coincide with the anchor product in expectation.

[40] Consumer observes the same set of characteristics, i.e. price, ratings, numbers of ratings and images for the products on the recommendation widgets.

feasible to do backward induction. Instead of solving the problem by backward induction, we use a heuristic near-optimal approximation to the solution given the descriptive evidence in the data. To decide whether and what to search, we assume that consumer follows an index strategy. When consumer lands on a ranking page, we assume that she constructs a utility index for the products in her awareness set A_{it}

$$z_{ijt} = \mu_{ijt} + \eta\kappa_{ijt} - c_{ijt},\ j \in A_{it} \tag{6}$$

where z_{ijt} is the utility index of consumer i for product j at time t, $\mu_{ijt} = \mu_{it}(X_j)$ is consumer i's prior mean payoff of product j at time t, $\kappa_{ijt} = \kappa_{it}(X_j, X_j)$ is the prior uncertainty about the j's payoff, c_{ijt} is the cost of clicking product j at time t.[41] If consumer decides to click on a product in her awareness set, she clicks on a product with highest index: $j^* = \arg\max_{j \in A_{it}} z_{ijt}$.

This index policy is similar to the Upper-Confidence Bound algorithm widely used in multi-armed bandit literature. We prove the near-optimality of the algorithm in our setting in Appendix B. The index policy captures that consumer has higher index for products with higher prior mean payoff (μ), but may also explore products with higher uncertainty.[42] This search behavior aligns with the standard exploration-exploitation tradeoff at the core of bandit literature.

Note that consumer may choose to view additional products before clicking. Consumer does not observe characteristics of the products outside her awareness set: J/A_t. However, Assumption 1 states that consumer has rational expectations over products outside her awareness set and she knows the correct distribution from which the observable characteristics are sampled.

Assumption 1 (Rational expectations)**.** *Consumers do not know the full set of products available on the platform, i.e., they have limited awareness. However, we assume that they know the correct distribution of all the products that are not in their awareness set.*

Therefore, consumer can construct an *expected* utility index summarizing the expected maximum utility that she believes she can find outside her awareness set. Formally, consumer constructs:

$$E[max(z_{ijt}) - c_s(r_{it})],\ j \in J/A_{it} \tag{7}$$

where expectation is taken with respect to the correct distribution of the observable characteristics of the products outside the awareness set. Two features are worth mentioning. First, consumer has beliefs over the maximum utility index she gets from clicking one of the products outside her awareness set. The net utility is calculated in the same way as in Equation 6 and already includes the clicking cost. Second, consumer incurs additional scrolling cost of $c_s(r_{it})$ to view products. Similar to Greminger (2022), we model the scrolling cost as a function of the position in the rankings reached so

[41] We abbreviated the $\mu_{ijt},\ \kappa_{it}(X_j, X_j)$ for simplicity.

[42] It is an empirical question whether consumer likes exploring uncertain products ($\eta > 0$) or dislikes it ($\eta < 0$) or is completely indifferent ($\eta = 0$).

far, r_{it}. This means that scrolling costs are allowed to change depending on the number of products consumer viewed so far.

Consumer decides to view the products instead of clicking if the expected maximum utility index from viewing products outside awareness set is higher than the maximum utility index in the awareness set

$$E[\max_{j \in J/A_{it}}(z_{ijt}) - c_s(r_{it})] > \max_{j \in A_{it}} z_{ijt} \tag{8}$$

Recall that the full discrete choice problem that the consumer solves is choosing between (i) leaving, (ii) purchasing clicked item, (iii) viewing and (iv) clicking. Formally, at time t, consumer chooses:

$$max\{ \underbrace{u_0 = 0}_{\substack{\text{outside option} \\ \downarrow \\ \textbf{Leave}}}, \underbrace{\hat{u}_i}_{\substack{\text{best utility} \\ \text{observed so far} \\ \downarrow \\ \textbf{Purchase}}}, \underbrace{max_{j \in A_{it}} z_{ijt}}_{\substack{\text{choose highest} \\ \text{index product} \\ \downarrow \\ \textbf{Click}}}, \underbrace{E[\max_{j \in J/A_t} z_{ijt} - c_s(r_{it})]}_{\substack{\text{view more} \\ \text{products} \\ \downarrow \\ \textbf{View}}}\} \tag{9}$$

5.4. Belief updating

At each point in time, consumer keeps track of the following state variables:[43] (i) current mean payoffs $\mu_t(X)$, (ii) current covariance matrix $\kappa_t(X, X')$, (iii) best product observed so far $\hat{\jmath}$, (iv) utility of the best product observed so far $\hat{u}$, (v) awareness set at the beginning of time t, A_t, and (vi) the set of products that she hasn't viewed yet J/A_t, (vii) the set of products she hasn't clicked on yet. We explain the transition of the mean payoffs and the covariance function, which are the moments of the Gaussian Process function $m(X)$ from which the product payoff function is drawn.

After each click consumer observes the utility of product j, u_j, and updates her beliefs about all the remaining $J-1$ products. Consistent with the Gaussian process specification, posterior mean utilities on the remaining $J-1$ products are updated as follows:

$$\underbrace{\mu'(X_{-j|j})}_{\textit{posterior means}} = \underbrace{\mu(X_{-j})}_{\textit{prior means}} + \underbrace{\frac{\kappa(X_{-j}, X_j)}{\kappa(X_j, X_j) + \sigma_\xi^2 + \sigma_\varepsilon^2}}_{\textit{weights}} \underbrace{(u_j - \mu(X_j))}_{\substack{\textit{deviation of observed} \\ \textit{utility from the prior}}} \tag{10}$$

where $\mu'(X_{-j|j})$ is the $(J-1) \times 1$ vector of posterior means on the yet unclicked $J-1$ products; $\mu(X_{-j})$ is the $(J-1) \times 1$ vector of prior means on these products; $\kappa(X_{-j}, X_j)$ is the covariance between the payoffs of products $-j$ and j; $\kappa(X_j, X_j)$ is the variance of the payoff of product j; σ_ξ^2 and σ_ε^2 are the uncertainties in the distribution of ξ_j and ε_j, respectively. Intuitively, the posterior mean payoff of a product -j is its prior mean payoff plus a weighted deviation of the actual observed utility from the prior mean of j, $(u_j - \mu(X_j))$. Weights are directly proportional to prior covariance between j and -j. If $\kappa(X_{-j}, X_j) = 0$, i.e. the clicked product j and some other product $-j$ are

[43] We drop the i subscripts for convenience

unrelated, then the posterior on -j is not updated at all. If $\kappa(X_{-j}, X_j)$ is high then the posteriors will be updated more for -j. In the example on Figure 15, one could argue that receiving low utility on wooden chair may downgrade consumers' beliefs about other wooden chairs but does not change consumers' beliefs about leather chairs.

Posterior covariances are updated as follows:

$$\underbrace{\kappa'_{-j|j}}_{\substack{\textit{posterior} \\ \textit{covariance matrix}}} = \underbrace{\kappa_{-j|j}}_{\substack{\textit{prior} \\ \textit{covariance matrix}}} - \underbrace{\kappa_{-j,j} \overbrace{(\kappa_{j,j} + \sigma_\xi^2 + \sigma_\varepsilon^2)}^{\substack{\textit{total uncertainty} \\ \textit{about product } j \textit{ payoff}}}{}^{-1} \kappa_{-j,j}^T}_{\textit{reduction in uncertainty}} \tag{11}$$

where $\kappa'_{-j|j}$ is the $(J-1) \times (J-1)$ posterior covariance matrix, $\kappa_{-j|j}$ is the $(J-1) \times (J-1)$ prior covariance matrix. The intuition behind Equation 11 is that the posterior uncertainty about product relationships decreases by the term $\kappa_{-j,j}(\kappa_{j,j} + \sigma_\xi^2 + \sigma_\varepsilon^2)^{-1}\kappa_{-j,j}^T$, which is positive and increasing in the prior covariance between products j and -j. If products are not related at all, then there is no decrease in the uncertainty of product -j's payoff. If products have high prior covariance, the uncertainty is going to decrease because consumer revealed the true utility of a similar product j.

We want to emphasize the role of these updating rules in consumer's decision problem described in the previous subsection. At each time t, consumer uses the *current* mean payoff and covariance beliefs when constructing the utility index:

$$z_{ijt} = \mu_{ijt} + \eta\kappa_{ijt} - c_{ijt},\ j \in A_{it} \tag{12}$$

This utility index is used both for in the clicking and the viewing decisions. After every click, consumer updates her beliefs and her utility indices are updated accordingly.

Assumptions. In the model, we make the following main assumptions. First, we assume that it is costless for the consumer to navigate back to the ranking page from the product page. Second, we assume that consumers have perfect recall within and across sessions. Thus, consumers remember and keep track of all the products they have viewed previously. Third, we assume that consumer do not forget any information they obtained across sessions:[44] this allows us to model multi-session search via propagating posteriors from the previous session as priors to the next session.

5.5. Model parametrization

In this part, we explain how we estimate the model using detailed clickstream and pixel-level data. Recall that the utility of consumer i from purchasing product j is given by

[44] A more involved model could use power prior that allows for a forgetting factor. See Ibrahim, Chen, Gwon and Chen (2015). Alternatively, one could incorporate the forgetting specification from Mehta, Rajiv and Srinivasan (2004). We do not do this for computational reasons.

$$u_{ij} = m_i(X_j) + \xi_j + \varepsilon_{ij} \tag{13}$$

where $m_i(X_j)$ is drawn from a Gaussian Process with prior mean payoff $\mu_i(X)$ and $\kappa_i(X, X')$, $\xi_j \sim N(0, \sigma_\xi^2)$ and $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$. We further parametrize the prior mean and covariance functions as follows. Prior mean function over product payoffs is a linear function of observable product characteristics:

$$\mu(X) = \alpha + X\beta_i \tag{14}$$

where X is a vector of observable product characteristics, β_i are consumer preferences over observable characteristics. Similar to Berry, Levinsohn and Pakes (1995), we allow for consumer heterogeneity through random coefficients such as $\beta_i \sim N(\beta, \Omega)$: β is the mean preferences and Ω is the variance matrix.

To get positive-definite variance-covariance matrix, one of the common specifications in Gaussian process is to parametrize the prior κ matrix as a squared exponential kernel:

$$\kappa(X_j, X_k) = exp\left(- \sum_a \frac{(X_{ja} - X_{ka})^2}{\rho_a} \right) \tag{15}$$

where $X_{ja} - X_{ka}$ indicates the difference between the value of attribute $a = \{price, rating, \#ratings, image\}$ for products j and k, ρ_a is the learning parameter along the dimension of attribute a. This specification implies that the covariance between products characterized by the observable vectors X_j and X_k depends on the sum of the distance between these vectors along each attribute scaled by the learning rate ρ. For example, if all $\rho_a = 0$ then the covariance between products j and k is zero, i.e., they are unrelated. However, if $\rho_a \neq 0$, and the distance between attributes $X_{ja} - X_{ka}$ is low then the product payoffs are highly related.

To simplify the estimation procedure, we assume that conditional on viewing the product, clicking cost is constant c_0 and there is a logit error term ψ_{ijt}, which accommodates any potential idiosyncracies in clicking costs across consumers, products and time.

$$c_{ijt} = c_0 + \underbrace{\psi_{ijt}}_{Type1EV} \tag{16}$$

If consumer decides to view additional products, the scrolling cost at time t is specified as:

$$c_s(r_t) = c_s \cdot log(r_t) \tag{17}$$

where c_s is the constant part, and r_t is the product rank reached so far at time t.

5.6. Estimation

Parameters to be estimated. The main parameters to be estimated are preference parameters in the prior mean function, α, β and Ω; learning rates ρ_a; baseline clicking cost c_0 and scrolling cost c_s; the exploration parameter η in the utility index function; product fixed effects ξ_j; and the variances σ_ξ^2, σ_ε^2.

In the clickstream data, for each consumer at each point in time we observe the ordered list of products, i.e., product rankings that are served to the consumer. We also observe the vector X_j for each product. Pixel-level data tells us which products were viewed by a consumer at each point in time. Thus, for each consumer at each point in time we observe (i) the awareness set at time t, A_t, (ii) products that she has not viewed yet J/A_t, (iii) products that she has not clicked yet.

Given initial parameter values, for consumer i we can draw $\beta_i \sim N(\beta, \Omega)$ and calculate prior mean payoffs and variance-covariance functions as in Equations 14 and 15. Next, given clicking and scrolling cost consumer constructs utility indices in Equations 6 and 7. In the data, we observe whether consumer decides to click or to view the product. Upon clicking, consumer reveals product's utility. We draw the payoff function $m_i(X) \sim GP(\mu_i(X), \kappa(X, X'))$. Given the initial $\sigma_\xi, \sigma_\varepsilon$, we can construct the utility as in Equation 13. After each click, we update consumer's posterior beliefs according to Equations 10 and 11. Next, we observe whether consumer decided to click further, view additional products, leave, or purchase the clicked product.

To estimate the model, we construct the likelihood function of the observed search paths and purchased options. Given the assumption of logit error terms on the clicking costs (Equation 16), the probability that the consumer chooses to search product j conditional on being in state $\mathcal{S}$ is

$$P(j_{it}|S) = \frac{exp(E[max(\hat{u}, u_j)|\mathcal{S}] - c_0)}{exp(\hat{u}) + \sum_{l \in J} exp(E[max(\hat{u}, u_l)|\mathcal{S}] - c_0)} \tag{18}$$

where $\hat{u}$ is the best utility searched so far. This structure nicely follows because we assumed logit cost error terms. Had we observed the entire state space for each consumer including the drawn utility payoffs, writing the likelihood function would be straightforward: the likelihood of consumer i searching for T_i periods would be:

$$L_i(\{j_{it}\}_{t=0}^{T_i} | \{\mathcal{S}\}_{t=0}^{T_i}, \theta) = \prod_{t=0}^{T_i} P_i(j_{it}|\mathcal{S}_{T_i}) \tag{19}$$

However, since we do not observe the drawn utilities, we have to integrate them out. Recall that due to the Gaussian structure, the distribution of the drawn utilities is $G(u_i) = N(\alpha + X_j\beta_i + \xi_j, \Sigma_i)$, where diagonal elements of Σ_i are $\kappa(X_j, X_j) + \sigma_\varepsilon^2 + \sigma_\xi^2$ and off-diagonal elements are $\kappa(X_j, X_{j'})$. Given that in addition to the utility draws we have to integrate out the random coefficients $(F(\beta_i))$, the likelihood function is as follows:

$$L_i(\{j_{it}\}_{t=0}^{T_i}, \hat{j}_i|\theta) = \int\int L_i(\{j_{it}\}_{t=0}^{T_i}, \hat{j}_i|\{\mathcal{S}_t\}_{t=0}^{T_i}, \theta)dG(u_i)dF(\beta_i) \tag{20}$$

During the estimation, we maximize the products of the individual likelihood functions in Equation 20 across all consumers. Due to the size of the data and potential incidental parameters problem when estimating large number of product fixed effects, we use Batch Stochastic Gradient Descent algorithm (Keskar et al., 2016) that samples consumers in batches of 500-1500 and minimizes the negative likelihood function.

5.7. Identification

Prior mean and variance parameters. The probability that each product is searched first identifies the prior mean parameters, β and α, and the total variance of prior beliefs. To explain identification of the variance of random coefficients, we use standard argument for the discrete choice model identification (Keane, 1997). If we observe more variation in the attributes of the searched products across individuals than within individual search paths, this would indicate higher heterogeneity in random coefficients β_i.

Price parameter. Estimating price parameter would be prone to endogeneity because more popular products could be priced higher. To address this concern, we use a period of time when the platform ran price experiments in the category. Price experiments randomly varied prices of products from -12% to +12% as shown in Figure 16. Only a subset of products were part of the experiment. However, we still use the existing experimental variation to address the price endogeneity.[45]

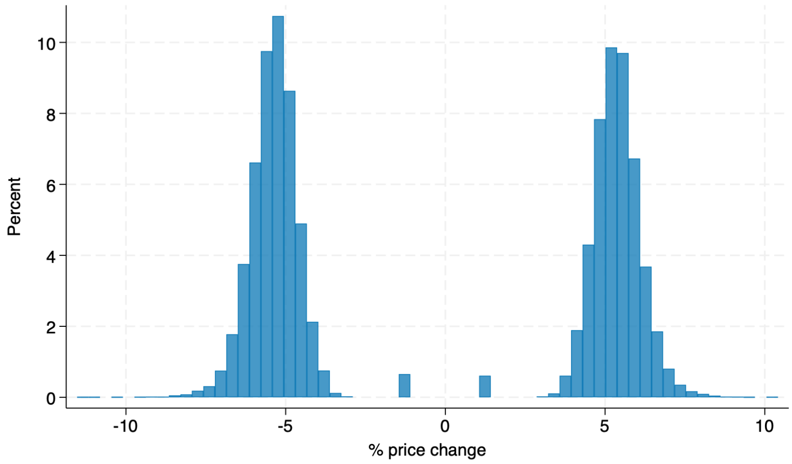
Clicking and scrolling costs. Experimental variation in the rankings and pixel-level data allows us to identify the clicking and scrolling cost. Baseline scrolling can be identified from the number of scrolls that consumer makes in the data. However, the clicking cost can no longer be identified simply from the number of searches that consumer makes. The reason is that consumer has to view the product to click on it. Therefore, the experimental variation in the product rankings allows us to identify the clicking cost.

Product fixed effects. The probability that product j is purchased, conditional on being clicked identifies product fixed effects, ξ_j. If a product is rarely purchased compared to the others with similar observable attributes, then it must be that $\xi_j < 0$.

Learning parameters (ρ_a) are identified from the observational data. Suppose product j has negative fixed effect, $\xi_j < 0$. Given the covariance matrix, we know which products are most similar to product j in the observable characteristic space. Probability of clicking on a product that is

[45] Alternatively, we could estimate the average price elasticities directly from the experiment and create a moment condition so that the implied price elasticity from the model would be arbitrarily close to the experimental price elasticity. This could be accommodated at the additional computational cost.

FIGURE 16: Price experiments



Notes. This figure shows the distribution of % price changes during the price experiments. The distribution is bimodal because experiment involved both the random increase and the decrease of prices.

similar to j should be lower in case consumer exhibits learning, i.e., $\rho_a > 0$. Similarly, the probability of clicking on a product that is similar to j with $\xi_j > 0$ should be higher under learning.

Exploration parameter is also identified both from the observational and experimental data. Suppose product j has negative fixed effect, $\xi_j < 0$. If there is a product k that is very similar to product j, then under learning framework consumer has to have lower probability of searching product k, as explained above. However, consumer could also stay in that region especially if product payoff uncertainty $\kappa_{kk'}$ is high. Therefore, while jumps in attribute space identify the learning parameter, the reluctance to jump when sampling a product with a negative fixed effect identifies the exploration rate parameter.

5.8. Estimation Results

We estimate the model on two samples as required by the counterfactuals. Recall that in the counterfactuals, we change the product rankings by mimicking the privacy restrictions of interest. First counterfactual deletes first-party data, i.e., that of consumers who arrive directly to the website. Second counterfactual deletes third-party data belonging to consumers who arrive from advertising channels. Thus, there are two samples that we estimate the model on: (i) consumers who arrive directly to the website and were first-party cookie-recognized, and (ii) consumers who arrive from advertising channels. Table E14 and Table E15 in the Appendix show the results of the t-test confirming that the search and purchase behavior of these two samples of consumers are different. Thus, for the validity of the counterfactuals we estimate the model twice.

Table 12 shows the estimation results for the cookie-recognized consumers, and Table 13 shows the estimation results for the advertising-based consumers. Qualitatively the results are similar

TABLE 12: Estimation results for cookie-recognized consumers

	Estimates	
	β	*st.err.*
Price ($)	-0.814	(0.003)
Rating	0.682	(0.002)
# Ratings	1.839	(0.013)
Image (x)	-0.268	(0.009)
Image (y)	0.899	(0.004)
Scrolling cost ($)	0.113	
Clicking cost ($)	0.200	
ρ_{price}	1.826	(0.004)
ρ_{rating}	0.023	(0.301)
$\rho_{\#ratings}$	1.405	(0.019)
$\rho_{image(x)}$	0.871	(0.008)
$\rho_{image(y)}$	1.290	(0.004)
Log-likelihood	4,934	
# Consumers	9,500	

across both samples: consumers dislike high prices and like products with higher ratings and higher number of ratings. When it comes to images, higher x and y correspond to modern chairs as is illustrated in Figure G26 . Therefore, the model predicts that consumers in both samples prefer chairs more similar to modern or traditional styles. Scrolling cost is almost twice smaller than the clicking cost in dollar terms, which is intuitive. Consumers seem to be learning along all dimensions except product ratings (all other ρ's are positive and significant).

There are several differences between two samples worth mentioning. First, consumers who arrive from advertising are less price sensitive ($\beta_{price} = -0.430$ versus -0.814 among cookie-recognized consumers). Second, consumers arriving from advertising have lower search costs (both clicking and scrolling). These patterns are consistent with the Table E15 provided in the Appendix, where consumers who arrive from advertising search more and purchase more expensive products.

TABLE 13: Estimation results for advertising-based consumers

	Estimates	
	β	*st.err.*
Price ($)	-0.430	(0.004)
Rating	0.817	(0.000)
# Ratings	2.005	(0.020)
Image (x)	-0.109	(0.001)
Image (y)	1.720	(0.029)
Scrolling cost ($)	0.051	
Clicking cost ($)	0.095	
ρ_{price}	1.901	(0.001)
ρ_{rating}	0.050	(0.602)
$\rho_{\#ratings}$	1.783	(0.004)
$\rho_{image(x)}$	0.180	(0.002)
$\rho_{image(y)}$	1.302	(0.003)
Log-likelihood	15,209	
# Consumers	8,000	

Thus, overall, the estimates make sense. To validate the model, we checked how well the model fits the data moments. First, we test the model fit using the observational data moments, namely, the awareness set size of the consumers (Figure J30). We also test the model fit by comparing predicted and data patterns during the Chrome event that occurred in 2020. During this event consumers' search costs increased substantially and we confirm that the model can predict the data patterns well (Figure J31).

6. Counterfactuals

This section shows the results of the counterfactuals we ran to evaluate the impact of privacy policies on the personalization outcomes. First, we describe the changes that occur in each counterfactual. Next, we explain the simulation procedure and present the results.

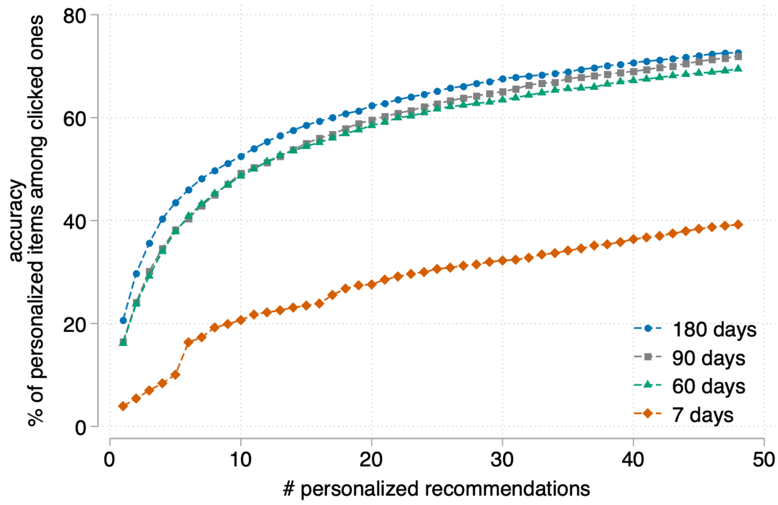
I. First-party data restrictions. Recall that one of the big changes was Safari regulation that automatically resets cookies after seven days of consumer inactivity. Thus, if consumer arrives to the website more than seven days later than the initial session, the platform will not recognize her because first-party cookies were reset. To mimic Safari's 7 day cookie reset policy, first, we re-trained personalization algorithm where the input data consisted only of the most recent 7-day searches for each consumer. That is, if consumer's inter-session arrival time was more than 7 days, we split the data and created a new customer identifier. We then input the fragmented data into the model and re-trained it. For completeness, we compare the 7-day model with the re-trained models where we kept 60, 90 and 180 days of data for each consumer. Figure 17 shows the accuracy of the resulting personalization algorithm. The x-axis shows the number of personalized recommendations shown to the consumer. The y-axis shows the predictive accuracy of the model, where accuracy is defined as the percent of personalized items among the items that consumer eventually clicked on. Note that we use the offline evaluation approach standard in Computer Science to plot these graphs: we fix the set of items consumer clicked on and evaluate whether the newly trained model would have recommended those items. Two points are worth mentioning. First, 7-day model exhibits significantly worse performance than all the other models. Second, there seems to be diminishing returns from data because 60, 90, and 180-days models perform similarly when they are allowed to show more recommendations.

In the first counterfactual, we change the personalized rankings using 7-day model assuming same fraction of Safari users in the counterfactuals as in the real data.[46,47] We fix the estimated set of parameters for cookie-recognized consumers (Table 12) and simulate how consumers respond to the personalized rankings generated using 7-day model.

[46] In the data, the share of Safari users is approximately 40%.

[47] It might be useful for the reader to think of this as a linear regression $y = \alpha + \beta X + \varepsilon$, where 7-day model is characterized by the parameters α, β. We input consumers' browsing histories as X's and the model gives the predicted ordered list of rankings.

Figure 17: Training accuracy by training data size



Notes: This figure shows the algorithm accuracy when we use 180, 90, 60, and 7 days of data. The x-axis is the number of personalized recommendations shown. They y-axis shows the algorithm accuracy defined as the % of personalized items among the clicked ones.

II. Chrome 2024 restriction on third-party cookies. In the second counterfactual, we concentrate on the effect of third-party cookie restrictions on recognition when consumers arrive from the display advertising channel. We mimic the Chrome restrictions by de-recognizing consumers who arrive from display advertising. Recall that it is 26% of the traffic in our sample (Figure 2).

Note that if consumers are completely not recognized then they would see non-personalized recommendations and the comparison of personalized versus non-personalized rankings is already captured by the experiment. Instead, we simulate two more interesting situations. First, we re-train the personalization algorithm using the fragmented data where we de-recognize consumers who arrive from Chrome display advertising channel. The algorithm re-training generates counterfactual rankings. Next, in the simulations, we assume that consumers arrive from display advertising channel and, therefore, are anonymized under Chrome restrictions. Their first session is non-personalized, and in subsequent sessions they get personalized recommendations using the re-trained algorithm.

Second approach is as follows. Note that Chrome plans to offer an alternative solution where platform may not be able to track consumers using third-party cookies, but will be able to get access to their aggregate interests. In addition to the simulations explained above, we also simulate a situation where Chrome does not show the platform who the consumer is but may share aggregated data based on consumers' browsing history. For instance, Chrome may indicate that consumer is interested in modern chairs or glam style chairs. The benefit of the aggregated information is that platform can keep personalizing, and, from the regulatory perspective, the platform does not know the price point consumer is interested in.[48]

[48] However, in Section 4, we showed that the platform benefits in long-term from showing better consumer-product

To simulate the aggregate information, we use a simple heuristic rule that clusters consumers into the product styles they are interested in. The major styles of dining chairs are: modern & contemporary, traditional, scandinavian, posh & luxe, industrial, french country, farmhouse, and coastal. Using a consumer's browsing history, we calculate the number of clicked chair styles among all viewed products, and classify the consumer into the cluster (style) that got the highest share of clicks.[49] We then personalize recommendations by showing a mix of bestseller and smaller seller's products from the chair style consumer is most interested in. We then simulate how consumers will respond to this type of recommendations.

Both approaches lead to qualitatively similar, but quantitatively different results, so in the subsequent discussion we show the results for the first approach, and we include the results from the second approach in the Appendix K.

III. Probabilistic Recognition Algorithm. Third counterfactual evaluates an algorithm proposed in Korganbekova and Zuber (2023) that aims at helping platforms adapt to privacy restrictions. The idea behind the algorithm is as follows. We use a device's behavioral data (e.g., clickstream data and purchase behavior) and IP address information to predict the association between the device and the existing customer identifier. We use XGBoost algorithm to classify consumers.

Counterfactual simulations. The simulation procedure for all three counterfactuals is similar. We fix the estimated set of parameters. The first and third counterfactuals use the estimates for the cookie-recognized consumers (Table 12) and the second counterfactual (advertising) uses the estimates from Table 13. Each consumer sees non-personalized bestseller recommendations during the first session. They decide between clicking, viewing, leaving, and purchasing. Suppose consumer searched for some products and left. We assume that consumer's probability of re-visiting the website, i.e., multi-session search, is equal to the fraction of clicked products among viewed ones. This is motivated by the empirical patterns in the data and serves as a proxy for consumers' interest level. One could argue that the opposite could be true: consumers who clicked a lot and left are less likely to return because they made up their minds. To account for that we add noise by allowing no re-visit with probability $\varepsilon \in U(0, 1)$. When consumer re-visits the website, we have their browsing history and can generate rankings according to the counterfactuals. Table 14 summarizes all three counterfactuals.

In *the first counterfactual*, we generate the product rankings using 7-day personalized model assuming the same fraction of Safari users as in real data, and simulate how consumers respond to them. In *the second counterfactual*, we use re-trained algorithm where the data is fragmented as a result of Chrome restrictions. In *the third counterfactual*, we proceed as follows. We take

matches, instead of pushing consumers towards higher margin items.

[49]Note that accounting for the share of clicks among viewed products is a more accurate measure than to simply calculate the number of clicks. We could also give more weight to more recent clicks but wanted to keep the counterfactual simple.

TABLE 14: Summary of counterfactuals

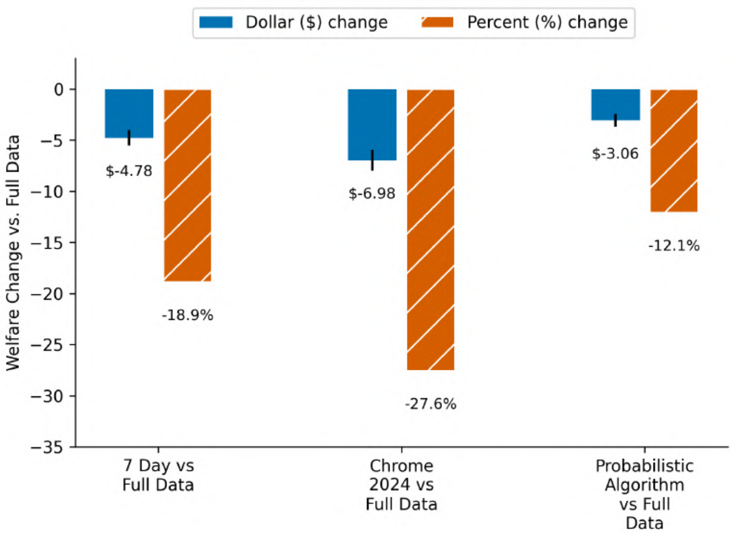
Counterfactual 1. Safari blocks first-party cookies	
Estimates used: Table 12	
1.1	Use 7-day personalized ranking algorithm.
Counterfactual 2. Chrome 2024 blocks third-party cookies	
Estimates used: Table 13	
2.1	Fragment the data by de-recognizing consumers who arrived from display advertising.
2.2	Show personalized rankings using re-trained model.
Counterfactual 3. Probabilistic Identity Recognition	
Estimates used: Table 12	
3.1	Identify cookie-recognized consumers who searched for multiple sessions.
3.2	Apply probabilistic recognition algorithm to associate devices with the consumer identifiers.

cookie-recognized consumers who browsed using multiple devices. We take their first sessions until the device change as given. As they re-visit the website using a new device, we run a separate probabilistic identity recognition algorithm to predict the association between the new device and the unique consumer identifier. We know the ground truth association between the devices and the consumers but conceal it to evaluate the algorithm. The algorithm produces a probability distribution indicating the probability that a device is associated with a consumer identifier. We take the consumer identifier with the highest predicted association and show the personalized rankings based on the associated consumer identifiers' browsing history. Note that in this counterfactual we use business-as-usual personalization algorithm. Next, we evaluate the consumers' actions given the new set of rankings.

The main outcomes of interest are (i) consumer welfare, (ii) consumers' search and purchase outcomes, (iii) seller revenue, and (iv) platform's revenue and profit. To have a common comparison benchmark, we compare consumer welfare to the welfare gains from the personalized rankings. We ran simulations with 10,000 consumers, where for each preference parameter and Gaussian Process draw, we generate and show personalized rankings. Consumer's welfare from a ranking is defined as the utility obtained from the item purchased under that ranking, net of total clicking and scrolling costs incurred during search. We average consumer welfare across multiple rankings and then calculate the difference between the personalized rankings and non-personalized rankings welfare. We find that the welfare from personalized rankings amounts to $25.3 per purchase. See Table K23 in the Appendix for full results.

Figure 18 shows the dollar and percent changes in consumer welfare as a result of privacy restrictions. The comparison benchmark is the personalized rankings. We find that consumers lose $4.78 after first-party cookie blocking, $6.98 after third-party cookie restrictions, and $3.06 if we use probabilistic recognition versus full data personalization. Thus, in percentage terms up to 28% of welfare gains from personalization are lost as a result of privacy restrictions. It is worth mentioning that (i) consumers are still better off than in the non-personalized condition despite the decrease in welfare, and (ii) our probabilistic algorithm can mitigate up to 56% of welfare losses.

FIGURE 18: Counterfactual results: consumer welfare

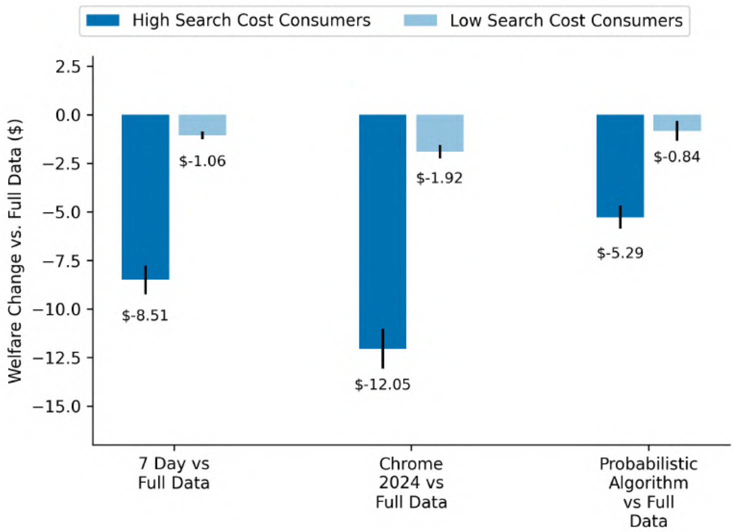


Notes: This figure shows the results of the counterfactual simulations on consumer welfare by the three counterfactuals ran. Chrome counterfactual is using approach 1 and approach 2 results can be found in the Appendix.

We breakdown the changes in welfare by the losses from the match value (utility), clicking costs, scrolling costs, and decreases in purchase probability for all three scenarios in Table K23 . To illustrate, the welfare losses after Chrome restrictions are driven by the decrease in purchase probability (-24.65%), significant decrease in the match value conditional on purchase (-54.11%), and increase in the scrolling costs (+32.53%). Partially, the losses are offset by the fact that consumers click less and, therefore, there are savings in the clicking costs (+11.28%). Given that the biggest driver of the welfare losses is the decrease in match value, this implies that consumers find it hard to find and, subsequently, purchase items that are more relevant towards their particular taste. Having said that, we also may be underestimating the welfare losses because we do not account for potential hassle costs in case consumers have to return the product post-purchase.

To investigate the heterogeneity across consumers, we re-ran simulations focusing on different types of consumers. Namely, we breakdown the losses in welfare depending on the level of consumers' search costs. We re-simulate the search and purchase behavior of the consumers with below and above median search costs. Figure 19 shows that consumers with high search costs will be most hurt by the 7-day Safari policy, losing \$8.51 compared to the \$1.06 lost by consumers with lower search costs. Similarly, above median search cost consumers will incur significantly higher welfare losses compared to those of consumers with lower search costs, i.e., \$12.05 versus \$1.92. However, the proposed ranking algorithm can significantly lower the welfare losses both for high search cost and low search cost consumers.

FIGURE 19: Counterfactual results: consumer welfare



Notes: This figure shows the results of the counterfactual simulations on consumer welfare by the three counterfactuals ran. Chrome counterfactual is using approach 1 and approach 2 results can be found in the Appendix.

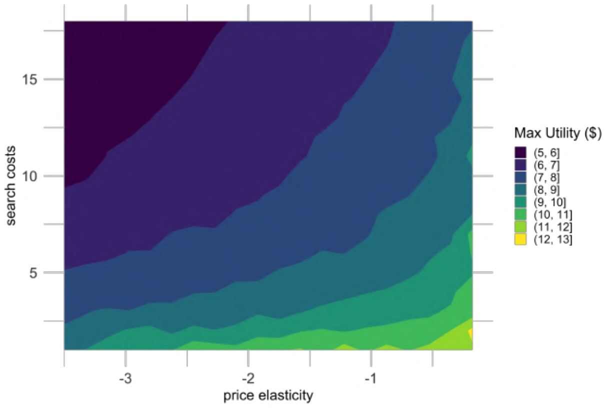
Next, we investigate the heterogeneity depending on both consumers' search costs and price sensitivity. We find that blocking third-party cookies will hurt consumers who are more price responsive and have high search costs (top left dark part of Figure 20a). Maximum achievable utility in this category is $5-6. The mechanism is as follows: the algorithm is unable to pick up that consumers are price responsive because it uses aggregate data, and, therefore, shows higher priced items. Because consumers have high search costs they are more likely to leave the website without purchasing (in which case, the utility is zero). Alternatively, they buy an item with lower utility than they would otherwise get in the full personalized condition. Consumers who have low price sensitivity and low search costs get the highest possible utility among all groups (right bottom part of Figure 20a).

Figure 20b shows that using the probabilistic recognition algorithm, we can increase welfare for the consumers in the mid range: the dark blue region in Figure 20a moves from $(6,7] region to $(7,9] range. Less price responsive groups with lower search costs are also better off. The takeaway is that while the algorithm can benefit consumers that have lower search costs and are more prone to re-visit the website, it is hard to help more vulnerable sets of consumers: more price sensitive and those with high search costs, because they leave without arriving back. Thus, these results call for alternative regulation that would take into account that privacy regulation hurts vulnerable groups of consumers more than others.

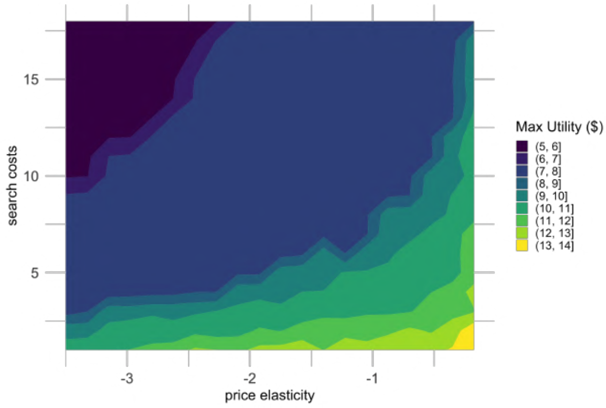
Next, we evaluate the impact of privacy restrictions on the seller outcomes. We divide sellers into two groups based on the historical revenue earned: 10th percentile-revenue sellers and 90th

Figure 20: Heterogeneity in consumer welfare by search costs and price elasticities

(a) Chrome 2024 event



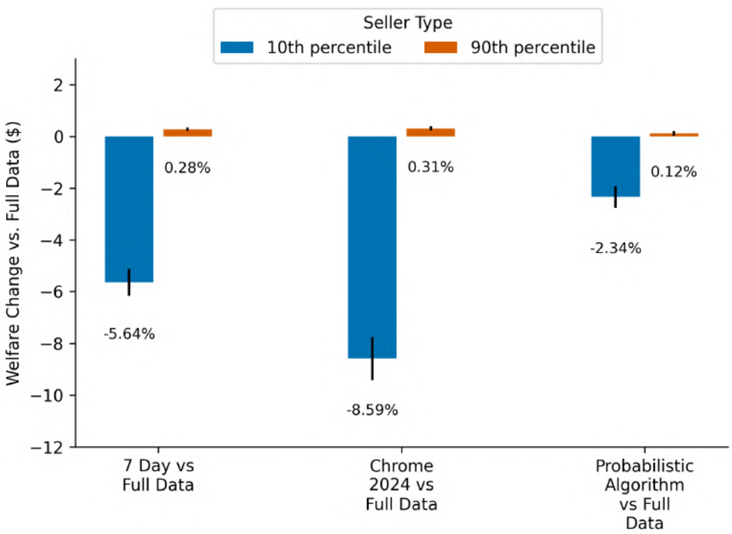
(b) Probabilistic Recognition



Notes. This figure shows the heterogeneity analysis to investigate the impact of privacy restrictions and alternative algorithm on consumer welfare. We simulate the consumer welfare by different price elasticity and search costs group.

percentile-revenue sellers (Figure 21). We find that privacy restrictions do not affect large sellers' revenue as much as they affect smaller sellers. First-party cookie restrictions decrease smaller sellers' revenue by 5.64%, and third-party Chrome restrictions decrease revenue by 8.59%.

Figure 21: Counterfactual results: seller outcomes

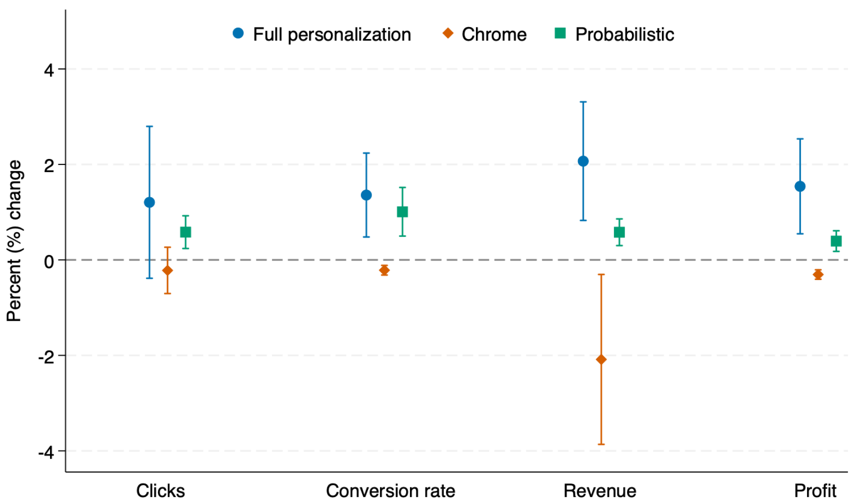


Notes: This figure shows the results of the counterfactual simulations on seller outcomes. Blue bars correspond to sellers that were in the 10th percentile by historical revenue, and orange bars are the 90th percentile sellers. The x-axis shows the three counterfactuals and the y-axis shows the % revenue change as a result of privacy restrictions.

The mechanism is as follows: with distorted data, personalization algorithms do not show relevant items and sometimes tend to resort to showing large sellers' products. Therefore, the privacy restrictions lead to the decrease in smaller sellers' revenue. Our proposed algorithm performs better because in the cases where it correctly predicts the association between devices and consumers, it will be equivalent to the full personalization algorithm. The takeaway from this part of the analysis is that privacy restrictions disproportionately hurt smaller sellers on the platform, and this is something current regulation may want to fix.

Finally, we investigate the consumers' search/purchase outcomes which directly translate to platform's revenue and profit. We use the platform's detailed cost data to calculate profits. Figure 22 shows three sets of results: (i) comparison between full personalization and non-personalized rankings (full personalization), (ii) comparison between Chrome outcomes and non-personalized rankings (Chrome), and (iii) probabilistic algorithm versus non-personalized rankings (probabilistic). Note that in this part of the analysis we use the non-personalized rankings as a benchmark to put the experimental results in perspective. Thus, blue bars are plotted experimental results from Section 4, and the rest are calculated based on the model.

FIGURE 22: Counterfactual results: consumer and platform outcomes



Notes: This figure shows the shares of traffic that arrive from different sources. The total doesn't sum up to 100, because one consumer can arrive through multiple channels. Majority of consumers would be recognized via first-party cookies given that they arrive from Google product ads or Direct traffic. For the consumers who arrive from display advertising, the platform relies on third-party cookies.

First, we see that both click and conversion rates, and revenue decrease significantly in the Chrome condition (orange bars in Figure 22). However, probabilistic recognition algorithm can recover substantial parts of the benefits from personalization (green bars), although all metrics are still lower compared to the full personalization condition. We also see that platform profits are less impacted than their revenues. The reason is that bestseller products are typically located in platform-owned distribution centers, which make their shipping costs lower. Therefore, for the platform it is cheaper to ship bestseller products rather than smaller sellers' products. Therefore, although we see significant decreases in revenue as a result of Chrome restrictions, the profit is impacted less. Moreover, we also have evidence that larger (bestseller) products tend to give more discounts and allowances on the wholesale costs, which additionally helps maintain profits even under privacy restrictions (Table E13).

There are three main takeaways from the counterfactual analysis described in this section. First, privacy restrictions hurt consumers who are more price responsive and have higher search costs. Second, privacy restrictions disproportionately hurt smaller sellers who rely on personalization algorithms to gain prominence on the website. Finally, platform revenue is hurt by the privacy restrictions, while its profit is relatively not impacted because of the different cost structure for large and small sellers.

6.1. Discussion

There are several concerns that could arise related to the counterfactuals. In the first counterfactual, we use the 7-day model to generate rankings. One could argue that the platform could keep using the entire historical data in order to generate rankings. However, we view the counterfactual as the way to evaluate the long-term impact of these privacy restrictions. In the long-term, the value of the historical data may fall either due to different product assortment or changes in consumer preferences, which is why the platform will use the available *more recent* data.

In the second counterfactual, the underlying assumption is that platform keeps advertising. However, Chrome restrictions could affect the ability to advertise to begin with, in which case consumers would not see the platform's ads at all. However, since Chrome is offering differential privacy and other Privacy Sandbox-based solutions,[50] we assume that the platform will be able to advertise in some form. For instance, it would not be able to show re-targeted ads featuring products similar to the browsed ones, but it could show generic platform ads.

Finally, with all counterfactuals, one could worry that consumers may start authenticating in case they see irrelevant products. This is one of the limitations of the study where we do not know how consumers' authentication decisions will change as a result of privacy restrictions. Moreover, platforms may offer other solutions to make authentication process easier, such as biometric login. As a result, more consumers will login voluntarily and the platform will continue collecting the data directly from consumers without relying on cookies. These are platform actions that we cannot take into account in one paper, but this could open an interesting future stream of work that could help platforms mitigate the negative consequences of privacy restrictions.

7. Conclusion

In this paper, we empirically study the effects of personalization and privacy restrictions on the retail platform, its consumers and sellers. To do that, we use large-scale field experiments ran with Wayfair, their detailed clickstream data and the platform's personalization algorithm. First, we use the experiments to quantify the reduced-form measures of welfare gains from personalization on

[50] See the Privacy Sandbox article for more details.

consumers, sellers and the platform. Second, we develop a novel multi-session consumer search model in the presence of personalized recommendations to evaluate how current and upcoming privacy restrictions affect consumers, sellers and the retail platform. Finally, we evaluate how probabilistic recognition solutions may help mitigate the negative consequences of privacy restrictions.

Experimental results suggest that personalization benefits consumers, smaller sellers and the platform. Consumers in the personalized condition buy more expensive but higher quality products, and they are 10% less likely to return the purchased product. Second, we show that smaller sellers are 15% more likely to be shown higher up on the product ranking pages, and therefore generate substantial part of their revenue from personalized impressions. These results suggest that personalization leads to better matches of consumers and sellers, and that it gives chance for the smaller sellers to grow their businesses.

To evaluate how alternative privacy policies impact consumers and sellers, we develop a structural model of consumer search and learning in the presence of recommendations. We re-train platform's personalization algorithm with distorted data mimicking privacy restrictions of interest to generate counterfactual recommendations. Next, we use the model to simulate how consumers' search and purchase behavior change under counterfactual recommendations. The results imply that a consumer gets $25 welfare gain from personalization per purchase. However, privacy restrictions, such as Chrome blocking third-party cookie tracking or Safari blocking first-party cookie tracking will decrease these welfare gains by up to 50%. More price responsive consumers and smaller sellers are hurt the most.

To the best of our knowledge, this is the first paper that empirically studies the impact of current and upcoming privacy restrictions on personalization algorithms and their subsequent effects on consumers, sellers, and platforms' outcomes. The question has significant importance because personalization plays key role in navigating consumers through thousands or even millions of products that platforms carry. Our results suggest that privacy restrictions disproportionately hurt smaller sellers and more price responsive consumers. This calls to alternative privacy regulations that address privacy concerns without unduly burdening small businesses or hindering consumer experiences. We show that probabilistic recognition algorithms can help platforms mitigate the negative consequences of the privacy restrictions, striking a balance between privacy and personalization.

References

Aridor, Guy, Yeon-Koo Che, William Nelson, and Tobias Salz, "The Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR," *SSRN Electronic Journal*, 2020. (Cited on page(s) 5)

Auer, Peter, "Using confidence bounds for exploitation-exploration trade-offs," *Journal of Machine Learning Research*, 2002, *3* (Nov), 397–422. (Cited on page(s) 6)

Bell, Sean and Kavita Bala, "Learning visual similarity for product design with convolutional neural networks," *ACM Transactions on Graphics*, jul 2015, *34* (4), 1–10. (Cited on page(s) 12, 79, 81)

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890. (Cited on page(s) 33)

Bottou, Léon, Frank E Curtis, and Jorge Nocedal, "Optimization methods for large-scale machine learning," *SIAM review*, 2018, *60* (2), 223–311. (Cited on page(s) 80)

Bronnenberg, Bart J., Jun B. Kim, and Carl F. Mela, "Zooming In on Choice: How Do Consumers Search for Cameras Online?," *Marketing Science*, sep 2016, *35* (5), 693–712. (Cited on page(s) 6, 26, 73)

Chen, Yuxin and Song Yao, "Sequential search with refinement: Model and application with click-stream data," *Management Science*, 2017, *63* (12), 4345–4365. (Cited on page(s) 27)

Choi, Hana and Carl F Mela, "Monetizing online marketplaces," *Marketing Science*, 2019, *38* (6), 948–972. (Cited on page(s) 5)

Coey, Dominic and Michael Bailey, "People and cookies: Imperfect treatment assignment in online experiments," in "Proceedings of the 25th International Conference on World Wide Web" 2016, pp. 1103–1111. (Cited on page(s) 5)

Compiani, Giovanni, Gregory Lewis, Sida Peng, and Will Wang, "Online search and product rankings: A double index approach," *Available at SSRN 3898134*, 2021. (Cited on page(s) 5)

Covington, Paul, Jay Adams, and Emre Sargin, "Deep neural networks for youtube recommendations," in "Proceedings of the 10th ACM conference on recommender systems" 2016, pp. 191–198. (Cited on page(s) 4)

Eksombatchai, Chantat, Pranav Jindal, Jerry Zitao Liu, Yuchen Liu, Rahul Sharma, Charles Sugnet, Mark Ulrich, and Jure Leskovec, "Pixie: A system for recommending

3+ billion items to 200+ million users in real-time," in "Proceedings of the 2018 world wide web conference" 2018, pp. 1775–1784. (Cited on page(s) 4)

Gibbard, Peter, "Search with Two Stages of Information Acquisition," *Available at SSRN 4331038*, 2022. (Cited on page(s) 5, 25)

Goldberg, Samuel, Garrett Johnson, and Scott Shriver, "Regulating privacy online: An economic evaluation of the GDPR," *Available at SSRN 3421731*, 2019. (Cited on page(s) 5)

Golovin, Daniel, Andreas Krause, and Matthew Streeter, "Online Submodular Maximization under a Matroid Constraint with Application to Learning Assignments," 2014. (Cited on page(s) 57)

Greminger, Rafael P, "Heterogeneous Position Effects and the Power of Rankings," *arXiv preprint arXiv:2210.16408*, 2022. (Cited on page(s) 5, 25, 30)

Hadsell, Raia, Sumit Chopra, and Yann LeCun, "Dimensionality reduction by learning an invariant mapping," in "2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)," Vol. 2 IEEE 2006, pp. 1735–1742. (Cited on page(s) 79, 80)

Hodgson, Charles and Gregory Lewis, "You Can Lead a Horse to Water: Spatial Learning and Path Dependence in Consumer Search," 2022. (Cited on page(s) 6, 26, 27, 28)

Honka, Elisabeth and Pradeep Chintagunta, "Simultaneous or sequential? Search strategies in the US auto insurance industry," *Marketing Science*, 2017, *36* (1), 21–42. (Cited on page(s) 5)

Ibrahim, Joseph G, Ming-Hui Chen, Yeongjin Gwon, and Fang Chen, "The power prior: theory and applications," *Statistics in medicine*, 2015, *34* (28), 3724–3749. (Cited on page(s) 32)

Johnson, Garrett A, Scott K Shriver, and Samuel G Goldberg, "Privacy and market concentration: intended and unintended consequences of the GDPR," *Management Science*, 2023. (Cited on page(s) 5)

Johnson, Garrett A., Scott K. Shriver, and Shaoyin Du, "Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry?," *Marketing Science*, jan 2020, *39* (1), 33–51. (Cited on page(s) 53)

Keane, Michael, *Marketing Letters*, 1997, *8* (3), 307–322. (Cited on page(s) 35)

Keskar, Nitish Shirish, Dheevatsa Mudigere, Jorge Nocedal, Mikhail Smelyanskiy, and Ping Tak Peter Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," *arXiv preprint arXiv:1609.04836*, 2016. (Cited on page(s) 35)

Kim, Jun B, Paulo Albuquerque, and Bart J Bronnenberg, "Online demand under limited consumer search," *Marketing science*, 2010, *29* (6), 1001–1023. (Cited on page(s) 5)

Korganbekova, Malika and Cole Zuber, "Probabilistic Identity Graphs for Personalization," *Working paper*, 2023. (Cited on page(s) 1, 4, 40)

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, 2017, *60* (6), 84–90. (Cited on page(s) 80)

Lin, Tesary and Sanjog Misra, "Frontiers: The Identity Fragmentation Bias," *Marketing Science*, may 2022, *41* (3), 433–440. (Cited on page(s) 5)

Linden, Greg, Brent Smith, and Jeremy York, "Amazon. com recommendations: Item-to-item collaborative filtering," *IEEE Internet computing*, 2003, *7* (1), 76–80. (Cited on page(s) 4)

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan, "Role of forgetting in memory-based choice decisions: A structural model," *Quantitative Marketing and Economics*, 2004, *2*, 107–140. (Cited on page(s) 32)

Mei, M Jeffrey, Cole Zuber, and Yasaman Khazaeni, "A Lightweight Transformer for Next-Item Product Recommendation," in "Proceedings of the 16th ACM Conference on Recommender Systems" 2022, pp. 546–549. (Cited on page(s) 4, 83)

Morozov, Ilya, "Measuring benefits from new products in markets with information frictions," *Management Science*, 2023. (Cited on page(s) 5)

Nemhauser, G. L., L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Mathematical Programming*, dec 1978, *14* (1), 265–294. (Cited on page(s) 57)

Seiler, Stephan, "The impact of search costs on consumer behavior: A dynamic approach," *Quantitative Marketing and Economics*, 2013, *11*, 155–203. (Cited on page(s) 5)

_ **and Fabio Pinna**, "Estimating search benefits from path-tracking data: measurement and determinants," *Marketing Science*, 2017, *36* (4), 565–589. (Cited on page(s) 5)

Sun, Tianshu, Zhe Yuan, Chunxiao Li, Kaifu Zhang, and Jun Xu, "The value of personal data in internet commerce: A high-stake field experiment on data regulation policy," *Available at SSRN 3962157*, 2021. (Cited on page(s) 5)

Ursu, Raluca M, "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions," *Marketing Science*, 2018, *37* (4), 530–552. (Cited on page(s) 5)

Vedaldi, A, Y Jia, E Shelhamer, J Donahue, S Karayev, J Long, and T Darrell, "Convolutional architecture for fast feature embedding," *Cornell University*, 2014. (Cited on page(s) 79)

Weitzman, Martin, *Optimal search for the best alternative*, Vol. 78, Department of Energy, 1978. (Cited on page(s) 5, 6)

Wernerfelt, Nils, Anna Tuchman, Bradley Shapiro, and Robert Moakler, "Estimating the Value of Offsite Data to Advertisers on Meta," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2022, (114). (Cited on page(s) 5)

Zhao, Yu, Pinar Yildirim, and Pradeep K Chintagunta, "Privacy regulations and online search friction: Evidence from GDPR," *Available at SSRN 3903599*, 2021. (Cited on page(s) 5)

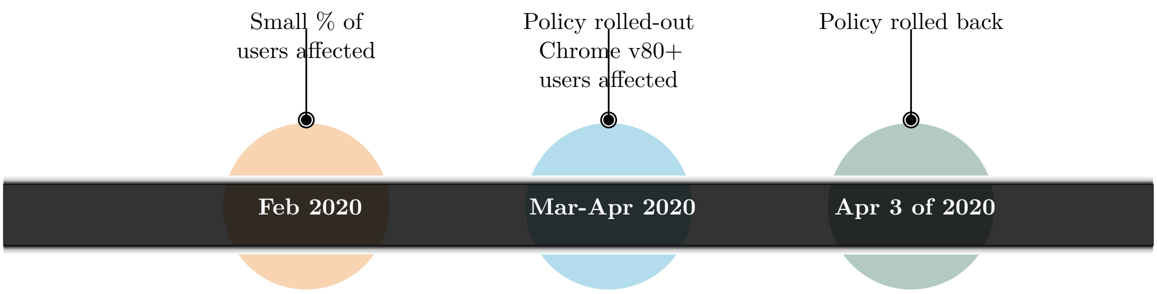
Online Appendix

Table of Contents

A. Current privacy policies and their impact

To validate the countefactuals, we identified a natural experiment: a change in browser policies that, while announced in advance, was rolled out unexpectedly, leaving the focal platform unprepared for it. Namely, in March of 2020 Chrome rolled out SameSite policy updates that blocked the use of cookies in third-party contexts and affected the traffic that was originating from advertising. The policy allowed to use cookies in the first-party context, but blocked all third-party requests. As a result, if consumer saw Wayfair ad on a third-party website, e.g. weather.com and clicked on it, as she navigates to wayfair.com, consumer's cookies were reset. Cookie resetting means that the consumer gets a complete new user identifier and Wayfair does not have access to consumer's browsing histories. As a result, Wayfair's ability to personalize is limited. The policy was short-lived in that Chrome rolled it back in April 2020, worried that a large number of advertisers depended on third-party cookies in Covid times.[51]. Despite the transient nature of the policy, it had significant effects on platform's ability to recognize the traffic and to personalize. Below, we describe the policy in more detail and show the impact on consumers and the platform. We use this natural experiment to check how well the model can replicate the observed data.

Figure A1: Timeline of the Chrome SameSite update releases



Notes: Chrome released the updates gradually for a fraction of Chrome Canary and Dev users starting October 2019, and in March of 2020, Chrome increased the target population affecting most Chrome users. However, on April 3, 2020, in light of global pandemic, Chrome decided to roll back the changes not to hurt advertisers who rely on third-party data. Afterwards, starting July 14, 2020, they rolled out the changes again with the rollout population gradually increasing on July 28, 2020 and on August 11, 2020 - the changes were rolled out to 100% of Chrome Stable users. Source: Official Chrome SameSite updates.

Chrome SameSite policy updates are secure-by-default changes that protect all cookies from external access unless otherwise specified by the user. Most Chrome version 80+ users were affected and didn't change the default setting.[52] Academic research also found that even when offered an opportunity to opt out from online advertising very few consumers choose to change the default settings (Johnson, Shriver and Du, 2020). As a result, during one month in March-April of 2020, the platform was limited in its ability to recognize Chrome users unless they logged in to the website

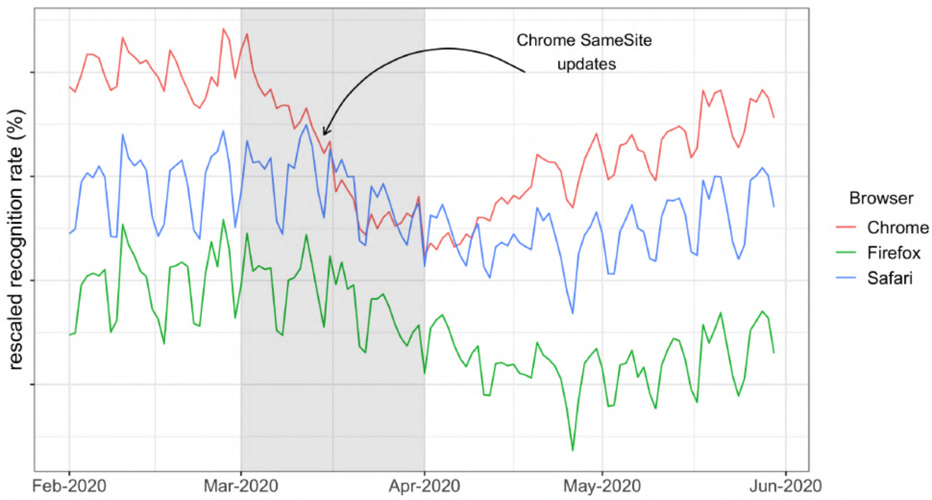
[51] For detailed policy releases, see the SameSite updates.

[52] As was documented for Apple's App Tracking Transparency (ATT) changes, consumers tend to stick with the default browser and app settings and do not change them. See Harvard Business Review: Apple Is Changing How Digital Ads Work. Are Advertisers Prepared?

voluntarily.

Figure A2 shows that the recognition rates on Chrome dropped drastically during the policy period. The x-axis of Figure A2 shows month-year. The roll-out of Chrome SameSite policy in March-April 2020 highlighted in grey, and y-axis shows scaled device recognition rates. We observe similar trends in the recognition rates before March 2020 (indicating that parallel trend assumption is satisfied). However, Chrome (red line) exhibits a drop in the recognition rates in March-April 2020, recovering only in April when the changes were rolled back.

FIGURE A2: Recognition rates per major browser (February 2020 - June 2020)



Notes. This figure shows the platform's recognition rates before, during and after the Chrome changes (February - June 2020). The shaded area is the period during which Chrome policy kicked in. The graph explicitly shows recognition on Chrome browser drastically decreased during that period and only started recovering in April 2020 (red line). The y-axis is hidden for data sensitivity reasons.

To further quantify the extent of recognition drop, we estimated difference-in-differences regression on the constructed browser-day level panel that documents daily recognition among different browsers. In Equation 21, j is the browser, t - day, $1\{j = Chrome\}$ is the indicator for Chrome, $1\{after\}$ is the indicator for the period between March 2, 2020 and different dates after April 3, 2020.

$$y_{jt} = \alpha + \beta_1 1\{j = Chrome\} + \beta_2 1\{after\} + \beta_3 1\{j = Chrome\}1\{after\} + \varepsilon_{jt} \tag{21}$$

Table A1, Column 1, shows that Chrome recognition rates dropped by 3.2 percentage points after the policy was implemented. The results are robust to different policy window definitions. Columns (2) and (3) of Table A1 show that under a narrower window, the effect of recognition rates is around 3.2-3.7 percentage points. We cannot disclose the intercept for data sensitivity reasons, but each percentage point in recognition amounts to the loss of data for millions of consumers.

TABLE A1: Difference-in-differences results for Chrome policy-related recognition rates

	Dependent variable:		
	Recognition rate		
	Feb 1-May 1, 2020	Feb 1-Apr 3, 2020	Feb 1 -Apr 15, 2020
	(1)	(2)	(3)
After	1.620^{***}	-0.268	0.302
	(0.481)	(0.500)	(0.527)
Treatment	7.303^{***}	7.868^{***}	7.471^{***}
	(0.442)	(0.623)	(0.631)
After$\times$ Treatment	-3.220^{***}	-3.785^{***}	-3.388^{***}
	(0.834)	(0.867)	(0.914)
Observations	330	180	195

Notes. This table reports the results of the difference-in-differences estimation (Equation 21) investigating the effect of Chrome SameSite policy on platform's recognition rates. Data is at the browser-day level. Intercept is hidden for data sensitivity reasons. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table A2 shows that the reduction in recognition rates resulted in lower platform revenue. Platform revenue decreased by 2.8 - 3.2 percent after the Chrome policy was rolled out.

TABLE A2: Effect of Chrome on revenue

	Dependent variable:	
	Scaled revenue	
	Feb-April 2020	Feb-May 2020
	(1)	(2)
After$\times$ Treatment	-0.032^{***}	-0.028^{**}
	(0.012)	(0.011)
Observations	180	330

Notes. This table reports the interaction terms from estimation of specification 21. We investigate the effect of Chrome SameSite updates on the platform revenue. Intercepts are hidden for data sensitivity reasons. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table A3 shows that the drop in revenue was driven by lower conversion rates after Chrome policy change (Column 4) and lower number of orders (Column 5). Consumers click more (Column 1) which could mean they incur higher search costs to find relevant items after privacy restrictions were introduced. We do not find significant differences in add-to-cart or basket page landing probability, but conversion rates drop significantly.

Finally, we evaluate the impact of the privacy restriction on the data that was

Data impact for the training models. Data can be used in two ways: first, sequence of products searched is important; second, sequence of events (actions) performed on the website by a consumer. We find that total rows of training data goes down by 5% and total number of products goes down by 2%. Thus, training data amount is definitely impacted even with one month of the policy changes.

Table A3: Impact of Chrome SameSite changes on consumer outcomes

	Purchase funnel				*Purchase outcomes*	
	(1) Clicks	(2) Add-to-cart	(3) Basket page	(4) Converted	(5) Orders	(6) Average revenue
Chrome	1.207***	-1.270***	-1.168***	-1.037***	-1.105***	8.258***
	(0.385)	(0.098)	(0.094)	(0.049)	(0.053)	(2.546)
After	0.748**	0.973***	1.073***	0.661***	0.780***	1.001
	(0.352)	(0.090)	(0.086)	(0.045)	(0.048)	(2.327)
After× Chrome	0.859*	-0.127	-0.160	-0.216***	-0.297***	-6.251*
	(0.498)	(0.127)	(0.122)	(0.063)	(0.068)	(3.291)
Observations	304	304	304	304	304	304

Notes. This table reports the results of the difference-in-differences estimation (Equation 21) investigating the effect of Chrome SameSite policy on consumers' search and purchase outcomes. Intercept is hidden for data sensitivity reasons. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table A4: Data impact as a result of Chrome SameSite changes

	Total rows	Log products
Treatment	3.282***	0.085***
	(0.251)	(0.008)
After	7.035***	0.141***
	(0.456)	(0.014)
After× Treatment	-3.123***	-0.051***
	(0.230)	(0.007)
Product FE	2,437,195	2,437,195

Notes. This table reports the results of the difference-in-differences estimation (Equation 21) investigating the effect of Chrome SameSite policy on the scaled number of rows and log products. Intercept is hidden for data sensitivity reasons. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

B. Proof of near-optimality of the UCB algorithm with search costs

The proof is built on the material in Golovin et al. (2014) and Nemhauser et al. (1978). The main difference is that in Engineering it is common not to account for potential search costs, but in our setting it is crucial to account for the search costs.

Consider a problem when consumer learns a payoff function $f(x)$, $x \in D$ where x is an action taken by a consumer. The goal is to maximize the function $f(x_t)$ with respect to action x_t, i.e. $\max_{x_t} f(x_t)$. In our setting, the action is the product to click on. Instantaneous regret is defined as $r_t = f(x^*) - f(x_t)$ and the cumulative regret is $R_T = \sum_{t=1}^{T} r_t$. Intuitively, regret represents the loss from clicking on a product that is different from the product that maximizes the payoff function $f(x_t)$. Consumer's goal is to minimize the cumulative regret R_T.[53]

Proposition 1. *Searching using Upper Confidence Bound to learn a function $f(x)$ leads to a finite regret bound of $O(\sqrt{T\gamma_T log|D|})$ with high probability $1-\delta$, $\delta \in (0,1)$. Regret bound is*

$$P[R_T \leq \sqrt{C_1 T \beta_T \gamma_T} \quad \forall T \geq 1] \geq 1 - \delta$$

where $C_1 = 8/log(1+\sigma^{-2})$ and T is the number of rounds of sampling an individual point.

Proof. Fix some $t \geq 1$ and $x \in D$. Conditional on all the utility values sampled so far, $u_1, ..., u_{t-1}$, $\boldsymbol{x}_1, .., \boldsymbol{x}_{t-1}$ are deterministic and subsequent payoff functions are drawn from $f(x) \sim N(\mu_{t-1}(\boldsymbol{x}), \sigma^2_{t-1}(\boldsymbol{x}))$. For a standard normal variable $r \sim N(0,1)$, we can show that

$$\begin{aligned} P[r > c] = \frac{1}{\sqrt{2\pi}} \int_c^\infty e^{-r^2/2} dr = \frac{1}{2\pi} \int_c^\infty e^{-r^2/2 + c^2/2} dr \\ = e^{-c^2/2} \frac{1}{2\pi} \int_c^\infty e^{-(r-c)^2/2} e^{-c(r-c)} dr \end{aligned}$$

Because $e^{-c(r-c)} \leq 1$ and the rest of the integral resembles Gaussian density integrated from c to ∞, we can write

$$e^{-c^2/2} \frac{1}{2\pi} \int_c^\infty e^{-(r-c)^2/2} e^{-c(r-c)} dr \leq \frac{1}{2} e^{-c^2/2}$$

Using $r = (f(\boldsymbol{x}) - \mu_{t-1}(\boldsymbol{x}))/\sigma_{t-1}(x)$ and $c = \eta - \dfrac{c_{t-1}(x)}{\sigma_{t-1}(x)}$, we get that for any given $x \in D$ and timepoint $t \geq 1$,

$$P[|f(\boldsymbol{x}) - \mu_{t-1}(\boldsymbol{x})| > \eta\sigma_{t-1}(x) - c_{t-1}(x)] \leq e^{-(\eta\sigma_{t-1}(x) - c_{t-1}(x))^2/2}$$

[53] We re-write the objective function in this way for tractability. Intuitively, minimizing cumulative regret means that consumer wants to reach the best possible product with minimal regret, e.g. time spent on clicking on 'wrong' products.

Applying the union bound over all $\boldsymbol{x} \in D$, we get

$$P\left\{ \cup_{\boldsymbol{x}\in D} |f(\boldsymbol{x}) - \mu_{t-1}(\boldsymbol{x}))| > \eta\sigma_{t-1}(x_t) - c_{t-1}(x_t) \right\} \leq \sum_{\boldsymbol{x}\in D} P[|f(\boldsymbol{x} - \mu_{t-1}(\boldsymbol{x}))| > \eta\sigma_{t-1}(\boldsymbol{x}) - c_{t-1}(x_t)]$$
$$\leq |D|e^{-(\eta - c_{t-1}(x_t)/\sigma_{t-1}(x))^2/2}$$

By defining the variables appropriately, we can apply union to all timepoints $t \in \mathcal{N}$ to get

$$P\left\{ \cup_{\boldsymbol{x}\in D} |f(\boldsymbol{x}) - \mu_{t-1}(\boldsymbol{x}))| > \eta\sigma_{t-1}(x_t) - c_{t-1}(x_t) \right\} \leq \sum_{t=1}^{\infty} P[|f(\boldsymbol{x} - \mu_{t-1}(\boldsymbol{x}))| > \eta\sigma_{t-1}(x_t) - c_{t-1}(x_t)]$$
$$\leq \sum_{t=1}^{\infty} \frac{\delta}{\pi_t} = \delta$$

If we change the inequality to upper bound the term $|f(\boldsymbol{x}) - \mu_{t-1}(\boldsymbol{x}))|$ for all $\boldsymbol{x} \in D$ and all $t \geq 1$, we can write

$$|f(\boldsymbol{x}) - \mu_{t-1}(\boldsymbol{x}))| \leq \eta\sigma_{t-1}(x_t) - c_{t-1}(x_t)$$

Next, we construct a bound for the instantaneous regret function: $r_t = f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)$. Upper Confidence Bound algorithm specifies choosing $\boldsymbol{x}_t$ as the argmax of $\mu_{t-1}(\boldsymbol{x}_t) + \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}_t)$ at each timestep. Thus, we have

$$\mu_{t-1}(x_t) + \eta\sigma_{t-1}(x) - c_{t-1}(x_t) \geq \mu_{t-1}(x^*) + \eta\sigma_{t-1}(x^*) - c_{t-1}(x^*) \tag{22}$$

This can be rewritten as

$$r_t = f(x^*) - f(x_t) \leq \eta\sigma_{t-1}(x^*) - c_{t-1}(x^*) + f(x_t) - \mu_{t-1}(x_t) \leq 2[\eta\sigma_{t-1}(x_t) - c_{t-1}(x)] \tag{23}$$

We now continue towards constructing regret bounds for the cumulative regret function. For a Gaussian process with a covariance matrix $\sigma^2\boldsymbol{I}$, the expression for the information gain can be written as

$$I(y_T; f_T) = H(y_T) - H(y_T|f_T) = H(y_T) - \frac{1}{2}log|2\pi e\sigma^2\boldsymbol{I}| \tag{24}$$

Since the determinant of the diagonal matrix is the product of the diagonal elements, we write

$$\frac{1}{2}log|2\pi e\sigma^2\boldsymbol{I}| = \frac{1}{2}\sum_{t=1}^{T} log(2\pi e\sigma^2) \tag{25}$$

We can expand out $H(y_T)$ as

$$H(y_T) = H(y_{T-1}) + H(y_T|y_{T-1}) = H(y_{T-1}) + \frac{1}{2}log(2\pi e(\sigma^2 + \sigma_{T-1}^2(x_T))) \tag{26}$$

We can write the variance term in the entropy expression as a sum of variances due to the fact that $x_1, ..., x_T$ are deterministic conditioned on y_{T-1}, and the conditional variance $\sigma^2_{T-1}(x_T)$ does not depend on y_{T-1}. Expansion of the entropy terms gives us

$$H(y_T) = \frac{1}{2}\sum_{t=1}^{T} log(2\pi e(\sigma^2 + \sigma^2_{t-1}(x_t))) \tag{27}$$

Substituting (27) and (25) into (24) gives us

$$\begin{aligned}
I(y_T; f_T) &= H(y_T) - H(y_T|f_T) \\
&= \frac{1}{2}\sum_{t=1}^{T} log(2\pi e(\sigma^2 + \sigma^2_{t-1}(x_t))) - \frac{1}{2}\sum_{t=1}^{T} log(2\pi e\sigma^2) \\
&= \frac{1}{2}\sum_{t=1}^{T} log\left(\frac{2\pi e(\sigma^2 + \sigma^2_{t-1}(x_t))}{2\pi e\sigma^2}\right) \\
&= \frac{1}{2}\sum_{t=1}^{T} log(1 + \sigma^{-2}\sigma^2_{t-1}(x_t))
\end{aligned}$$

Let's denote $\beta_t^{1/2} = \eta - \frac{c_{t-1}(x)}{\sigma_{t-1}(x)}$. We know that $r_t^2 \leq 4\beta_t\sigma^2_{t-1}(x_t)$, $\forall t \geq 1$ with probability $\geq 1-\delta$. Since β_t is nondecreasing for increasing $t \leq T$, we can write

$$4\beta_t\sigma^2_{t-1}(x_t) \leq 4\beta_T\sigma^2_{t-1}(x_t)$$

Using the restriction that $\kappa(x, x') \leq 1$ for all x, x' means that $\sigma^2_{t-1}(x_t) \leq 1$ for all t. For positive s,

$$\frac{\sigma^{-2}\sigma^2_{t-1}(x_t)}{log(1 + \sigma^{-2}\sigma^2_{t-1}(x_t))} \leq \frac{\sigma^{-2}}{log(1 + \sigma^{-2})}$$

$$\sigma^{-2}\sigma^2_{t-1}(x_t) \leq \frac{\sigma^{-2}}{log(1 + \sigma^{-2})} log(1 + \sigma^{-2}\sigma^2_{t-1}(x_t))$$

$$4\beta_T\sigma^2_{t-1}(x_t) = 4\beta_T\sigma^2(\sigma^{-2}\sigma^2_{t-1}(x_t)) \leq 4\beta_T\sigma^2 C_2 log(1 + \sigma^{-2}\sigma^2_{t-1}(x_t))$$

where $C_2 = \sigma^{-2}/log(1 + \sigma^{-2})$. We can combine the inequalities derived for the instantaneous regret and write

$$\begin{aligned}\sum_{t=1}^{T} r_t^2 \leq 4\beta_T \sum_{t=1}^{T} \sigma_{t-1}^2(x_t) &\leq 4\beta_T\sigma^2 C_2 \sum_{t=1}^{T} log(1+\sigma^{-2}\sigma_{t-1}^2(x_t)) \\ &= 8\sigma^2 C_2 \beta_T \left(\frac{1}{2}\sum_{t=1}^{T} log(1+\sigma^{-2}\sigma_{t-1}^2(x_t))\right) \\ &= C_1\beta_T I(y_T; f_T) \leq C_1 \beta_T \gamma_T\end{aligned}$$

where $C_1 = 8\sigma^2 C_2 = 8/log(1+\sigma^{-2})$.

Next, using Cauchy-Schwarz inequality, we get

$$R_T^2 = \left(\sum_{t=1}^{T}\right)^2 \leq \left(\sum_{t=1}^{T} r_t^2\right)\left(\sum_{t=1}^{T} 1\right) = T\sum_{t=1}^{T} r_t^2$$

Therefore, $R_T \leq \sqrt{T\sum_{t=1}^{T} r_t^2} \leq \sqrt{C_1 T \beta_T \gamma_T}$. This inequality holds with probability $\geq 1-\delta$. This concludes the proof of Proposition 1.

$\square$

C. Randomization Checks

This section provides randomization checks for the consumers that were part of the A/B tests we ran with the platform. All variables are scaled for data anonymity reasons. Table C5 shows that all p-values are large indicating that there are no significant a-priori differences between the treatment and control group consumers.

Table C5: Randomization check in dining chairs (return to p. 14)

	Non-personalized (1)	Personalized (2)	Difference (3)	p-value (4)
Historical Purchases ($)	886.98	937.77	-50.79 (-1.26)	0.21
Historical Quantity Bought	6.19	6.51	-0.32 (-1.22)	0.22
Estimated networth	348,657.63	346,654.85	2,002.77 (1.51)	0.13
Estimated income	70,847.86	70,779.64	68.22 (0.48)	0.63
Age	24.90	24.89	0.01 (0.52)	0.60
Home Value	175,067.06	175,177.83	-110.77 (-0.19)	0.85
Prices searched	87.24	87.29	-0.05 (-0.20)	0.84
Gender dummy (0,1)	0.85	0.85	0.00 (0.22)	0.82
Observations	319,783	315,484		

Notes. This table provides randomization check between the treatment and control groups in the experiment ran on the ranking pages. The numbers are scaled for data anonymity purposes.

D. Additional experimental results

These results are likely attenuated because not everyone will see personalized impressions. If we look only at the consumers that had history to personalize from then I'll see that the results are even higher. Note that because the experiment was randomized regardless of the historical patterns of consumers, we have a balanced sample across two groups. In the non-personalized group there are 1,799,908 consumers, while in the personalized group there're 1,806,831 consumers. The randomization checks are in the appendix. That means that around 36% of people in each group had historical searches on the website.

Table D6: Personalization experiment results: search and purchase outcomes (return to p. 16)

	Logistic				*OLS*		
	(1) Clicks	(2) Add-to-cart	(3) Basket page	(4) Converted	(5) Log(Revenue)	(6) Purchases	(7) Log(Profit)
Personalized	0.002	0.013***	0.011***	0.017***	0.009***	0.007***	0.006***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Intercept	1.016***	-1.271***	-1.382***	-2.275***	0.579***	0.253***	–
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Observations	9,818,022	9,818,022	9,818,022	9,818,022	9,818,022	9,818,022	9,818,022

Notes. This table reports the output from the estimation of equation 1 for all nine million consumers. Data is at the consumer-level. Columns (1)-(4) report the logistic specification and Columns (5)-(7) report the OLS specification results. Robust standard errors in parentheses. The intercept in profit Column (7) is hidden for data sensitivity reasons. Statistical significance: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D7: Effect of personalization on repeat visits and product returns (all consumers) (return to p. 19)

	Repeat purchases						*Returns*
	(1) 7 days	(2) 30 days	(3) 90 days	(4) 150 days	(5) 365 days	(6) 500 days	(7) product returns
Personalized	0.019***	0.012**	0.017***	0.014***	0.012***	0.017***	0.010***
	(0.007)	(0.005)	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)
Personalized× Personalized product							-0.104***
							(0.010)
Intercept	-2.083***	-1.203***	-0.305***	-0.374***	0.092***	0.303***	-2.764***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)
Observations	933,510	933,430	1,511,657	933,328	933,246	602,954	9,056,732

Notes. This table shows the effects of personalization on repeat purchases and product return rates for **all consumers** in the experiment. Columns (1) - (6) are estimated using logit version of 1. Data are at the consumer level. Column (7) is the estimation of Equation 3. Data are at the consumer-purchased product level. Consumers in the personalized group might buy the item that was part of the organic rankings and wasn't personalized to them and we control for that by interacting the treatment dummy with the indicator for whether the product was personalized. For statistical power, we've included all consumers who were shopping in dining chairs category and their visits to the same marketing category, i.e. dining chairs, chairs, Each column represents the set of people who purchased a dining chair and we check the probability they will purchased in 7, 30, 90 etc. days. Robust standard errors in parentheses. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

$$y_i = \alpha + \beta_1 treatment_i + \beta_2 filtered_i + \beta_3 treatment_i \times filtered_i + \varepsilon_i \tag{28}$$

Table D8: Personalization experiment results: search and purchase outcomes (consumers with browsing history) (return to p. 16)

	Logistic				*OLS*		
	(1) Clicks	(2) Add-to-cart	(3) Basket page	(4) Converted	(5) Log(Revenue)	(6) Purchases	(7) Log(Profit)
Personalized	0.000	0.014***	0.011***	0.018***	0.020***	0.017***	0.013***
	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Intercept	2.143***	-0.218***	-0.258***	-1.201***	1.443***	0.656***	–
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	3,121,011	3,121,011	3,121,011	3,121,011	3,121,011	3,121,011	3,121,011

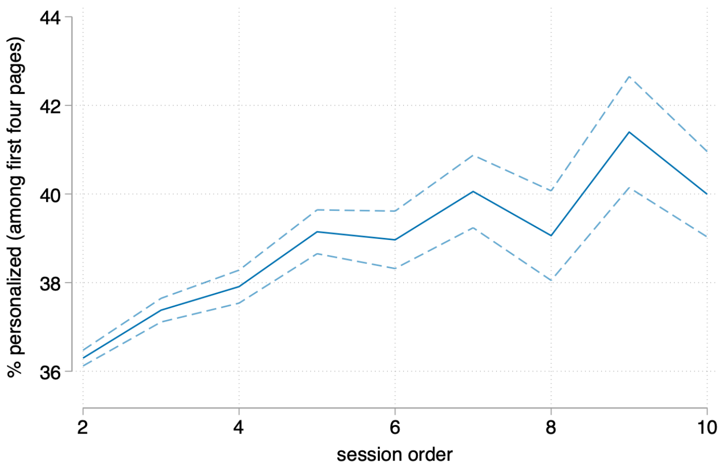
Notes. This table reports the output from the estimation of equation 1 for **the consumers with browsing history**. Data is at the consumer-level. Columns (1)-(4) report the logistic specification and Columns (5)-(7) report OLS specification results. the results of OLS regression on the experimental data. Robust standard errors in parentheses. Statistical significance: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D9: Further evidence that filtering doesn't change outcomes (return to p. 27)

	Logistic				*OLS*		
	(1) Clicks	(2) Add-to-cart	(3) Basket page	(4) Converted	(5) Log(Revenue)	(6) Purchases	(7) Log(Profit)
Personalized	0.010	0.013**	0.016***	0.015**	0.021**	0.026***	0.017**
	(0.013)	(0.006)	(0.006)	(0.006)	(0.009)	(0.008)	(0.007)
Filtered	0.143***	0.260***	0.230***	0.384***	0.617***	0.528***	0.501***
	(0.021)	(0.009)	(0.009)	(0.009)	(0.014)	(0.016)	(0.011)
Personalized× Filtered	-0.040	-0.007	-0.007	-0.001	0.001	-0.006	-0.004
	(0.029)	(0.012)	(0.012)	(0.013)	(0.020)	(0.023)	(0.016)
Intercept	2.957***	0.188***	0.096***	-0.962***	1.808***	0.975***	–
	(0.009)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)
Observations	635,267	635,267	635,267	635,267	635,267	635,267	635,267

Notes. This table reports the results of Equation 28. We investigate whether filtering plays an important part in changing consumer outcomes. Data is at the consumer level. Standard errors in parentheses. Intercept in Column (7) is hidden for data sensitivity reasons. Robust standard errors. Statistical significance: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure D3: Treatment Intensity (return to p. 15)



Notes. This figure shows the relationship between the session order and the % of the personalized products on the first four pages. Platform personalizes more as it collects more data on consumer clicks.

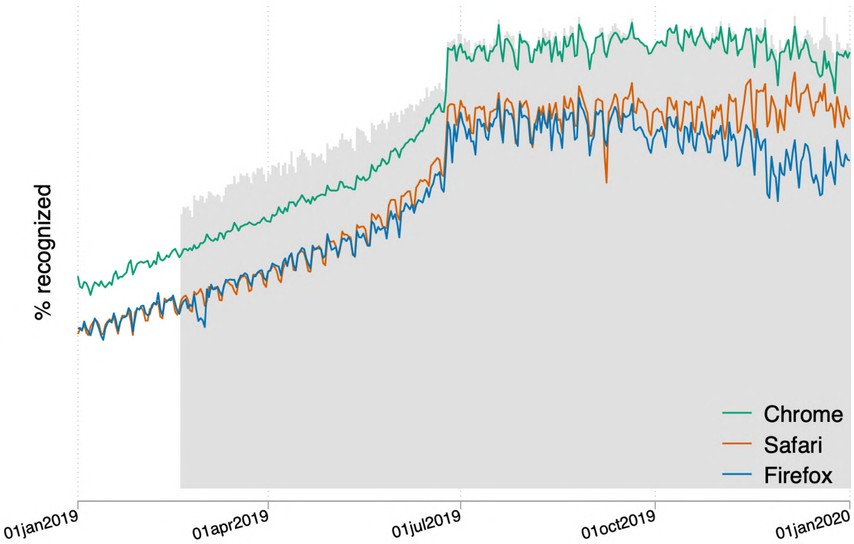
E. Additional descriptive statistics

Table E10: Number of devices and browsers used (return to p. 10)

	Observations	Mean	St.Dev.	Min	25%	50%	75%	Max
# Devices								
multi-session consumers	321,967	3.06	2.97	1.00	1.00	2.00	4.00	14.00
purchasing consumers	43,110	4.88	3.50	1.00	2.00	4.00	7.00	14.00
# Browsers								
multi-session consumers	321,967	1.31	0.54	1.00	1.00	1.00	2.00	6.00
purchasing consumers	43,110	1.02	0.15	1.00	1.00	1.00	1.00	3.00

Notes. This table reports the number of devices and browsers used by consumers who visited the websites for multiple sessions. Median consumer searches on two devices from the same browser, and purchasing consumers search more intensively and use four devices.

Figure E4: Recognition rates by browser in 2019 (return to p. 8)



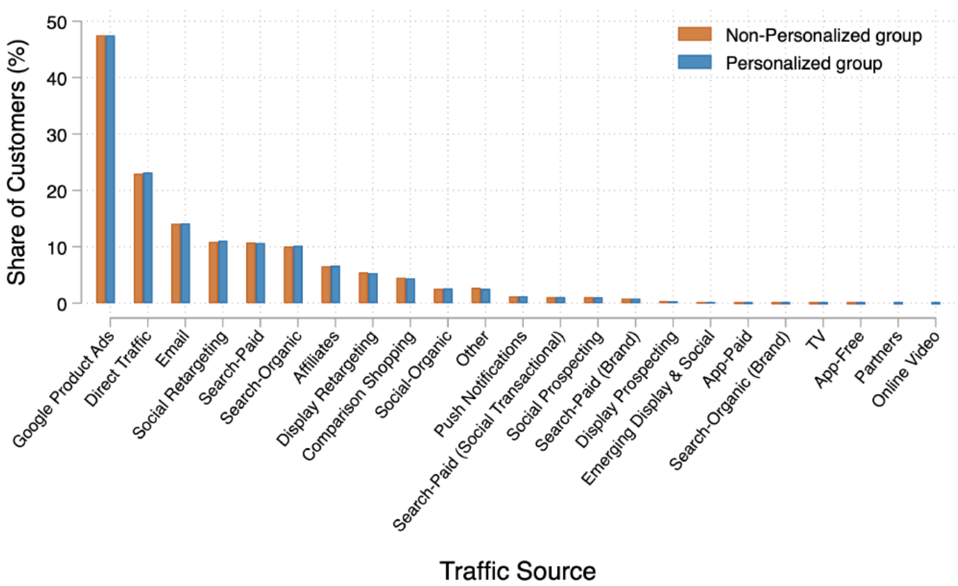
Notes. This figure illustrates the dynamics of recognition rates in 2019. The Safari ITP 2.1 that blocks first-party cookies was introduced on February 21 of 2019. However, we do not see significant differences in the recognition rates.

Table E11: Difference-in-differences effects of Safari privacy restrictions on recognition rates (return to p. 8)

	Recognition rate (%)		
	(1)	(2)	(3)
	Feb 21, 2019 to *Jan 31, 2019*	*Feb 21, 2019* to *Mar 21, 2019*	*Feb 21, 2019* to *Mar 1, 2019*
Safari	-5.284***	-4.693***	-4.766***
	(0.403)	(0.608)	(0.601)
After	16.899***	-12.162***	-12.876***
	(0.464)	(0.614)	(1.020)
After× Safari	0.593	-1.017	-0.326
	(0.682)	(0.827)	(1.179)
Observations	1,464	1,464	1,464

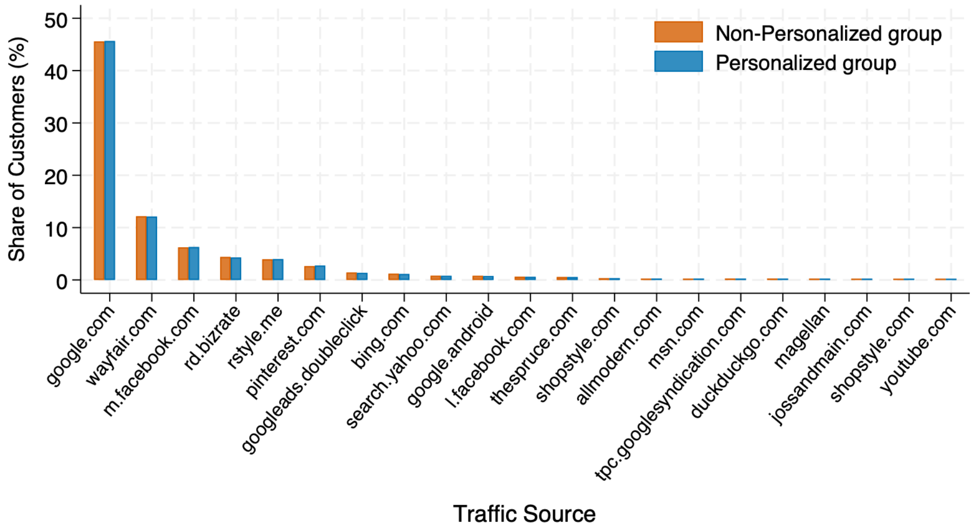
Notes: This table reports the results of the difference-in-differences estimation (Equation 21) investigating the effect of Safari policy on platform's recognition rates. Data is at the browser-day level. Intercept is hidden for data sensitivity reasons. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure E5: Consumer arrival by channel type (return to p. 27)



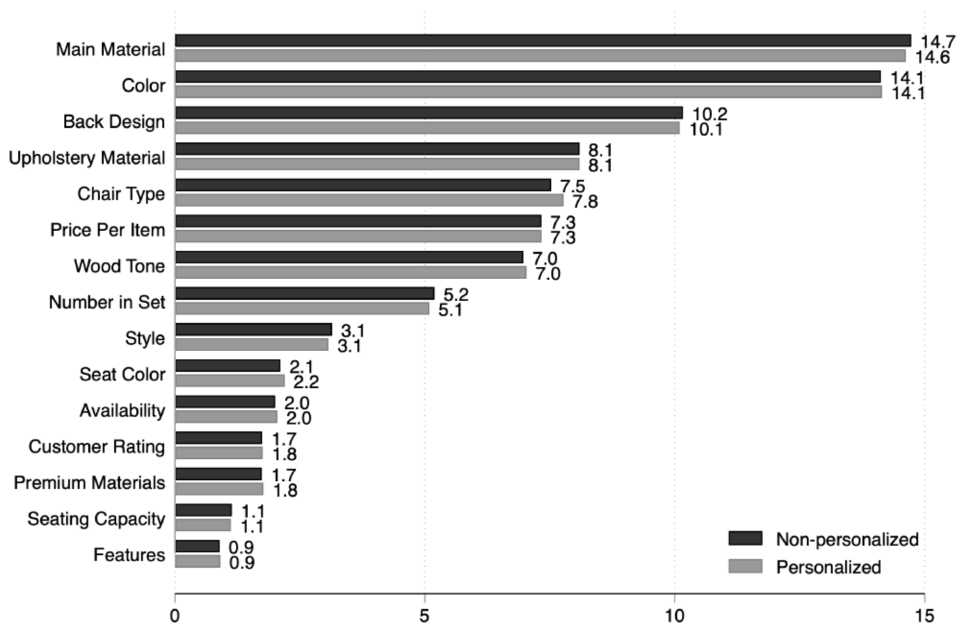
Notes. This figure shows the distribution of the inter-session arrival channels that consumers in the personalized and non-personalized groups use. The distribution seems balanced between two groups, which serves as an indication that personalization may not affect actions consumers take outside the website.

Figure E6: Consumer arrival by referral type (return to p. 27)



Notes. This figure shows the distribution of referral URLs consumers arrive from during re-visits. The distribution of websites seems to be balanced between personalized and non-personalized group, which serves as an indication that personalization may not affect actions consumers take outside the website.

Figure E7: Filtering behavior by personalized condition (return to p. 27)



Notes. This figure shows the distribution of the filters applied by consumers in the personalized and non-personalized groups. The figure highlights that conditional on filtering, consumers apply similar set of filters.

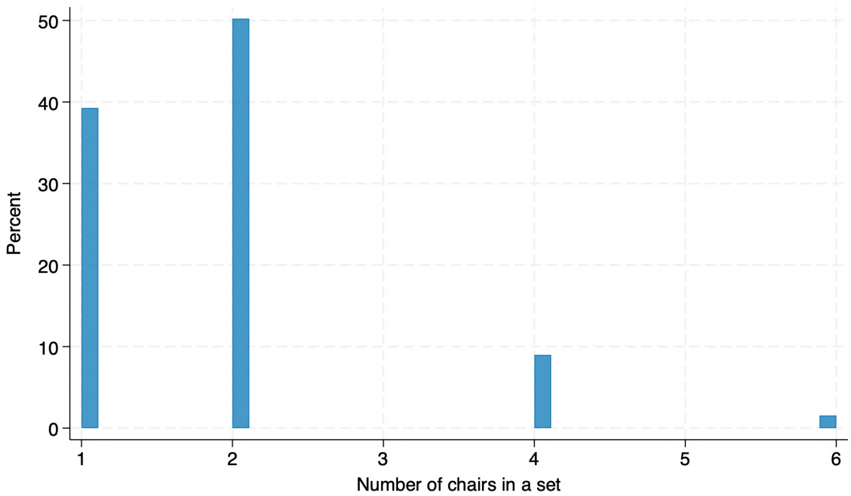
These graphs show that the personalization effect is consistent across different brand quintiles.

Table E12: Heterogeneity in product characteristics

	Observations	Min	Mean	Median	Max	St.Dev.
price ($)	35,873	4.46	431.99	349.99	2999.99	291.27
sort rank	35,873	14.81	514.96	470.70	5809.35	254.18
rating (1-5)	14,909	1.00	4.47	4.65	5.00	0.69
impressions	35,873	1.10	10.46	10.43	18.35	1.79
clicks	35,873	0.00	6.35	6.21	14.76	2.16
add-to-cart	35,873	0.00	3.83	3.71	12.79	2.57
quantity	35,873	0.00	1.99	1.39	11.42	2.20
revenue	35,873	0.00	5.29	6.90	16.12	4.56

Notes. This table reports summary statistics among dining chair products: prices, historical rank, rating, historical scaled impressions, clicks, add-to-cart rates, quantity sold, and revenue.

Figure E8: Number of dining chairs within a product



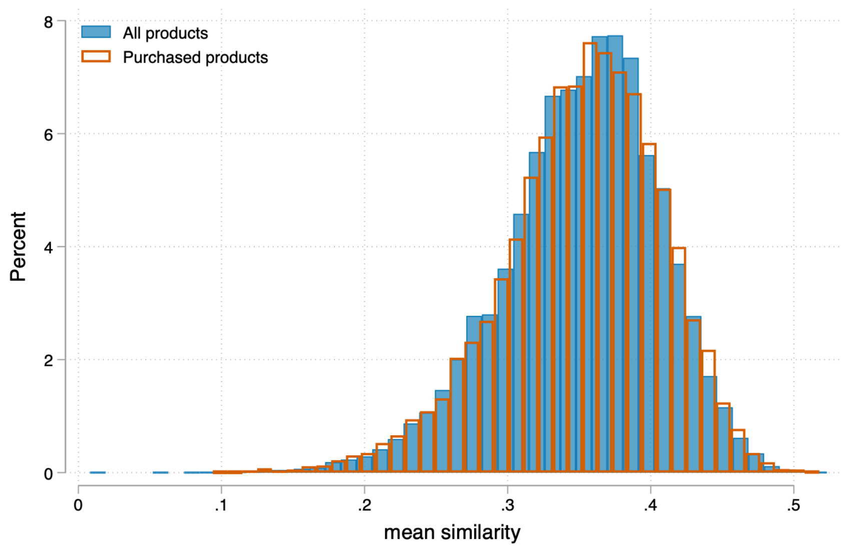
Notes: This figures shows the percent 1-, 2- 4, 6-chair sets among all chairs available on the website. The figure highlights that most products have either one or two chairs in the set, which is why there are opportunities for the repeat purchases of the products because typically consumers would need multiple chairs.

Table E13: Prices, wholesale costs and profits by seller mass marketness (return to p. 45)

	(1) Log price ($)	(2) Log wholesale cost ($)	(3) Log markup (%)
Mean cosine similarity	-0.0477***	-0.0499***	0.0240***
	(0.00810)	(0.00960)	(0.00491)
Observations	53,801	53,801	53,801

Notes: This table shows the results of estimation: $y_j = \alpha + \beta cosine_j + \varepsilon_j$. We investigate the correlation between the mean cosine similarity of a product and the retail price, wholesale cost, and the markup. We hide the intercepts for data sensitivity reasons in this table. The table shows that products that are more mass market (bestseller - have higher mean cosine similarity to other products) have lower prices, lower wholesale costs and higher markups for the platform. This is why profits are less impacted under privacy restrictions.

Figure E9: Purchased products versus existing products



Notes: This figure shows the distribution of products' mean cosine similarities among all products and those products that are purchased at least once. The distributions are close showing that consumers have heterogeneous preferences and buy all types of products. Products with high mean cosine similarity are more mass market, while low mean cosine similarity products are more niche.

Table E14: Differences in search and purchase behavior by authentication decision (return to p. 36)

	Logged in (1)	Cookie-recognized (2)	Difference (3)	p-value (4)
Clicks	4.25	2.70	1.55*** (111.52)	0.00
Clicked Prices ($)	173.77	177.03	-3.26*** (-5.96)	0.00
Add-to-carts	1.12	0.94	0.18*** (32.08)	0.00
Add-to-cart Prices ($)	163.79	165.21	-1.42 (-1.40)	0.16
Orders	0.55	0.52	0.02*** (8.05)	0.00
Purchased Prices ($)	152.82	145.51	7.32* (2.23)	0.03

Notes: This table shows the results of the t-test between the search and purchase behavior of people who log in versus those who are cookie recognized. All the differences are significant and indicate that consumers who login click and order more and are less price elastic.

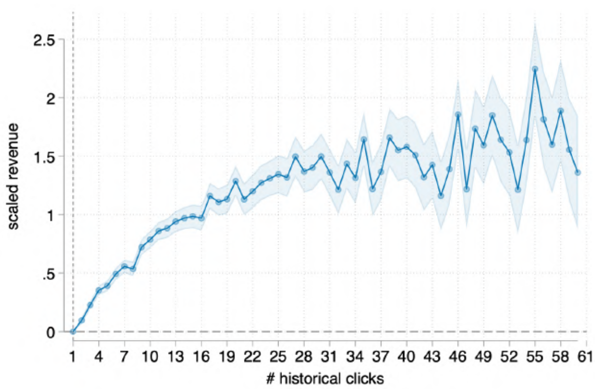
Table E15: Differences in search and purchase behavior by traffic source (return to p. 36)

	Display Advertising People (1)	Other (2)	Difference (3)	p-value (4)
Clicks	2.70	1.75	0.95*** (56.10)	0.00
Clicked Prices ($)	109.59	85.94	23.65*** (53.77)	0.00
Add-to-carts	0.24	0.16	0.08*** (35.20)	0.00
Add-to-cart Prices ($)	95.05	81.29	13.76*** (17.72)	0.00
Orders	0.02	0.02	-0.00*** (-3.49)	0.00
Purchased Prices ($)	83.17	76.56	6.61*** (3.44)	0.00

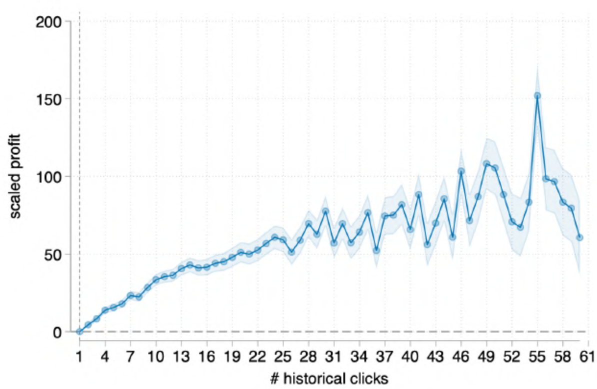
Notes: This table shows the results of the t-test between the search and purchase outcomes of consumers who arrive from display advertising and those who arrive from other sources. All the differences are significant indicating that consumers who arrive from display advertising are significantly less price elastic and potentially have lower search costs.

Figure E10: Diminishing returns of the data

(A) Scaled revenue

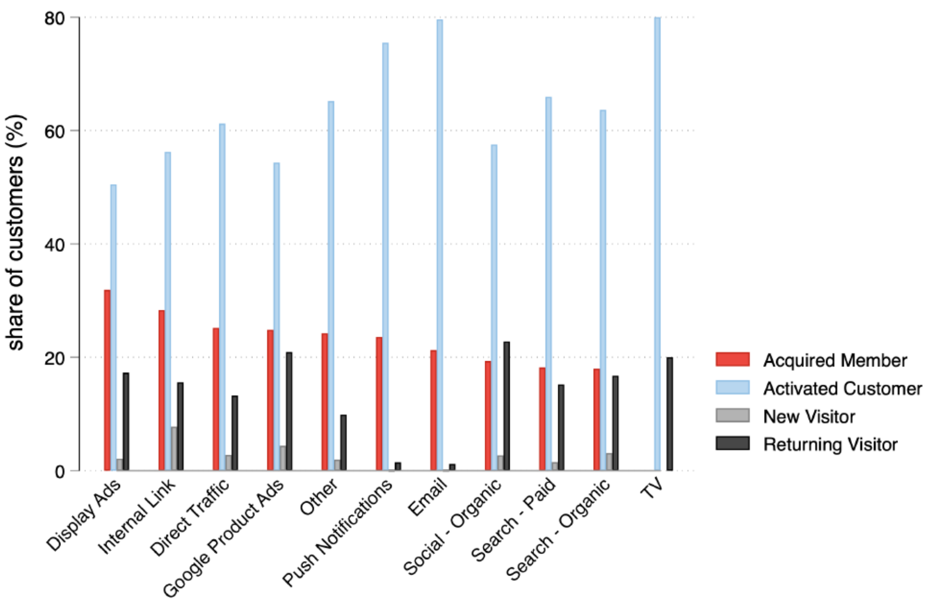


(B) Scaled profit



Notes. This figure shows the scaled revenue (A) and scaled profit (B) as a function of historical (pre-experiment) clicks made by consumers who participated in the ranking experiment. There is a positive and significant relationship between the amount of browsing history and scaled revenues and profits. However, there is diminishing returns of data because marginal data point brings less and less revenue/profit and both lines become flatter. Back-of-the envelope calculation would say that every click would bring approximately $2.69 before the economies of scale kick in and it drops to $0.23 after the maximum is reached.

Figure E11: Types of consumers arriving through different channels



Notes: This figures shows the distribution of consumer types arriving from different channels. Most channels have similar shares of consumer types except for TV, Email and Push Notifications. This is intuitive because it is mostly most active 'Activated Customers' that use these channels. The difference between Acquired Member and Activated Customer is the level of their activity: the latter is more active on the website.

Table E16: Correlations b/w purchased product and search products' prices by personalization condition

	Price of the Purchased Product	
	(1) Personalized Product	(2) Non-personalized Product
Avg. Searched Price	0.711***	0.490***
Min. Searched Price	0.241***	0.009**
Max. Searched Price	0.190***	0.002
Observations	2,123	64,615

Notes. This table shows the correlations between the purchased products' and searched products' prices by personalization status of the product. Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table E17: Main consumer metrics in the sample

	Total	**Clicked**	**Added-to-cart**	**Purchased**
Absolute	635,267	581,423	166,902	43,110
Percent terms	100.0%	91.5%	26.3%	6.8%

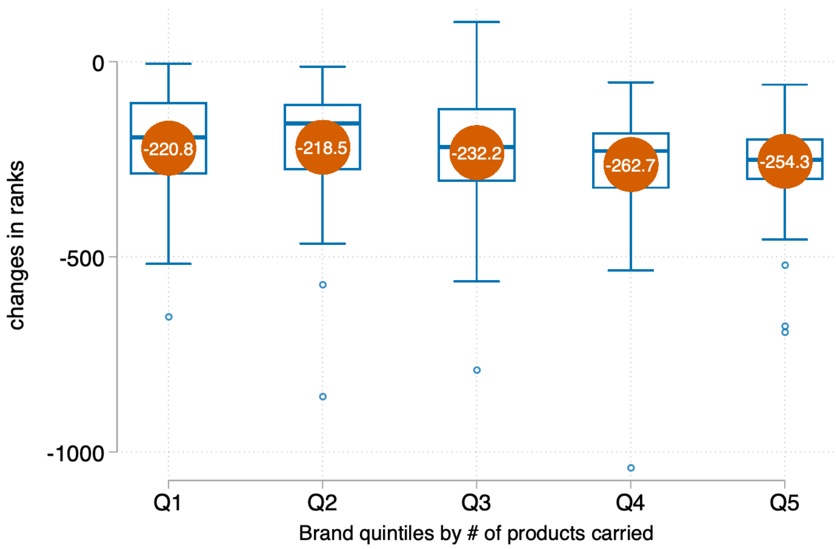
Notes. This table shows the absolute and relative (to total # of consumers) number of consumers who clicked, added to cart and purchased.

TABLE E18: Heterogeneity in dining chair prices

Main Material	Mean Price ($)	Median Price ($)	St.Dev. Price ($)	% of Products
Wood	524.26	389.99	413.96	33.33
Wicker / Rattan	523.81	449.99	354.77	1.86
Upholstered	499.86	379.99	398.71	54.89
Metal	300.24	222.00	248.16	2.92
Plastic / Acrylic	272.75	217.99	210.96	7.00

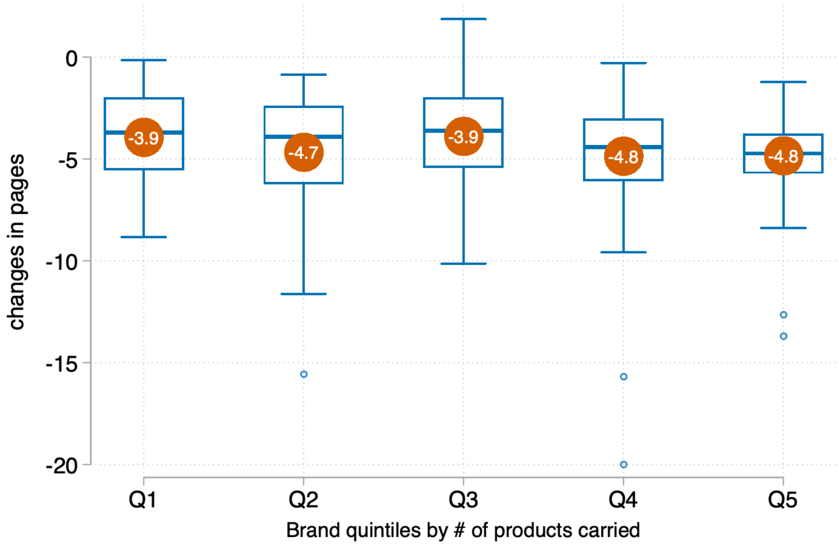
Notes. This table shows the heterogeneity dining chair prices across different chair materials.

FIGURE E12: Change in ranks by brand sizes



Notes. This figure illustrates the relationship between brand size and the changes in the rankings when product is shown as part of the personalized rankings. Brand size is defined as the total number of products offered by a brand. The $x-$axis represents the quintiles of brands categorized by their product count. The y-axis represents the differences in product ranks in personalized impressions versus non-personalized impressions. The numbers inside orange circles indicate the median rank change for each quintile.

FIGURE E13: Change in pages by brand sizes



Notes. This figure illustrates the relationship between brand size and the changes in the pages when product is shown as part of the personalized rankings. Brand size is defined as the total number of products offered by a brand. The $x-$axis represents the quintiles of brands categorized by their product count. The y-axis represents the differences in product ranking pages in personalized impressions versus non-personalized impressions. The numbers inside orange circles indicate the median page change for each quintile.

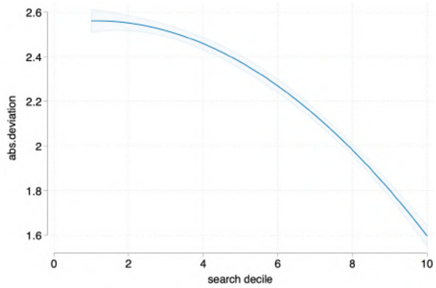
TABLE E19: Jumps in the search process (return to p. 26)

	(1) $\Delta prices$	(2) $\Delta cosine\ similarity$	(3) $\Delta width$	(4) $\Delta depth$
Standardized sales	6.588***	0.866***	-0.040***	-0.127***
	(0.560)	(0.019)	(0.012)	(0.022)
Purchased	-51.432***	1.202***	-0.617***	-0.769***
	(3.163)	(0.149)	(0.028)	(0.044)
#Ratings	-13.402***	-0.470***	-0.013	0.067***
	(0.609)	(0.022)	(0.016)	(0.025)
St.Dev. Ratings	-9.507***	0.085***	-0.021***	0.046***
	(0.816)	(0.028)	(0.007)	(0.011)
Rating	-7.697***	-0.021	-0.103***	-0.115***
	(1.246)	(0.044)	(0.010)	(0.016)
Intercept	264.160***	18.438***	2.429***	2.847***
	(6.107)	(0.215)	(0.051)	(0.079)
Customer FE	Yes	Yes	Yes	Yes
N	4,798,732	2,917,274	820,034	826,132

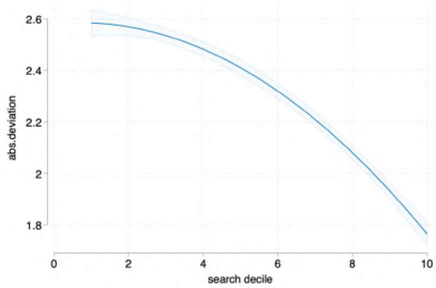
Notes: We calculate the step size in quantitative attributes (e.g. price, cosine similarity with the previously clicked product, width and depth of the chair): the difference between the attribute value of the previously searched product and that of next product. $|\Delta y_{ijt}| = |y_{ijt} - y_{ijt-1}| = \alpha + \beta_1 sales_{ijt-1} + \beta_2 purchased_{ijt-1} + \beta_3 \#ratings_{ijt-1} + \beta_4 st.dev.ratings_{ijt-1} + \beta_5 rating_{ijt-1} + \varepsilon_{ijt-1}$, where i is the consumer, jt is the product searched at search instance t, y-variables are absolute step size between attribute value of the previously searched product and the current one, $sales_{ijt} - 1$ is the standardized sales of product searched at search instance $t-1$ (previously searched one), $purchased_{ijt} - 1$ is the indicator for whether consumer i purchased product j searched at $t-1$, $\#ratings_{ijt}$ - number of ratings product j carries, $st.dev.ratings_{ijt}$ - standard deviation in the ratings of the previously searched product and rating is the absolute rating of the product that we control for.

Figure E14: Convergence patterns in consumer search

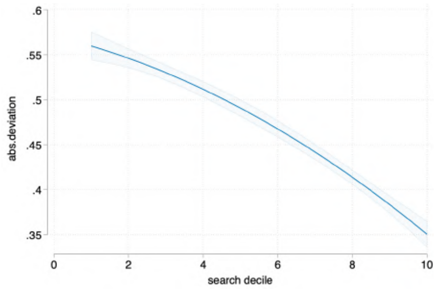
(A) chair height



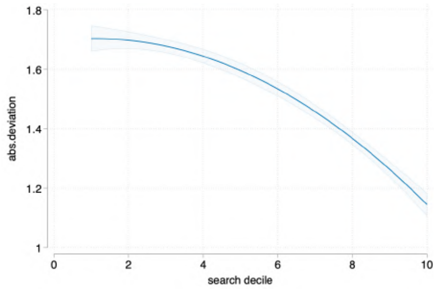
(B) chair depth



(C) number of chairs in a set

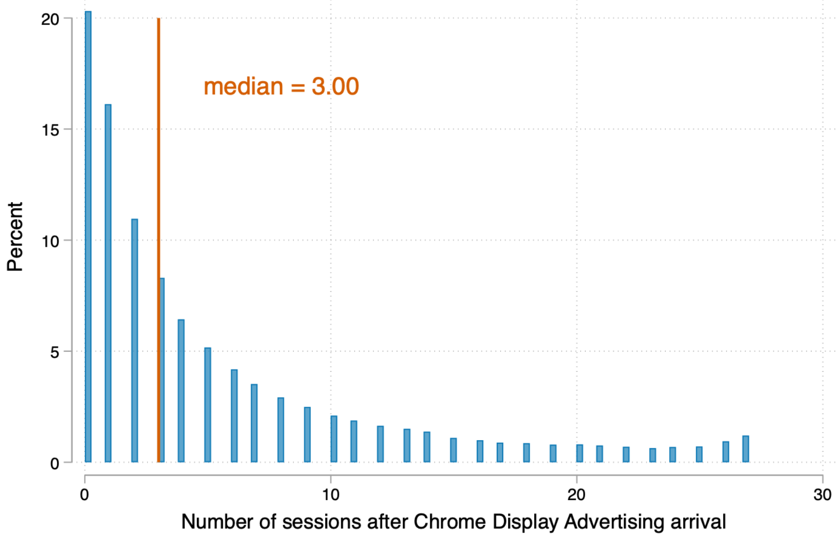


(D) seat width



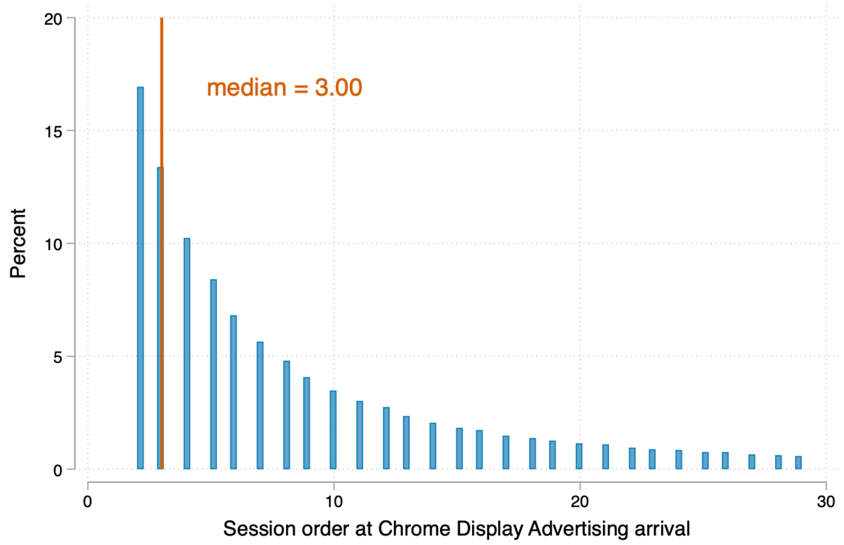
Notes. This figure shows that consumers exhibit spatial learning behavior documented in Bronnenberg, Kim and Mela (2016). The x-axis shows consumer's search decile (progression), and y-axis shows the absolute deviation of the searched product attribute from the chosen (purchased) product's attribute.

Figure E15: Sessions after Chrome Display Advertising arrival



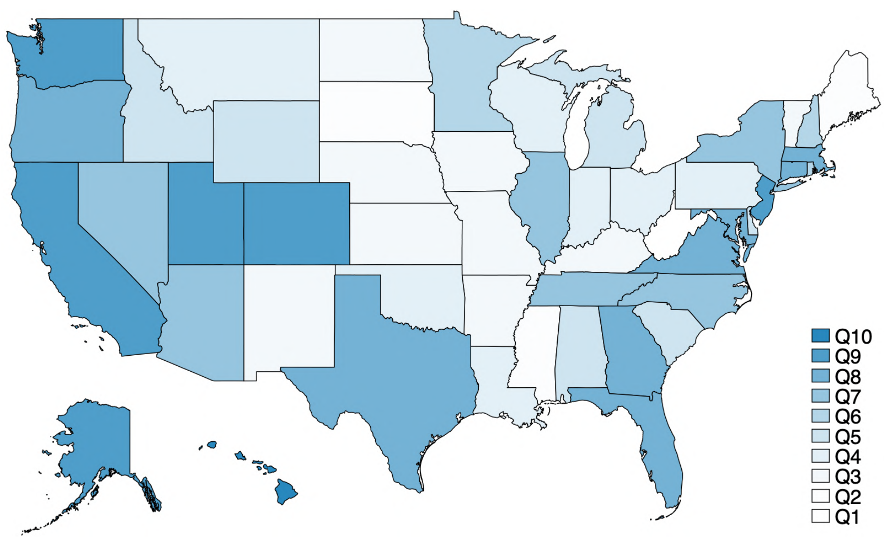
Notes. This figure illustrates the number of sessions consumers search for after arriving from Chrome Display Advertising channel. Median consumer keeps searching for additional three sessions after arriving from the advertising channel.

Figure E16: Session order at Chrome Display Advertising arrival



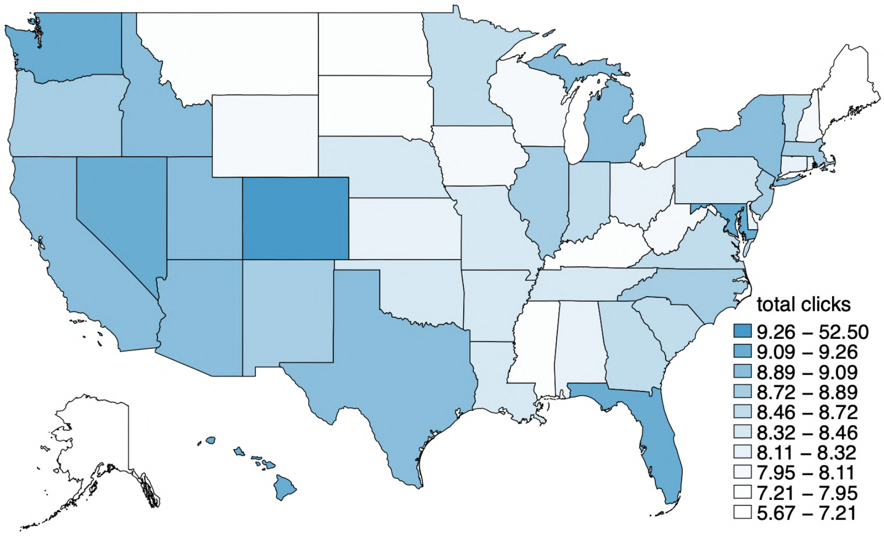
Notes. This figure illustrates the session order at which consumers arrive from Chrome display advertising. Median consumer who arrives from Chrome display advertising, does so at session number three.

Figure E17: Price inelastic consumers by state



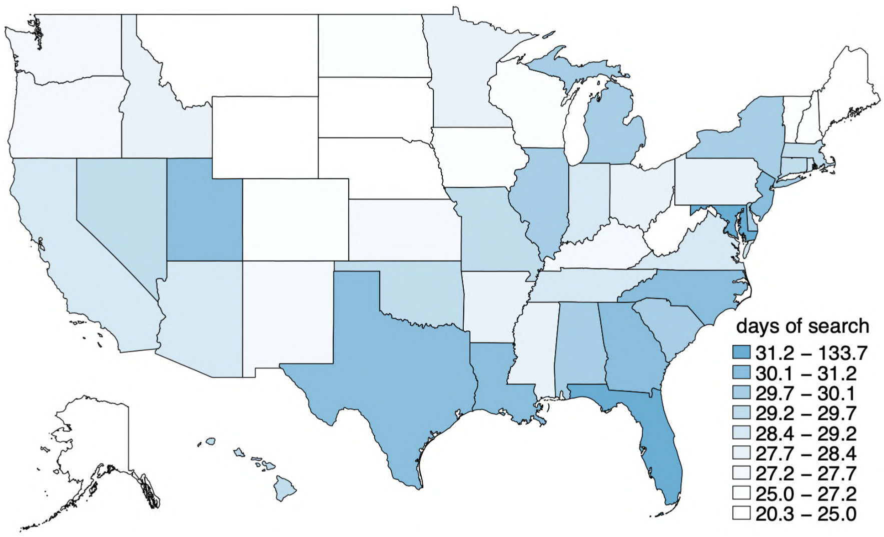
Notes. This figure illustrates the implied price elasticities based on consumers' previous clicking and purchase behavior.

FIGURE E18: Total clicks by state



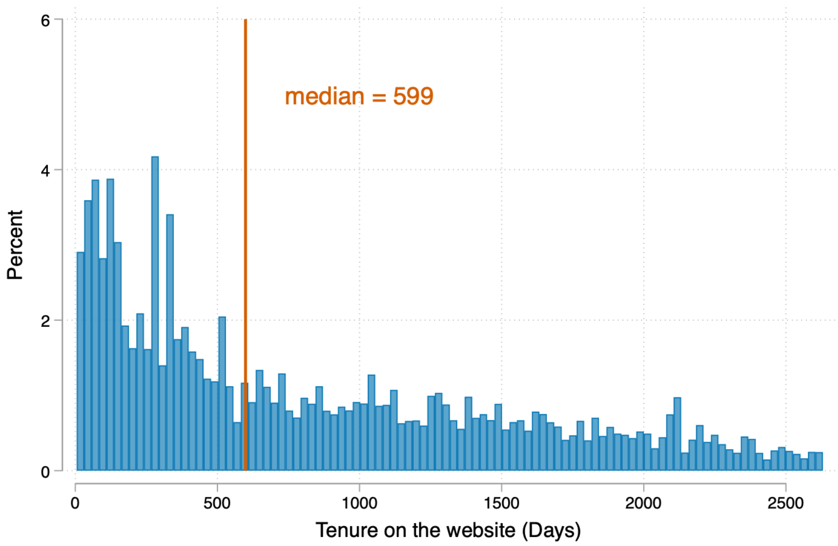
Notes. This figure illustrates the number of total clicks based on consumers' previous clicking and purchase behavior.

FIGURE E19: Days of search by state



Notes. This figure illustrates the days of search based on consumers' previous clicking and purchase behavior.

FIGURE E20: Tenure on the website



Notes. This figure illustrates the distribution of days products existed on the website.

F. Belief updating

F.1. Prior mean updating

Table F20 below gives an intuition behind belief updating process that occurs as a result of Gaussian process specification. Suppose consumer clicked on product j and observed true utility u_j. Given the observed utility, the model allows the consumer to update her beliefs about the remaining products (Equation 29). Consumer first calculates how much the observed utility from product j deviates from her prior belief $(u_j - \mu(X_j))$ and then updates posterior beliefs about other products according to the weights implied by Gaussian distribution.

$$\underbrace{\mu'(X_{-j|j})}_{posterior\ means} = \underbrace{\mu(X_{-j})}_{prior\ means} + \underbrace{\frac{\kappa(X_{-j}, X_j)}{\kappa(X_j, X_j) + \sigma_\xi^2 + \sigma_\varepsilon^2}}_{weights} \underbrace{(u_j - \mu(X_j))}_{deviation\ of\ observed\ utility\ from\ the\ prior} \tag{29}$$

Table F20: Posterior updates of mean utilities

Case	**Updating**	**Explanation**
$u_j = \mu(X_j)$	$\mu'(X_{-j\|j}) = \mu(X_{-j})$	observed utility is equal to the prior, and, consequently, posteriors are not updated.
$u_j > \mu(X_j)$ and $\kappa(X_{-j}, X_j)$ high	$\mu'(X_{-j\|j}) > \mu(X_{-j})$	if products X_{-j} and X_j are highly related, and consumers got positive utility signal, then they will update positively the posterior mean on X_{-j}.
$u_j < \mu(X_j)$ and $\kappa(X_{-j}, X_j)$ high	$\mu'(X_{-j\|j}) < \mu(X_{-j})$	if products X_{-j} and X_j are highly related, and consumers got negative utility signal from X_{-j}, then they will think that the products similar to the observed one have low utility and will decrease posterior mean utility beliefs.
$u_j > \mu(X_j)$ and $\kappa(X_{-j}, X_j)$ low	$\mu'(X_{-j\|j}) \leq \mu(X_{-j})$	if products X_{-j} and X_j are not quite related or negatively related, then getting positive signal from X_j either does not affect the beliefs about the dissimilar products or may decrease them.
$u_j < \mu(X_j)$ and $\kappa(X_{-j}, X_j)$ low	$\mu'(X_{-j\|j}) \geq \mu(X_{-j})$	if products X_{-j} and X_j are not quite related or negatively related, then getting negative signal from X_j either does not affect the beliefs about the dissimilar products or may increase them.

F.2. Prior covariance updating

In the model, consumer updates both the posterior mean beliefs about product qualities, as well as the strength of the relationship between products. The updating process follows Equation 30. Every click on a product may reduce uncertainty about product qualities, which is reflected in the posterior covariance matrix. Table F21 gives an intuition for different updating scenarios.

$$\underbrace{\kappa'_{-j|j}}_{posterior\ covariance\ matrix} = \underbrace{\kappa_{-j|j}}_{prior\ covariance\ matrix} - \underbrace{\kappa_{-j,j}\ \overbrace{\left(\kappa_{j,j} + \sigma_{\xi}^{2} + \sigma_{\varepsilon}^{2}\right)}^{\substack{\text{total uncertainty} \\ \text{about product } j \text{ payoff}}}{}^{-1}\kappa_{-j,j}^{T}}_{reduction\ in\ uncertainty} \tag{30}$$

Table F21: Posterior updates of covariance matrix

Case	**Updating**	**Explanation**
$\kappa_{-j,j}$ are zeros or low	$\kappa'_{-j\|j} = \kappa_{-j\|j}$	if observed product X_j is unrelated to product X_{-j}, i.e. $\kappa_{-j,j} = 0$ then corresponding posterior covariances of product $-j$ with other products do not change.
$\kappa_{-j,j}$ are high	$\kappa'_{-j\|j} < \kappa_{-j\|j}$	if observed product X_j is highly related (similar) to product X_{-j}, i.e. $\kappa_{-j,j}$ is high, then corresponding posterior covariances of product $-j$ with other products will decrease. The intuition is that observing a product decreases uncertainty about the payoffs of other products similar to the observed one.

G. Lower-dimensional product representation: embeddings

We use Deep Learning, namely Convolutional Neural Networks (CNNs), to construct image similarity. The goal of the model is to represent each product as a numerical vector with 512×1 dimensionality. To train the model, we need product images and ground truth training data that has true labels of product similarity. For training, we use more than 1.5 million images generously provided by Wayfair. Each product has multiple pictures associated with it (Figure G21) and this provides ground truth labels. That is, if the images correspond to the same product, then in the pairwise comparison the label is one, and zero otherwise. The goal of the trained Convolutional Neural Network is to project the images to a lower dimensional space in a way that would place images of the same product close together, while all the dissimilar images will be located at a distance.

Figure G21: Different images of the same product

Figure G22 illustrates this: the goal is to learn a lower-dimensional representation of the image such that the image in Panel (A) and the matched image of the same product in Panel (B) map to the same point in the embedding space, while image in Panel (C) which corresponds to a completely different product is separated by at least a margin m in the embedding space.

Formally, we use the Convolutional neural networks (CNNs) (see Hadsell, Chopra and LeCun (2006) and Bell and Bala (2015) for more details). A convolutional neural network is a function f that maps each image I into an embedding position x, given parameters θ: $f_\theta : I \to x$. The goal of CNN is to solve for the parameter vector θ such that the produced embeddings x are such that similar images are placed nearby, while dissimilar images are more distant.

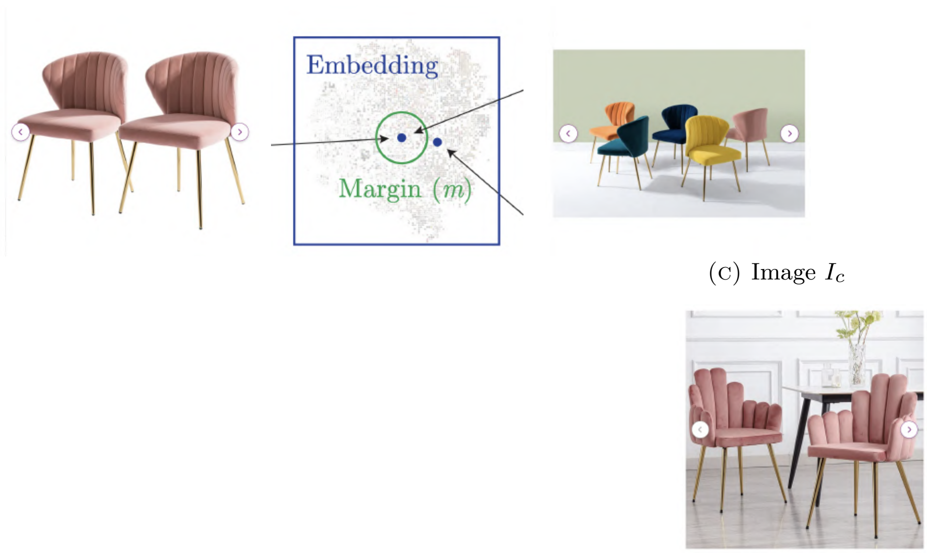
Suppose there are three images, I_a, I_b, I_c. Images I_a and I_b are two different images of the same object (Panels (A) and (B) of Figure G22), while image I_b and I_c correspond to two completely different objects (Panels (B) and (C) of Figure G22). Via CNN f_θ, the images are mapped into the embeddings x_a, x_b, x_c. If the model is trained well and produces 'good' embeddings, then the correct (positive) pair x_a and x_b should be close together, while the incorrect (negative) pair x_b and x_c is further apart.

The objective function for this kind of task is called contrastive loss and was proposed by Vedaldi, Jia, Shelhamer, Donahue, Karayev, Long and Darrell (2014). We minimize the loss function defined

FIGURE G22: Idea behind embeddings

(A) Image I_a (B) Image I_b

(C) Image I_c



Notes: The figure illustrates the idea behind image embeddings. Products in Panel (A) and (B) are very similar, and the model's task is to put these products close in the embedding space. Meanwhile, product in Panel (C) should be more distant from (A) and (B) in the embedding space.

as

$$L(\theta) = \sum_{x_a, x_b} \underbrace{L_b(x_a, x_b)}_{\substack{\textit{Penalty if similar images} \\ \textit{are far away}}} + \sum_{x_b, x_c} \underbrace{L_c(x_b, x_c)}_{\substack{\textit{Penalty if dissimilar images} \\ \textit{are nearby}}} \tag{31}$$

where $L_b(x_a, x_b) = ||x_a - x_b||_2^2$ and $L_c(x_b, x_c) = max\{0, m^2 - ||x_b - x_c||_2^2\}$. The objective (loss) function consists of two parts: $L_b(x_a, x_b)$ penalizes a positive pair (x_a, x_b) if the embeddings of the images of the same product are too far apart, and $L_c(x_b, x_c)$ penalizes a negative pair (x_b, x_c) if the images of two different items are closer than the margin m.

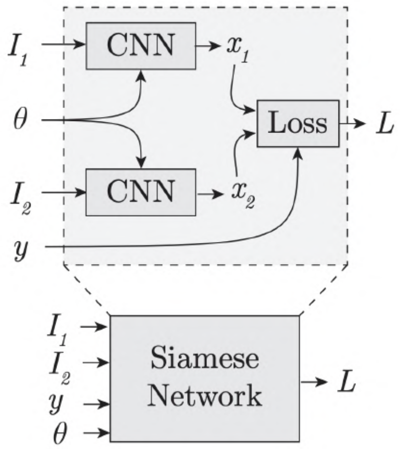
We minimize the objective function in Equation 31 with respect to parameters θ using stochastic gradient descent with momentum (Krizhevsky, Sutskever and Hinton (2017), Bottou, Curtis and Nocedal (2018)):

$$v^{(t+1)} \leftarrow \mu \cdot v^{(t)} - \alpha \cdot \frac{\partial L}{\partial \theta}(\theta^{(t)}) \tag{32}$$

$$\theta^{(t+1)} \leftarrow \theta^{(t)} + v^{(t+1)} \tag{33}$$

where v is the momentum sequence, $\mu \in [0, 1)$ is the momentum and $\alpha \in [0, \infty)$ is the learning rate. For computational reasons, to efficiently compute the loss function L and the gradient $\frac{\partial L}{\partial \theta}$, we follow Hadsell, Chopra and LeCun (2006). Their approach is to construct a siamese network, which is the two copies of the CNN that share the same parameters θ. The network takes as an input two images I_1, I_2, θ and the indicator variable for whether the images are a positive pair $(y = 1)$ or a negative pair $(y = 0)$, and outputs the loss function L (Figure G23).

Figure G23: Siamese network



Notes. This figure illustrates the architecture of the Deep Neural Network that we used to create image embeddings. the Source: Bell and Bala (2015).

For training, we used more than 1.5 million images of approximately 320,000 products. We took all images available for fifty[54] thousand dining chair products and supplemented them with products from other furniture categories, e.g. sofas, office chairs, ottomans, beds. Using products from other categories ensures better training of the image recognition model.

Different images of the same product constitute a positive pair, and the model should deliver embeddings (vectors) that are similar in some distance metric, e.g. cosine similarity. For negative examples, for each product, we sampled other products from the same category and other categories.

The output of the model is the 512×1 vector that characterizes each product. Next, we can calculate cosine similarities between different products' vectors to identify how similar or dissimilar products are. Figures G24 and G25 show some examples of products that have high cosine similarity (> 0.9) and the examples of dissimilar products (< 0.1), where we calculated cosine similarities using the model output. Figure G26 a condensed output of the model in two-dimensional space, where we used UMAP to project 512×1 images into two-dimensional space.

Figure G24: Examples of products with high cosine similarity

(a) Cosine similarity: 0.9294

(b) Cosine similarity: 1.0000

Notes. This figure shows the examples of the products with high cosine similarities. Typically, products that have cosine similarity higher than 0.9 are visually very similar.

[54]At each point in time, there are approximately thirty thousand dining chairs on the website, but historically there were alltogether fifty thousand chairs and we use all the available pictures.

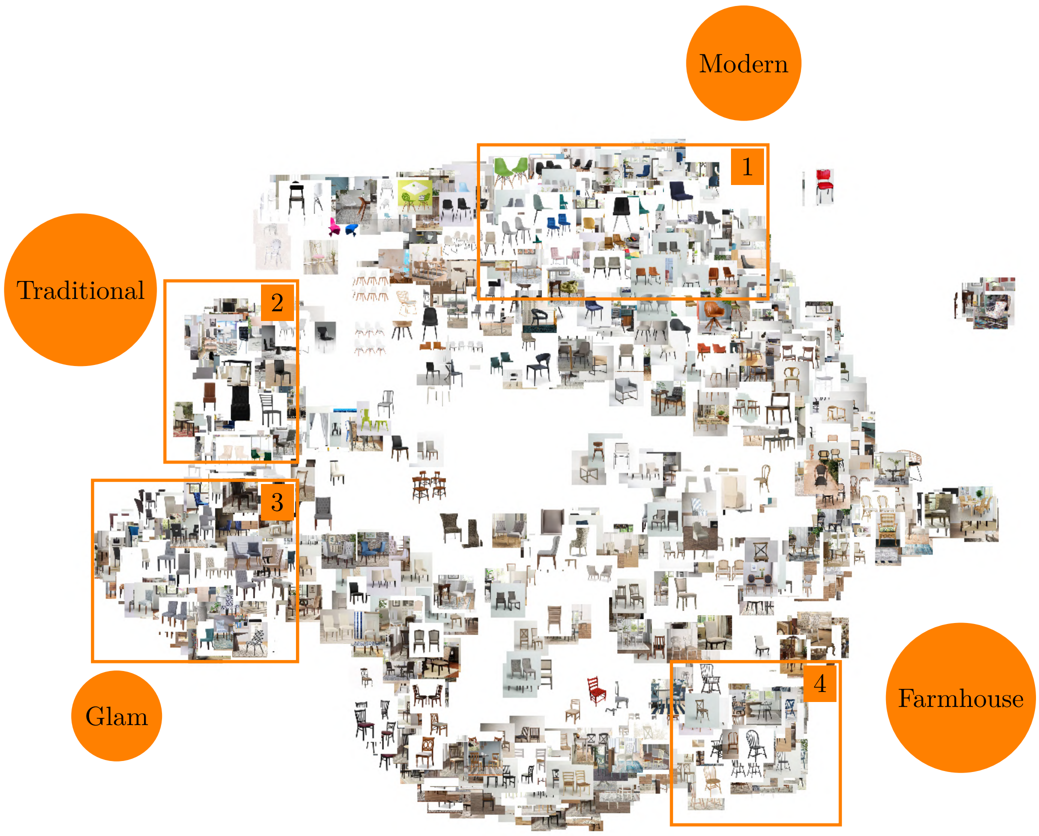
Figure G25: Examples of products with low cosine similarity

(A) Cosine similarity: -0.0001

(B) Cosine similarity: 0.0074

Notes. This figure shows the examples of the products with low cosine similarities. Typically, products that have cosine similarity lower than 0.5 are visually very different.

Figure G26: Output of embedding images in two-dimensional space (return to p. 37)

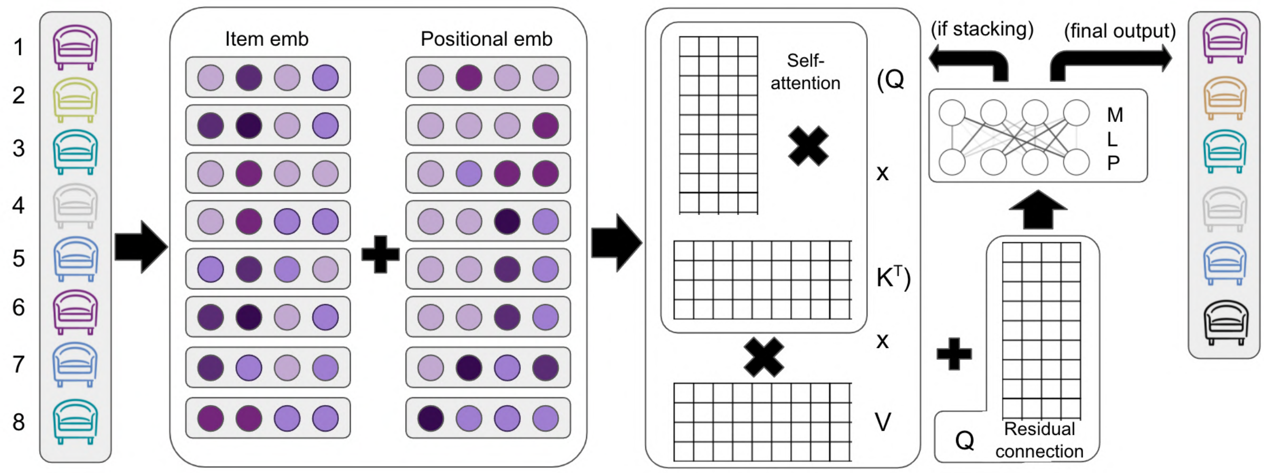


Notes: This figure shows a partial output of the image embedding model. The actual image emebdding is 512×1 vector. We used UMAP (Uniform Manifold Approximation and Projection) to reduce the dimensionality to get two-dimensional vectors.

H. Personalized rankings algorithm

On the ranking pages, Wayfair uses a sequential recommendation model to learn most recent consumer preferences and to provide recommendations that better match consumer tastes. The input to the model is simple: it is a sequence of items that a consumer has browsed previously within a category. During the sample period, the algorithm leveraged only within-category browsing and trained separate models **for each product category**. Model architecture is depicted on Figure I29. The input to the algorithm is an ordered sequence of consumer searches. First, the model learns similarity between different products and represents each product in a numerical vector form, i.e. item embedding. Similarly, the information on the order of searches is stored in the positional embeddings. Next, combined lower-dimensional representation of consumer searches stored in the item and positional embeddings is passed into subsequent layers of the model. There is a self-attention layer that learns the relationship between different items and decides which lower-dimensional attributes to store in the item embeddings. Next, the resulting summed embeddings are passed through the fully-connected layer with a sigmoidal (logistic) activation and binary cross-entropy loss. The final output of the model is the list of scores for all items in the training set. These scores are then used to provide the top n recommendations. The recommended products can include both previously-unseen and previously-seen products. See Mei, Zuber and Khazaeni (2022) for more details.

Figure H27: Architecture of the personalization algorithm



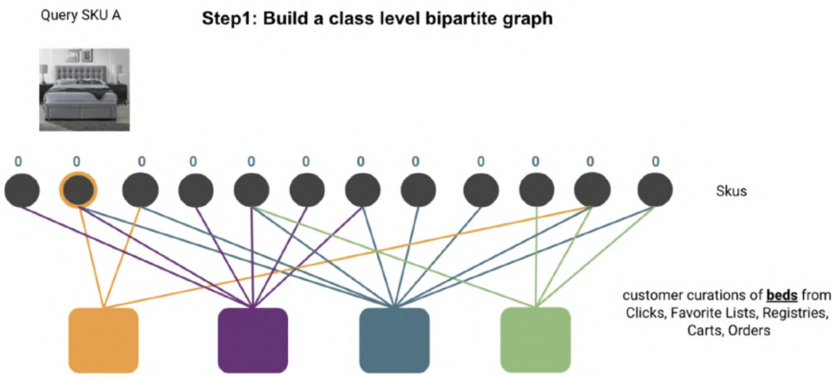
Notes: This figure shows the architecture of the personalization algorithm used by Wayfair. The input to the algorithm is an ordered sequence of consumer's browsing history, e.g. consumer browsed for eight chairs. This sequence of searches is encoded into item and positional embeddings and the result if passed through the multi-level perceptron layer. The final output of the model is a list of scores for all items in the training set. Next, products are ranked according to the scores to provide top n recommendations. Source: Mei, Zuber and Khazaeni (2022).

I. Recommendations on the product pages

On the product pages, consumers will see additional product recommendations. These recommendation widget is called compare similar items. It consists of five products, the first product is the anchor product itself, and there are four additional products. On Wayfair, consumers can click on a product, add it to cart, add it to favorite board and finally purchase.

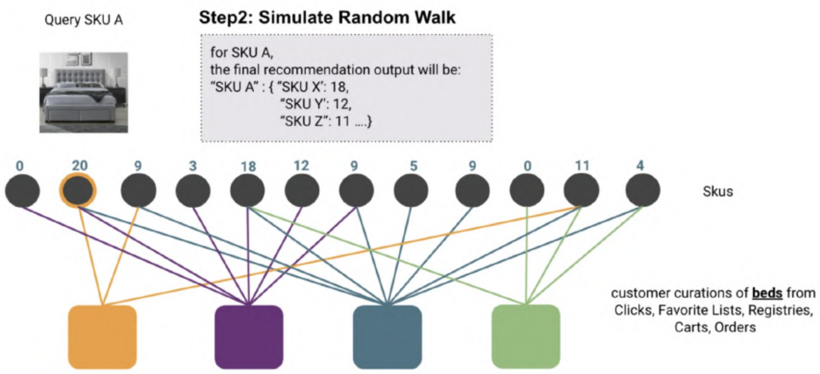
To create these recommendations, Wayfair uses data on how consumers co-purchase, co-add-to-cart and co-click the products. First, they create a bipartite random graph that calculates the visit counts to every product, using consumer clicks and add-to-carts. Next, the algorithm simulates a random walk to predict which products consumers are more likely to click on.

Figure I28: Building Bipartite graph algorithm



Notes: The platform builts the bipartite graph at the category level. The random walk starts by initializing the visit counts for every SKU to zero.

Figure I29: Simulation of random walk



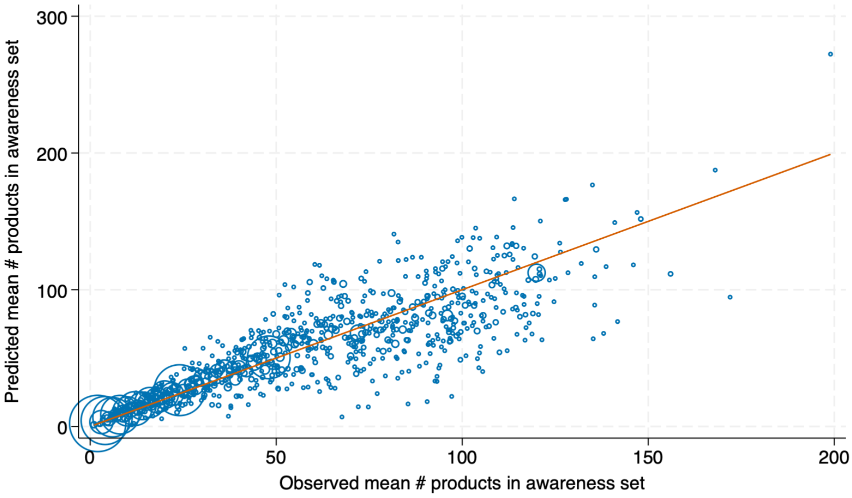
Notes: During the random walk process, we record the visit count for all SKUs in the graph. The top K SKUs with the highest visit count will become the final recommendation.

J. Model fit and validation

This section shows the results of the model fit analysis. First, we test the model fit by comparing the observed and predicted awareness set sizes, i.e. the number of products consumers *viewed.* Figure J30 shows the observed awareness set size on the x-axis and the predicted measure on the y-axis. The size of the circle indicates the number of people with a given awareness set size. The model predicts the awareness set particularly well, which ensures some credibility in the model.

We chose awareness set size as the model fit measure because it captures one of the main aspects that we model: viewing. If the model predicts the awareness sets well that means it captures search behavior sufficiently well.

Figure J30: Model fit: awareness set size (return to p. 38)



Notes: This figure plots the simulated and observed data moment: size of the consumers' awareness sets. The size of the dot is proportional to the number of consumers in each group. The orange line is the the 45-degree line.

Next, we check the model fit on other consumer-level and seller-level metrics. Table J22 shows that the model can predict the number of clicks, prices searched and the ranking page position of the purchased product, and the share of different sellers well.

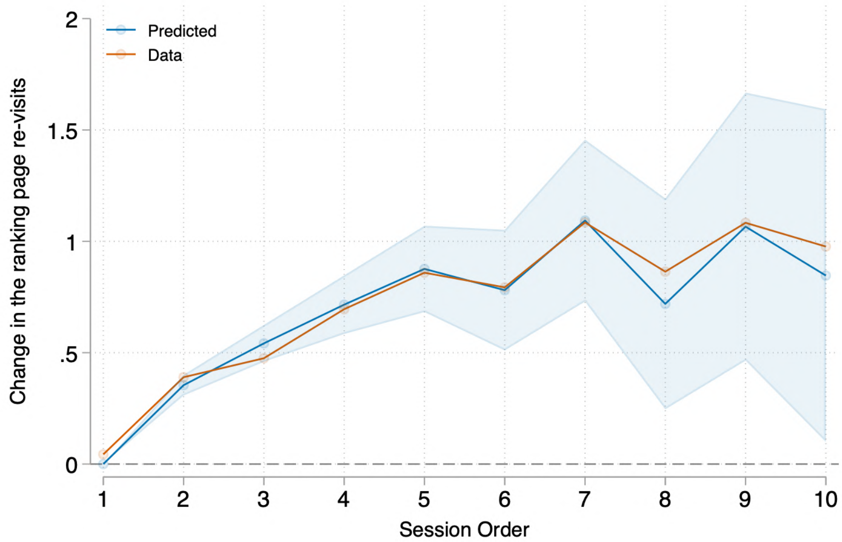
Table J22: Model fit: consumer outcomes

	Data	Predicted
Consumers		
# clicks	3.42	3.17
awareness set size	59.46	60.32
price searched ($)	$234	$239
position purchased	49	47
Sellers		
share of 10th percentile sellers	100%[1]	95%
share of 90th percentile sellers	100%[1]	97%

[1] Share normalized to 100% for data sensitivity reasons.

Finally, we validate the model using the data from Chrome privacy policy that was introduced in 2020 (Appendix A). The model predicts the number of times consumer leaves the product page and re-visits the ranking page. We chose this metric because it captures the main components of the model, i.e., viewing products and clicking and then navigating back to the ranking page. The orange line on Figure J31 shows the increase in the number of re-visits of ranking page during the Chrome event. The predicted and observed data moments match well, further validating the model.

FIGURE J31: Model fit: search behavior during Chrome SameSite updates (return to p. 38)



Notes: This figure plots the simulated and observed data moment: the increase in the ranking page re-visits during Chrome event. The red line corresponds to the observed data and blue line is the predicted data moment. The shaded area indicates the 95% confidence interval across simulation draws.

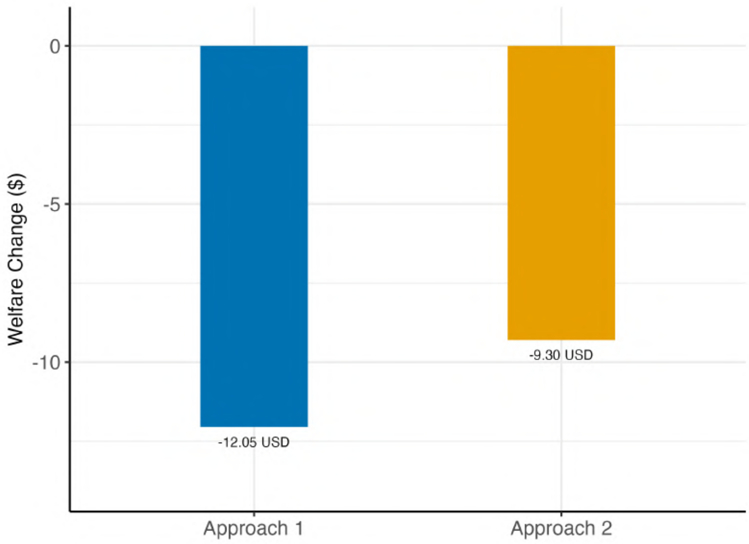
K. Counterfactual Results

Figures K32 and K33 show the comparison between two approaches to simulating Chrome restrictions. In Approach 1, we fragment the data and re-train the personalization algorithms using fragmented data. Next, we re-train platform's personalization algorithm using the fragmented data. This is the approach that we report in the main text.

In Approach 2, we assume that Chrome will indicate aggregate clustering group for the consumer interests, e.g., modern style. To simulate that, first, we segment consumers into clusters (styles) of interest based on their browsing history. We calculate the number of clicked products of a certain style among the viewed ones and classify consumer into a cluster that received highest click-through-rate. Next, we show personalized recommendations of a style of interest mixing small and large sellers' products in the rankings.

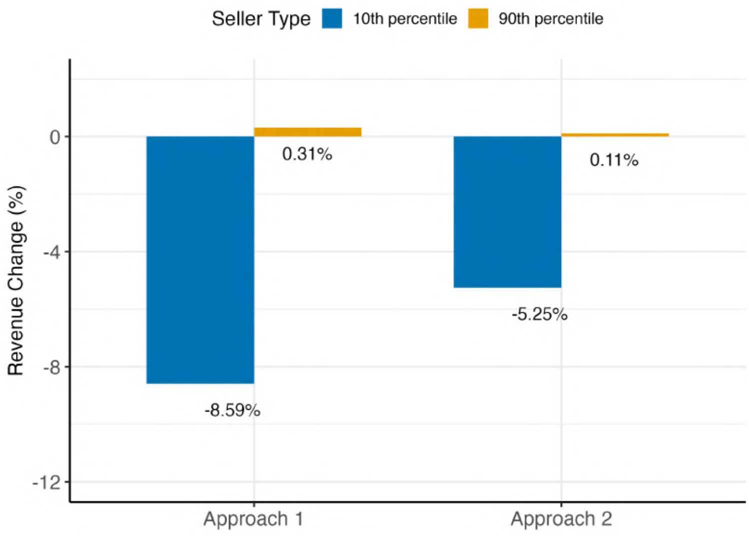
Both approaches give qualitatively similar results but we report the Approach 2 in the Appendix because it makes a strong assumption about the kind of information that is revealed to to the platforms. Moreover, we apply a simple heuristic approach and show a mix of small and large sellers' products. In reality, the platform might use a complete different algorithm given the identified clusters. Approach 1 is a scenario that is more realistic to happen and requires less heuristic choices compared to Approach 2.

Figure K32: Comparison of two approaches for Chrome counterfactuals: consumer welfare



Notes: This figure compares the counterfactual welfare measures between two approaches of simulating Chrome restrictions.

Figure K33: Comparison of two approaches for Chrome counterfactuals: seller revenue



Notes: This figure compares the counterfactual seller revenue measures between two approaches of simulating Chrome restrictions.

Table K23: Counterfactual results (return to p. 42)

	(1) 'Full' data vs Bestseller	(2) 7 day vs 'Full' data	(3) Chrome 2024 vs 'Full' data	(4) Algorithm vs 'Full' data
Consumer				
Δ welfare ($)	+25.3 vs +16.4 (Δ = $8.9)	-4.78	-6.98	-3.06
purchase probability	+12.09%	-17.52%	-24.65%	-23.45%
match value	+48.70%	-33.45%	-54.11%	-30.36%
scrolling cost savings	+38.37%	-43.24%	-32.53%	-35.42%
clicking cost savings	+0.83%	-5.79%	+11.28%	-10.77%
Seller				
revenue (10th percentile)	+10.23%	-5.64%	-8.59%	-2.34%
conversion prob (10th percentile)	+1.64%	-2.30%	-2.84%	+1.23%
revenue (90th percentile)	+0.24%	+0.28%	+0.31%	+0.12%
conversion prob (90th percentile)	+1.00%	+1.5%	+1.10%	+0.78%
Platform				
Δrevenue (%)	+2.19%	-1.15%	-1.94%	-0.82%
Δprofit (%)	+1.25%	-0.21%	-0.39%	-0.13%

Notes. This table reports the full **counterfactual simulation** results. Column (1) shows the changes in consumer, seller and platform outcomes under personalization using full data versus non-personalized rankings. Column (2) compares using Safari-style 7 day cookie reset policy with the full personalization case. Column (3) compares forthcoming Chrome 2024 policy to the full personalization. Column (4) compares the proposed algorithm with the full data scenario.