Behavioral advertising and consumer welfare: An empirical investigation

Preliminary Draft, June 2021 - Please do not cite
Please contact authors for updated version

Eduardo Abraham Schnadower Mustri, Idris Adjerid, and Alessandro Acquisti [1]

Abstract

We investigate the impact of behavioral advertising on consumer welfare in a within-subjects online experiment. While the vast majority of empirical work on the impact of online advertising focuses on click and conversion rates of behaviorally targeted ads, we propose a counterfactual approach, in which online consumers are presented with alternative offers: products associated with targeted ads they were served online, competing products, and random products. Participants are asked to compare these alternatives along a variety of metrics. Thus, we assess consumer welfare implications of behavioral advertising *comparatively*, in an ongoing online experiment that captures differences in participants' purchase intentions and other product characteristics which can affect consumer utility.

1. Introduction

The value that consumers derive from behavioral advertising has been more often posited than empirically demonstrated. Prior work has shown that behaviorally targeted ads tend to receive higher click-through rates than non targeted ones (Bleier & Eisenbeiss, 2015; Yan et al., 2009), suggesting that the former can reduce consumer search costs. Other than through that cost reduction, however, little is known about the manner and extent to which behaviorally targeted ads affect consumers' welfare. For instance: what is the relationship between products associated with targeted ads and other factors that also affect the utility a consumer derives from a product - such as its price, quality, or novelty? We present the preliminary results of an ongoing online experiment in which participants were presented with a) products associated with targeted display advertisements that had been served to them while they navigated, using their own devices, a variety of sites; b) competitor products, obtained from search results; and c) randomly selected products, from ads that had been served to *other* participants in the experiments and that therefore were unlikely to be closely related to the participant's interests. We compare purchase intentions, as well as participants' perceptions of price fairness, quality, product relevance, and novelty,

[1] Eduardo Abraham Schnadower Mustri: eschnado@andrew.cmu.edu; Idris Adjerid: iadjerid@vt.edu; Alessandro Acquisti: acquisti@andrew.cmu.edu..

across the advertised, competitor, and randomly selected products, to evaluate comparatively the consumer welfare implications of behaviorally targeted advertisements.

Online advertising has become an essential part of the Internet economy. It composes 54% of total advertising spending across all media, and it was expected to reach $332 billion worldwide in 2020 (Cramer-Flood, 2020). "Online targeted advertising" (OTA) allows ad exchanges to tailor ads to a consumer. Some of the targeting occurs by analyzing the user's past behavior while browsing online, and this type of targeting is known as Online Behavioral Advertising (OBA). This contrasts with traditional mass media advertising, in which each ad targets a large mass of consumers, which is expected to include a good number of individuals that are interested in the advertised product. A significant proportion of online display ads today are behaviorally targeted (Fisher, 2019; Samuel et al., 2021). The technologies used for OBA have raised significant privacy concerns (Chiasson et al., 2018), and there is an ongoing debate of whether these technologies indeed benefit consumers.

Industry representatives have argued that OBA is beneficial to consumers, as targeted ads may be more relevant to consumers' interests. OBA can also provide to consumers information they would not have obtained otherwise, or would have required significant effort to obtain, allowing them to discover relevant products and services, saving time and money (Dehling et al., 2019). As noted, empirical evidence shows that OBA increases click-through rates (Bleier & Eisenbeiss, 2015). This increase in click-through rates (CTR) can be used to support the argument that OBA is beneficial for consumers: if consumers are more likely to click (and then purchase) products advertised via OBA, it means that they derive some utility from the ads. Thus, increases in CTR have been linked to increases in consumer welfare (Jeziorski & Segal, 2015). In addition, since behaviorally targeted ads are typically paid by advertisers more than non-targeted ones, they are likely to enhance the ability of content providers to provide their services online for free (Schuman et al., 2014) - something that consumers value (the Digital Advertising Alliance found that eighty-five percent of individuals surveyed preferred a free, ad-supported Internet as opposed to paying for services: DAA, 2020).

On the other hand, critics have pointed out that consumers may receive disutility from OBA due to their perceived instrusivenes, which can raise privacy concerns or cause annoyance (such as getting an ad for something already purchased or already researched eliminated from purchase considerations: Dehling et al., 2019). As a matter of fact, as of 2019, up to 47% of US consumers were estimated to use ad blockers (McCue, 2019), suggesting that a significant proportion of online visitors are willing to dispense with the purported benefits of ads, be them targeted or not.

More importantly, from the perspective of this manuscript, the relationship that behavioral targeting has with quality, price, or novelty of product offers is not well understood. For instance: How does quality compare, on average, between products shown in targeted ads to others in the market? How often do these ads show new products and undiscovered brands to consumers? What about prices of products that advertisers targeted to a consumer, versus the price of a competitor the consumer may have found through searches?

The answers to these questions are not obvious, on theoretical grounds, because different (even opposite) dynamics concerning the relationship between sellers' motivation to target ads and consumer welfare are plausible. For instance, sellers may have incentives to advertise their most profitable products - which may, or may not, align with the most preferred by consumers (Hagiu & Jullien, 2011; Zhang & Katona, 2012; Acquisti, 2014). Similarly, smaller vendors with higher margins may find it more profitable to use behavioral targeting than larger vendors that sell in bulk and pass on cost savings to consumers (Chen & Stellaert, 2014) - but larger vendors may, in theory, have access to more consumer information that allows them to target individuals' preferences and needs more efficiently. In fact, depending on whether vendors are able to target by price and preference, and how differentiated consumers are across those levels, targeting may either have a positive or negative effect on consumer welfare (Marotta et al., 2021).

The answers to the above questions are not obvious on empirical grounds, either, because, to the best of our knowledge, empirical testing of those theories is lacking. To be clear, much empirical attention has been devoted in recent years to online (and specifically behavioral) advertising. For instance, empirical studies have suggested that, in addition to click-through rates, targeting can lead to increased purchase intentions (Van Doorn & Hoekstra, 2013; Bart et al., 2012), or purchase probability (Manchanda et al., 2006; Lewis & Reily, 2009). However, most studies have focused on the specific offers that consumers clicked or purchased, rather than on counterfactual alternatives that exist and could have been preferred. In other words, existing studies have not specifically focused on the differences in product characteristics between products that are behaviorally targeted and other products in the market that the recipient of the ad may also have access to. For any consumer, clicking on an ad and purchasing the advertised product carries an implicit opportunity cost: the value of the competitor product the consumer could have searched for, or the value of any random product the consumer could have bought with the amount she spent, instead, on the advertised product. In short, we lack *comparative* investigations of different dimensions of the possible welfare implications of targeted ads.

We address this gap by conducting an online within-subjects experiment. The experiment is ongoing, and we present here the results obtained from a first batch of 181 participants. The experiment consists of two stages. In the first stage, we ask participants to provide URLs from ads they see while browsing randomly selected websites. In a second stage, which takes place some days after the first, participants are shown the products from ads that had been presented to them while they were browsing, competing products obtained from a search, and randomly selected products. For each product, participants are asked questions that capture their purchase intentions, perceptions of product quality, price fairness, relevance, and novelty of product type, brand and vendor. We evaluate the consumer welfare implications of behavioral advertising in comparative terms: first, by comparing participants' perceptions associated with targeted products against those associated with a competitor product, we assess how well the ad performs in terms of quality, and price fairness; second, by comparing a targeted ad to an ad targeted to a different person, we assess how well OBA performs in precisely targeting consumers.

The rest of this paper is organized as follows: In section (2) we discuss related literature and the theoretical background of the investigation; in section (3) we present our empirical approach; section (4) describes our experimental design; section (5) presents the results; section (6) discusses our limitations; and in section (7) we discuss the implications of our findings and conclude.

2. Background

Our work contributes to the large body of literature on online advertising. Targeted advertising has existed in some form since before the advent of the Internet through direct channels such as mail or telephone (Roscitt & Parkett, 1988), but wasn't applied as deeply and thoroughly until the emergence of modern tools for online tracking. Our paper analyzes a form of targeting called online behavioral targeting, in which a consumer is targeted based on previous browsing behavior. The targeting occurs in the context of display ads, which appear along the content of websites. These ads can be negotiated directly, through human negotiation, or with a process called programmatic advertisements, in which advertisers bid automatically for a placement in the ad based on the website's data and the user's profile. It is estimated that 87% of the ad placements are negotiated in this way (Fisher, 2019; Samuel et al., 2021). For brevity, this review will concentrate on studies that relate to online targeted advertising, which has had significant contributions in both theory and empirical research.

Among theoretical papers, it is generally assumed that targeted advertising brings better matching of consumers to products. However, how this affects consumer welfare depends on the nuances of each model. Some have suggested that if consumers make the voluntary decision to provide personal information to advertisers, only those who benefit from it would do so, and therefore targeted advertising should be strictly beneficial to consumers (Chen & Stellaert, 2014; Picker; 2009). However, this is problematic, since often consumers are not fully informed of the implications of targeted advertising and may not even be aware of what is happening (Goldfarb & Tucker, 2011a). Iyer et al. (2005) show that firms can increase their profits by targeting customers with strong preference for the product as opposed to comparison shoppers, to reduce price competition. However, this reduced price competition may still lead to consumer welfare improvements despite increases in prices, due to the improved matching (Esteban & Hernandez, 2007; Gal-Or & Gal-Or, 2005). It has also been suggested that either vendors or advertising platforms may have incentives to not target accurately and/or show consumers less preferred options if they are more profitable (Hagiu & Jullien, 2011; Zhang & Katona, 2012; Acquisti, 2014). Amaldoss & He (2010) indicate that when consumers' valuations are low, targeted advertising leads to higher prices, as those few consumers who are less price sensitive have more relative importance for sellers' profits, while high valuations cause targeted advertising to have the opposite effect, as companies compete to attract a higher volume of consumers. Others, however, have argued that targeting benefits consumers as long as the targeting is based on product preferences and not on consumer valuations of the product, since revealing reservation prices may result in an overall negative impact on consumer welfare (Marotta et al., 2021; Varian, 1996). Kshetri (2014) proposes that, by allowing consumers to buy within their affordability range using price discrimination, targeted advertising can help eliminate deadweight losses. On the other hand, Kshetri also argues that highly customized offers can be "unpleasant, creepy and frightening". Some models have incorporated annoyance and privacy concerns into their models. For example, Johnson (2013) indicates that whether targeting benefits or annoys consumers depends on the increase in advertising volume it can produce, and whether such increased volume annoys consumers more than it brings them value. Similarly, Gal-Or et al. (2018) suggest that, depending on how users derive value from the ads, their level of privacy concerns and the cost of improving targeting, the equilibrium level of targeting may be higher or lower than the consumer welfare maximizing level. All these models show that the impact that targeted advertising has on consumer welfare is highly nuanced. There is no simple answer as to whether it is beneficial or detrimental for consumer welfare.

Empirical studies have linked targeted display advertisements to increased click-through rates (Bleier, & Eisenbeiss, 2015; Yan et al., 2009), purchase intentions (Van Doorn & Hoekstra, 2013; Bart et al., 2012), and purchase probability (Manchanda et al., 2006; Lewis & Reily, 2009). These effects, too, are not without nuance. For instance, they are higher for more trusted vendors but lower for less trusted ones (Bleier, & Eisenbeiss, 2015), especially if they disclose their data collection practices (Aguirre et al., 2015). Furthermore, making the user feel in control over their privacy increases ad performance (Tucker, 2014). On the other hand, when an ad is both targeted and obtrusive, ad performance may suffer (Goldfarb & Tucker, 2011b). Also, If targeting is too intense, it can be perceived as intrusive and negatively affect purchase intentions (Van Doorn & Hoekstra, 2013; Bart et al., 2012). Similarly, being interrupted (which occurs frequently with display advertisements) may negatively affect consumer's attitudes towards the ad (Acquisti & Spiekerman, 2011; Duff & Faber, 2011). Another aspect that has been examined is that highly targeted advertisements seem to work better when the customer is further along the purchase funnel (that is, how close is the person to making a purchase) (Lambrecht & Tucker, 2013; Hoban & Bucklin, 2015). However, the experiments described above tend to observe participants' responses under a specific condition, and do not tend to capture the alternatives that may have been present at the time for each participant. Thus, they do not address the specific question we are asking in this paper - how do products shown in targeted ads compare to random products or those that can quickly be identified via search. Two studies showed that increased targeting to specific users can increase consumer welfare (Yao & Mela, 2011; Jeziorsky & Segal, 2015) by offering a better match. These conclusions however came from simulations. In addition, they were performed in a different environment (search advertising) from ours (display advertising). More recent developments have taken a different perspective: observing consumer's purchasing behavior with and without the presence of ad-blockers (Frik et al., 2020; Todri, 2020). Since ad-blockers prevent ads from being displayed, the changes in purchasing behavior can be considered as an indirect measure of how well ads perform. Frik et al., (2020) found mostly no impact of ad-blockers on consumers' economic outcomes. However, the study was based on a search environment and was performed on lab computers, with no behavioral targeting. In an observational study, Todri (2020) found that users' expenditures were reduced after installing ad-blockers, and that the decrease was higher for those brands that were new to consumers. The study, however, did not separate between behavioral and other types of advertisements. It does show, however, that advertisements are an important way for new vendors to reach consumers.

These studies show that behaviorally targeted advertisements have more relevance than non-targeted advertisements, and therefore can help reduce search cost. However, we cannot know whether the products in those advertisements have better prices, quality, or give higher overall satisfaction than other alternatives, and thus we cannot understand the welfare implications. In addition, these papers also show that annoyance and intrusiveness can have negative consequences.

Properly addressing the impact of behavioral targeting on consumer welfare is no simple matter. Because of the immense amount of products, vendors, and services that exist online, it is impossible for a consumer to gather all the information required to make a fully informed decision on the purchase of a product. Consumers therefore must selectively choose a few places to buy from or to search. When a display advertisement is clicked, it carries an implicit opportunity cost, as it constrains the amount of time they will be able to search for products through other channels If the product is not prefered, this may have a negative impact on the consumer. Therefore, just observing consumer behavior between subjects in different conditions is not enough to properly assess these welfare implications. While there have been studies that have addressed this issue in search advertising (Yao & Mela, 2011; Jeziorsky & Segal, 2015; Frik. et al., 2020), to our knowledge no study has applied this to behavioral display advertising.

Our study contributes to this literature by using a counterfactual approach in which online consumers are presented with alternative offers. Experimental participants are asked to assess these alternatives along a variety of metrics. Thus, we assess the welfare implications of behavioral advertising *comparatively*. In a within-subjects design, we present participants with a) products associated with display advertisements that had been previously served to them, against b) competitor products obtained from a search and c) randomly selected products. We compare participants' reactions to these products across several dimensions of relevance to consumer welfare: purchase intentions, quality, price fairness, relevance, and novelty.

3. Empirical approach

In economic theory, it is assumed that consumers derive an unobserved utility from consumption. Consumers make purchase decisions in order to maximize utility. Therefore, when a consumer has to choose between alternatives, it is assumed that the chosen one has higher utility than the other. This is known as revealed preferences.

In our paper we use the notion of revealed preferences to assess, comparatively, the welfare implications of behavioral advertising. We focus on two comparisons. First, we compare products associated with behavioral advertisements served to a participant against competing alternatives present in the market, in order to assess whether the product offered in behavioral advertising is offering a well preferred option of a given type of product. Second, we compare products associated with behavioral advertisements served to a participant against random products, to assess how much more relevant is the targeted offer and whether it is perceived to have better quality, more reasonable prices, and so forth.

We use products from search results for the following reason: the amount of information available online is huge, and consumers can only devote limited time to acquire goods or services online. Clicking on an advertisement deviates from an activity, and reduces the time that the user has available to look for alternatives. Therefore, it implies an opportunity cost. One way this time can be spent instead is using a search engine to find alternative options. If the product offered in the advertisement is good enough, then the extra time required to search would not be worth it and overall, the consumer would be benefited by the ad. If the ad, on the other hand, is poor, then the consumer is wasting valuable time by clicking on it. By using search results, we are therefore able to evaluate how well ads perform in providing the best possible match between product and consumer.

In addition, we compare the products that appeared in the ad to products that appeared in ads viewed by other participants. Since our searches are for similar products to those observed in the ad, we also want to assess whether the ad is providing products that are specifically targeted to the consumer's interest. When we compare between two similar products, we cannot assess whether the product type was the right kind of product to begin with. Using instead a product shown to another participant allows us to know how well the ad performs in this regard.

To compare the products, we first consider purchase intentions - a frequently used measure of consumer preferences in empirical literature, and which empirical literature has found to be correlated with actual purchasing behavior (Morwitz et al., 2007; Pavlou & Fygenson, 2006). There are, however, several factors that influence purchase intentions. We therefore also capture perceptions of quality, price fairness, relevance and familiarity with product type, brand and vendor. Quality, price and relevance directly affect the utility a consumer derives from a product: quality and relevance by increasing the value to the consumer, and price as the main cost incurred.

To measure quality, we ask participants to subjectively assess the quality of the product. Although participants most likely will not have any previous experience with many of the products, the expectation of quality is an essential part of the purchase decision. We do *not* use product ratings as a measure of quality, because they have been shown to have poor correlation with objective quality (Köcher & Köcher, 2018)[2].

We measure price *fairness*, in addition to the price itself, because in many cases the prices cannot be directly compared. Consider, for example, that a product from a high prestige brand is sold five times higher than a similar product by a relatively unknown brand. Even if they are similar products, based on their expectations of quality, the consumer may consider that the price of the prestige brand is reasonable and the price of the unknown brand is not.

We use measure familiarity for two purposes: first, since participants' previous experience will affect quality assessments, we need to have it as a control. Second, we want to gauge the informative value of behavioral advertising - that is, the extent to which behavioral ads bring new information to consumers. If a certain brand, vendor, or product type is new to consumers, it is obvious that they would not be familiar with them. Therefore we use familiarity also as a measure of novelty.

The main goal of our study is to compare each of the variables we measured between our three experimental conditions. The simplest way to do that comparison is with a difference in means of each of our variables across conditions. However, we need to control for both observed and unobserved participant effects, and account for the ordering of the products (they were presented to participants in randomized order). With a simple model with dummy variables we can estimate the differences in means:

$$Y_{ij} = \alpha + \beta_1 Competitor_{ij} + \beta_2 Random_{ij} + \delta X_i + \xi_j + \epsilon_{ij} \quad (1)$$

[2] A simple example can illustrate this. Imagine two pairs of headphones, one from a luxury brand with very high sound quality and another from an obscure brand with average quality. However, because the expectations around the luxury brand are much higher, even while being objectively better, they may get a lower star rating than the cheap ones. At any given price, a consumer would objectively be better off with the luxury brand. However, the star rating quality would make it appear as if it is the opposite. This illustrates why online product ratings are not a good measure to use for welfare implications.

Where Y is represents each of our constructs, *Competitor* and *Random* are dummies that are equal to 1 if the participant is in the specified condition, and ϵ is the error term. X is a vector of participant level controls and we also include sequence level fixed effects (the sequence $j \in [1, 9]$indicates the order in which products were presented).

As price fairness, quality, familiarity, and relevance have been shown to influence purchase intentions (PI) (Dursun et al., 2011; Laroche et al., 1996; Campbell, 1999; Alalwan, 2018), we expect the differences in purchase intentions across conditions to be driven by differences in those variables. We expect those variables to be the main drivers behind purchase Intentions. It is important, in addition to how each variable varies across conditions, to understand how much each of the other variables contributes to purchase intentions, as this would allow us to know how important price, quality, relevance and familiarity are in customer's decisions, in addition to how they differ across conditions, and understanding that relative importance of each variable would allow us to better understand the welfare implications of the differences we observe in model (1).

$$PI_{ij} = \alpha + \beta_1 Quality_{ij} + \beta_2 Price\ Fairness_{ij} + \beta_3 Relevance_{ij} + \beta_4 Fam_{ij} + \delta X_i + \xi_j + \epsilon_{ij}$$
(2)

In addition to estimating the above models, we analyze objective measures of vendor quality and prices. We use three measures of quality: website popularity obtained from SimilarWeb, customer rankings from SiteJabber (a popular online review platform, which is used by Google for measuring vendor quality: Google, n.d.), and rankings from the Better Business Bureau. Vendor quality is an important element of consumer welfare, and together with product quality encompases the whole experience that the consumer derives from a purchase. Since the random products were sampled from the original advertisements seen by participants, these objective measures can only be compared between the ads and the search result, as in expectation targeted ads and random ads should be equal. Also, these measures are uncorrelated with participants' characteristics, as they are measured externally. Since this is the case, differences of quality between conditions can be assessed with a simple difference in means.

For prices, we do two comparisons. First, we compare the mean prices between the ad and the search result using differences in mean of log(price). Since the ad and the search results represent similar products, the prices are expected to be comparable. However, because we don't have an objective measure of product quality, this comparison may be affected by omitted variable bias. To address this bias we also collect, for products that were sold by more than one vendor, different prices available online for the same exact product. We then compare the price obtained from the ad with the lowest price that was found through search results. Since prices are heterogeneous, to make them comparable we use logs of prices.

4. Experimental design

Our design consists in a within-subjects experiment in which consumer preferences are compared across the three conditions: products associated with display advertisement, products associated with search results, and randomly selected products (from advertisements shown to other participants).

Before the study began, we undertook a selection process to collect websites for participants to visit (Section 4.1: Website selection process). The intent of the selection protocol was to expose participants during the experiment to ads that would have a high likelihood of being behaviorally targeted. In Stage 1 of the experiment, recruited participants were randomly assigned to browsing a subset of the selected websites, and were asked to obtain URLS from ads that appeared on those websites (Section 4.2: Stage 1). Immediately after the completion of Stage 1, we used a combination of scripts and searches to collect information about the products associated with the ads that had been served to participants during Stage 1, as well as competitor products and random products (Section 4.3: Interstage process). A few days later, participants were invited to participate in Stage 2 (Section 4.4: Stage 2). During that Stage, participants were presented with offers from the three within-subject conditions, and were asked questions that captured our variables of interest, as well as control variables such as demographics and usage of technologies that may affect the types of ads they saw during the study. Figure 1 summarizes the process.

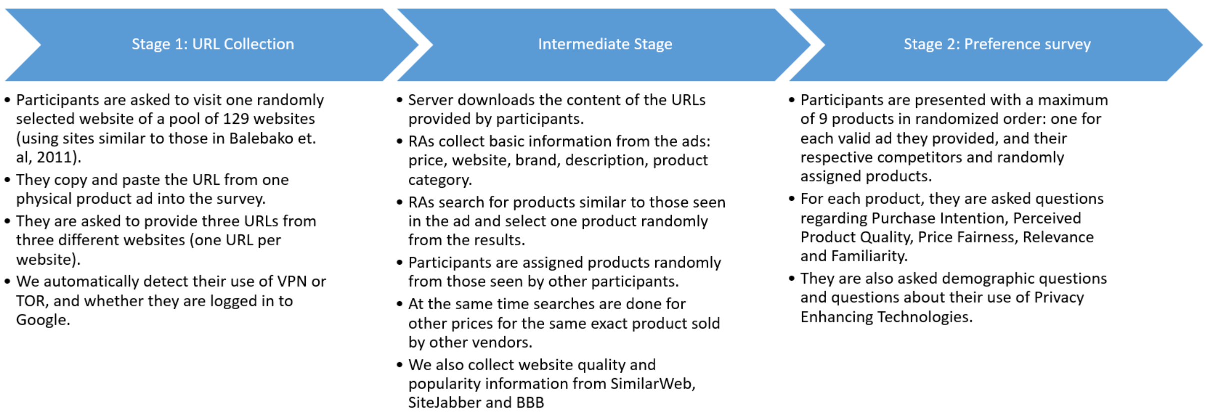

Figure 1: Summary of the experimental procedure.

4.1 Website Selection Process

Before we began the experiment, we identified websites that were appropriate for participants to visit and obtain ads from. We focused on websites in English that serve text (as opposed to video or audio) content, cater to wide audiences, and serve behavioral ads. Our website selection process was modelled after Balebako et al. (2012). We seeded the set of sites with two popular news and entertainment websites,[3] as those websites cater to broad audiences and broad topical interests, making them more likely to serve behavioral advertisements (Balebako et al., 2012). While Balebako et al. (2012) used a total of five news and entertainment websites, we expanded their approach to include a broader list of websites. We used the Amazon Alexa "audience overlap" tool to find such additional sites. The audience overlap tool provides a list of websites that have overlapping visitors with a given website. This generated two lists: one for news sites and one for entertainment sites. We merged them and went through each website individually. We included in the final list all the websites that served free text (as opposed to audio or video) content, served ads, did not require login to see (at least some) content, and showed in our tests to display behavioral ads. To determine whether a website frequently serves behavioral ads, each site was visited

[3] The seeding sites were www.cnn.com and www.tmz.com.

several times from different computers with different profiles and then the AdChoices Icon was clicked to capture information about why the ad was served. Examples of these are shown in Figure 2. The AdChoice information is mostly available for ads served by Google. For ads that did not provide this information, we observed how closely the ad matched the website's content. In addition, we added a set of smaller websites that we obtained from *Discuuver*, a random useful website generator that provides visitors with links to lesser known websites. Smaller websites were added to account for the fact that some large vendors may engage in wide publicity campaigns in well-known websites, while smaller websites may have to rely more on purely behavioral targeting applied by the ad networks. This process resulted in a list of 129 websites (see Appendix 1).

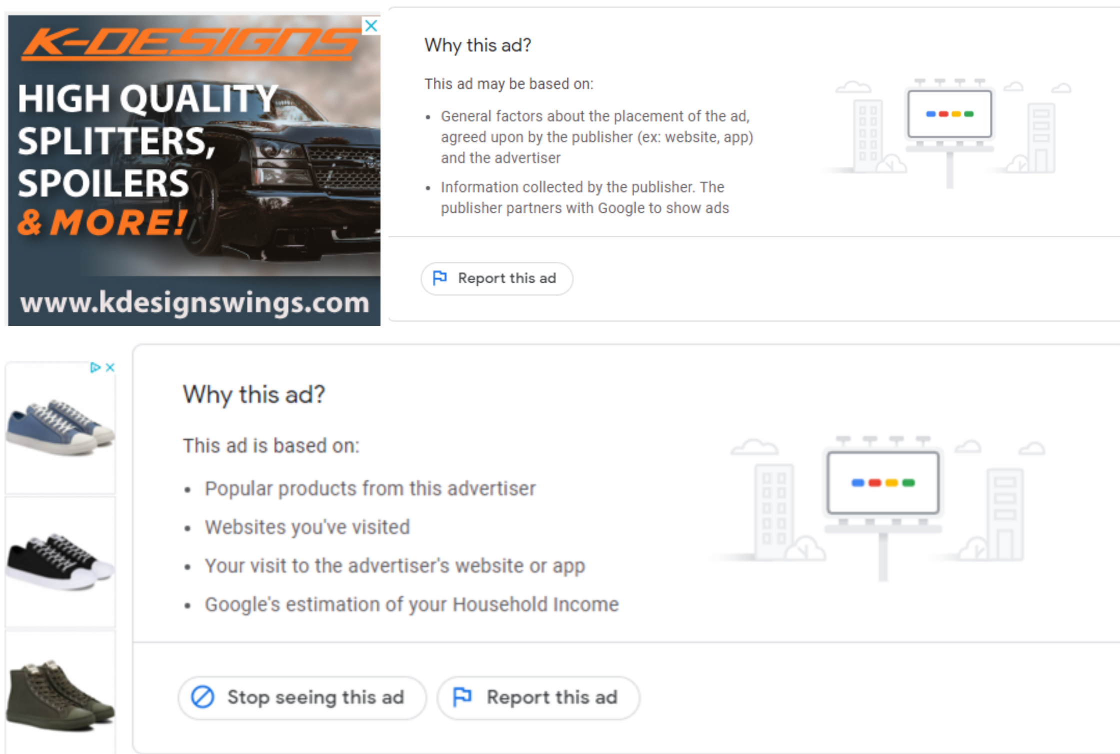

Figure 2: The ad on the top left appeared on an online trucking magazine. The profile used had never visited any website of the kind. The "Why this ad?" explanation suggests the ad is contextual. In the bottom panel we see an ad for shoes in a computer that previously visited several shoes shopping sites. The information under "Why this ad?" clearly states it is using behavioral data.

4.2 Stage 1

The actual experiment began in Stage 1. Participants filled out an online consent form. Thereafter, participants were presented with instructions. The instructions described what types of ads to look for while browsing during the experiment. We focused on physical products that can be purchased online, do not require subscriptions, preparation or customization, and can be delivered to their homes. We used these criteria since those kinds of products are the ones which, in Stage 2, participants could more easily assess by looking at their image and a small amount of additional textual information. After the instructions, participants were asked an attention check question. Then, they were shown a video tutorial

(1:46 total duration) that explained each step of the process. Finally, they were asked to actually perform the task.

The task consisted in participants being presented with a link to a randomly picked site from the selected set of websites (see Section 4.1). Participants were asked to visit the site with their own devices and browsers. On the site, they were asked to obtain the URLs of the first three ads they saw that met the criteria described above, and paste those URLs in the survey form. In case they had issues obtaining a valid ad from the site, they were offered a button that would allow the participant to request a new randomly selected site. We repeated this process three times, so that each participant provided a total of three links from three unique websites. Figure 3 shows an example of an ad being shown on the CNN website.

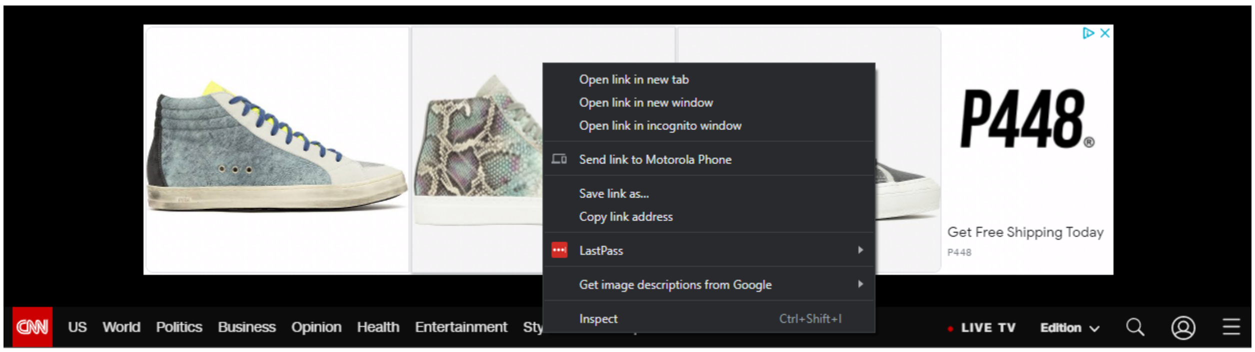

Figure 3: An advertisement on a website. The participant right-clicks on the advertisement for the menu to appear, and then must select "copy link address" to copy the URL. The participant then pastes the URL in our survey.

During this stage, we detected whether the participant was connecting through a VPN, Proxy or TOR (in addition, participants were also asked about their previous usage of these and other technologies in Stage 2) and whether they were logged in to Google. Although we asked participants to disable ad blockers for the task (since otherwise they would not be able to perform it), the presence of ad blockers in the system may change the way they are targeted when they are paused. Similarly, the use of VPN, Proxy or TOR services may affect the quality of targeting. We captured these data to include them as controls in our analysis, since we are also interested in how these technologies change what participants observe.

4.3 Interstage process

Right after each participant provided the three URLs, the html contents of the landing page were downloaded to one of our servers and the interstage process of collecting information about the products associated with the ad landing URL commenced. Once the server downloaded the URLs and the html contents, Research Assistants (RAs) received a notification to start collecting the following information from the pages stored in the server: product image, price, description, brand and name of the website that sells the product.[4] (If the server got blocked by the website, the RA downloaded the URL instead and uploaded a full page screenshot to the server.)

If the landing page showed more than one product, the RAs obtained the information for the one in the top left. The position was picked based on the assumption that vendors sort their products to increase purchase likelihood, and therefore the first product should be the most likely to be relevant. If the landing page did not lead to an offer selling a physical product, then the link was discarded and excluded from the study. Figures 4 and 5 show examples of landing pages with a single and multiple products, respectively.

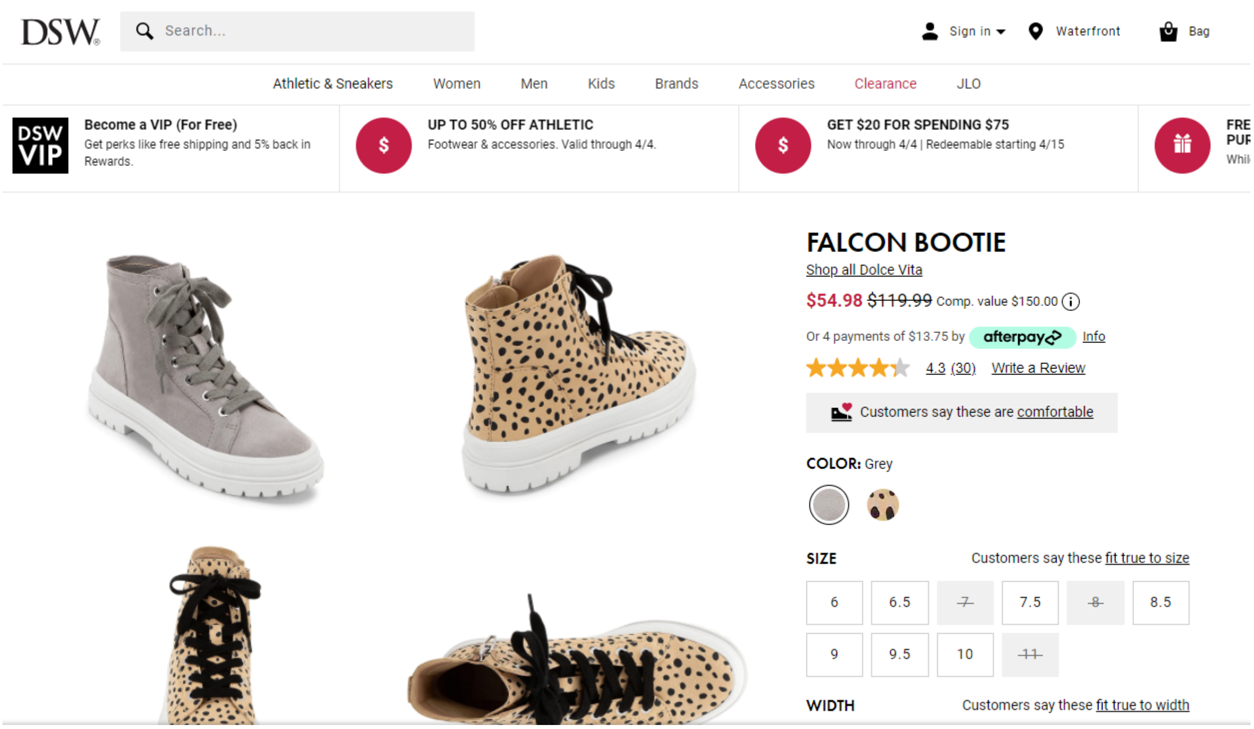

Figure 4: A landing page from an ad with a single product.

[4] In order to ensure that this would not introduce any bias in data collection, we conducted a pilot test using computers in different locations and obtained ads from a random sample of 50 of our websites and captured ads from each site. Multiple testers (research assistants) sent a link to a chat for all those present in an online meeting to visit, and in all cases we confirmed that the same URL when visited from a different location leads to the same exact landing page with the same products and the same prices. Therefore, we can be confident that what is downloaded by our server is what the user would have seen if they left-clicked on the ad instead of just copying the URL.

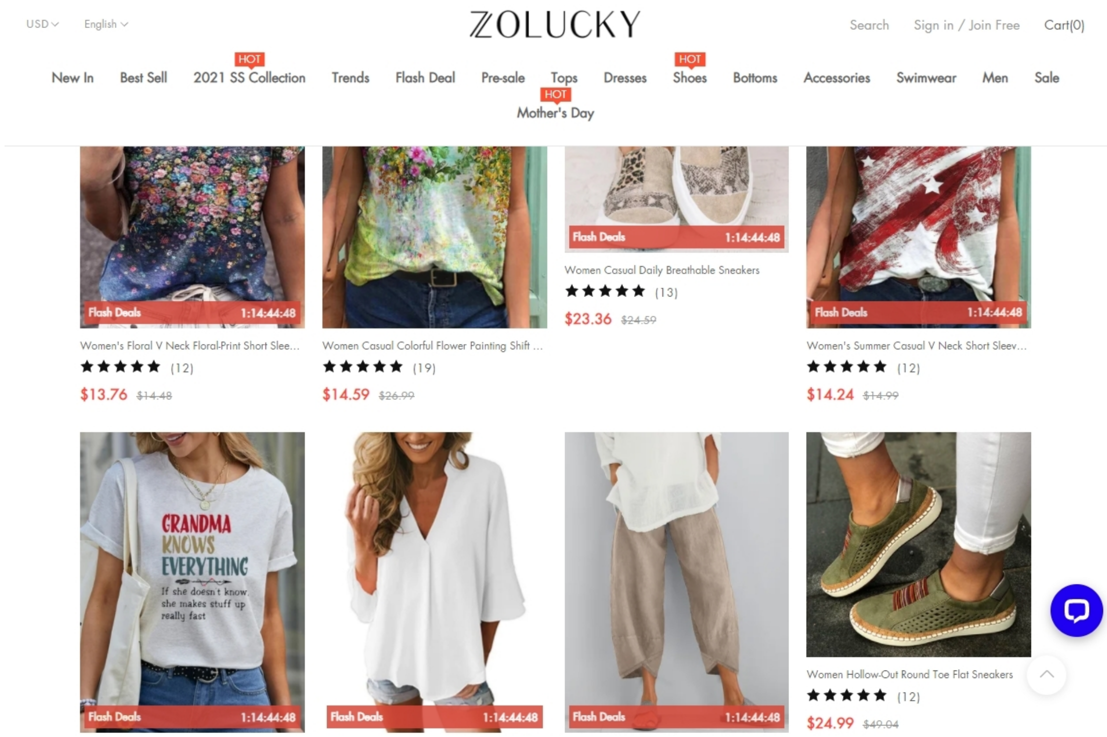

Figure 5: A landing page with multiple products. In these cases, the top left was selected to present to the participant.

Once the information for the product in the ad was obtained, the RAs obtained the competitor product by performing searches on a popular engine (Google). After the RAs finished this task, our server automatically also assigned, to each participant, a *random* product for each valid link they provided. They were obtained from ads provided by other participants.

To obtain the competing product, the RAs obtained the search term from the original product description. If they did not find competing products, they made small incremental changes to the search terms to make them less specific and repeated the search. They did so until they obtained search results that showed offers for varied brands and vendors for competing products. Appendix 2 details the full algorithm that was followed.

Then we randomized whether they used one of the organic search results on the first page or the first sponsored search result. If the landing page of the result was not usable (for example, it did not have offers for the product, showed as out of stock or sold in a foreign currency), then the next result to click would also be determined randomly from the remaining ones, and so on until a usable search result was found. We focus mostly on organic search results since those are based on relevance and do not constitute proper advertising, and therefore are completely different from display advertisements. However, we included a small amount of sponsored results in our sample to account for the fact that they may provide viable alternatives for consumers.

Once the RA landed on an appropriate page, they would collect the same information that was collected for products in the ads.

4.4 Stage 2

Participants who had taken part in the first stage of the study were invited to Stage 2 using the Prolific platform messaging feature. Stage 2 took place typically a week after Stage 1. The moderate time lapse addressed two goals: adding time in-between the stages to reduce the likelihood of recall effects from Stage 1; but not adding so much time as to increase attrition.

After being welcomed back to the study and reading a brief introduction, participants were presented with products from the three experimental conditions: products associated with ads that had been served to them while they browsed randomly selected sites in Stage 1; competing products; and random products (that is, products randomly chosen from the set associated with ads served to *other* participants in the study). Participants were presented with up to three such triads of products, and they were asked to evaluate each product independently and in random order.

For each product, the participant was shown its image, price, description, and brand - but (initially) no information about its vendor website. Participants were asked a series of questions derived from the literature, and were asked to respond on 7 point likert scales.

To measure purchase intention, we adapted a triad of questions from Shaouf et al. (2016) :

1. "Based on the information above, I became interested in making a purchase." (PI1)
2. "Based on the information above, I am willing to purchase this product from this website." (PI2)
3. "Based on the information above, I will probably purchase this product in the next month." (PI3)

All questions used as endpoints *Strongly disagree - Strongly agree*.

We used single likert scale questions adapted from previous studies to measure other perceptions of each product:

4. Quality: "Based on the information above, what is the overall quality of this product likely to be?" (Extremely low - Extremely high) (Kirmani, 1990);
5. Price fairness: "How reasonable is the price for this product?" (Very unreasonable - Very reasonable) (Kwak et al., 2015);
6. Relevance: "This product might be relevant to my needs" (Strongly disagree - Strongly agree) (Laczniak & Muehling, 1993)
7. Familiarity with product type: "How familiar are you with [product type]?" (Not familiar at all - Extremely familiar) (Darley, & Smith, 1995)
8. Familiarity with brand: "How familiar are you with [brand]?" (Not familiar at all - Extremely familiar) (Darley, & Smith, 1995)

After capturing a participant's answers to the above questions, we presented the same product information a second time, but added to it information about the vendor website. We then asked the participant to answer a modified purchase intention (PI4) question that makes reference to the website and a question about familiarity with the website:

9. “Based on the information above I will probably purchase this product from this website in the next month.” (PI4)
10. “How familiar are you with [vendor]?”

We did not include information on the website selling the product when we *first* presented it to the participant to isolate participants’ perceptions of the product from their perceptions of the vendor (which may affect the former). Figure 6 shows an example of how we presented the product to the participants.

Next, we asked participants about their use of Privacy Enhancing Technologies (PET). A variety of PETs can be used by consumers to block advertisements, reduce the amount of information that can be collected about them, or disguise specific information about them such as location. These technologies affect the effectiveness of targeted advertisements, since they may prevent companies from collecting data about users. Although participants were asked to temporarily disable ad blockers for the experiment, the ads they see may be affected by the fact that they had been blocking ads before. We control for this both using our automated detection metrics (see Section 4.2) and by asking participants about their use of these technologies.

The technologies we asked participants about included: 1) Browser extensions, which can be either ad blockers (prevents ads from loading) or anti trackers (prevents tracking consumers across websites) or both; 2) Networking based solutions, which can be used to disguise the location and IP Address of the user to reduce tracking. These include VPNs, Proxies and the TOR Browser; 3) Opt outs: These are provided by either individual companies or industry alliances for consumers to allow consumers to ask participating companies to not use tracking information to display advertisements, and can be effective in reducing the amount of behaviorally targeted ads (Belbako et al., 2012); specifically, we ask about opt outs for Google, Amazon, Facebook, and the DAA and NAI industry alliances; 4) Their current cookies settings: since settings within the browser related to cookies affect targeting, it is important to know if a participant has changed the cookie settings from the default settings.

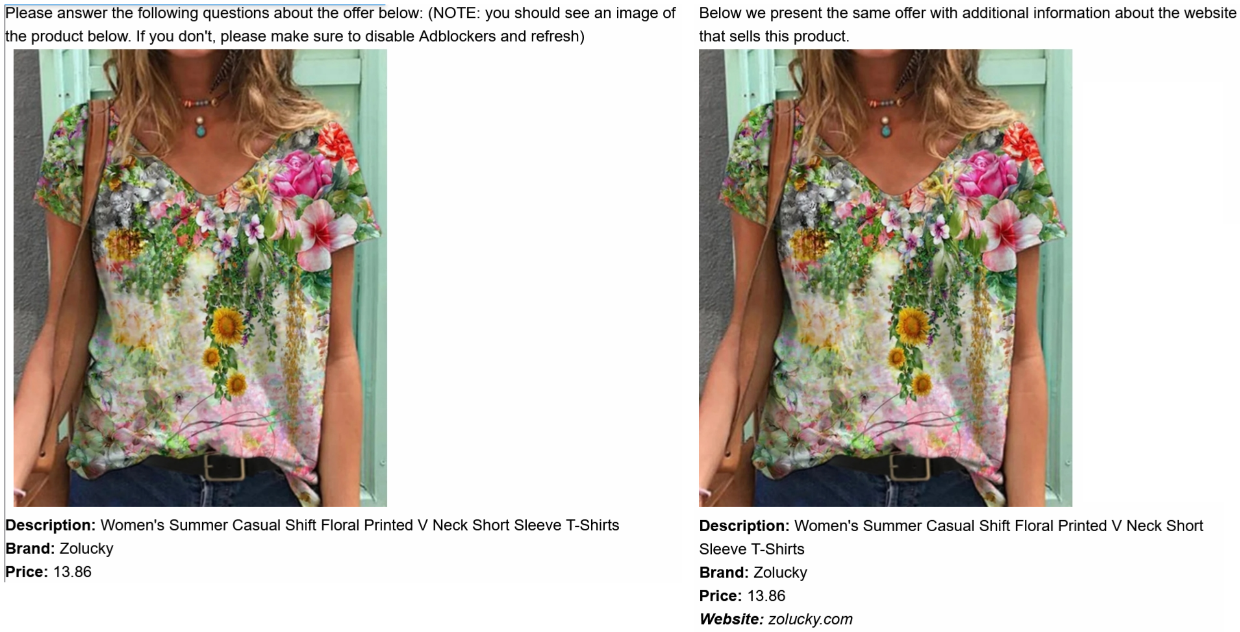

Figure 6: On the left side we observe how participants were presented with the product first without the name of the website and asked all the questions except PI4 and familiarity with the website. Then we presented the screen on the right side and asked the remaining two questions.

We also asked participants about the search engines and browsers they commonly use. Although we explicitly recruited participants that use Google Chrome as their main browser, we intended to control for potential usage of other browsers.[5]

Finally, we collected demographic information, such as age, gender, highest education level, employment and state of residence to include as controls.

The study was preregistered.

5. Preliminary Results

Participants were recruited using the Prolific Academic platform, which has been shown to be a reliable platform for research studies (Peer et al., 2017). In order to participate, prospective participants were required to reside in the US, have completed at least 100 assignments on the platform and have a 95% or higher approval rate. Participants were offered monetary compensation for their participation ($1.67 for stage 1, $2.83 for stage 2, $4.50 total, mean completion time 27 minutes). In addition, we used some recruitment criteria to decrease the likelihood of confounds affecting the types of ads participants would be exposed to in the study. In particular, we did not want participants who would be served less behavioral ads than others because of their browser. Thus, participants were required to have Google Chrome - the browser with the least restrictive default privacy settings - as their main browser. As Chrome is currently

[5] It would have been difficult to detect automatically if the user had changed their browser settings, or had opted out of tracking. Hence survey responses complement our detection tools. In addition, survey questions allow us to determine whether participants regularly use those technologies even if they were not active during the time they fill out the survey.

the most widely used browser (Statcounter, 2021), this restriction does not significantly limit the external validity of our results. Due to its default configuration, targeting will be the highest for Chrome users versus other browsers, and therefore allow us to better identify effects between conditions. Participants were also requested not to use incognito mode. Incognito mode starts a temporary session without any cookies, and all cookies created during this mode are deleted after the session ends. Because of this, the information available for targeting during incognito sessions is reduced, which would limit the effects of targeting. Both stages of the study were approved by Carnegie Mellon University's IRB.

Here, we present descriptive statistics from a first batch of 181 participants. The study is ongoing, and new batches are being recruited until we reach a sample of over 450 participants. The target sample size was determined via power analysis through simulations of our empirical model. The analysis determined that in order to obtain 95% power under several different scenarios, 1,000 ads were required to be collected. By assuming a 70% attrition rate from Stage 1 to Stage 2, and an 80% valid link rate (the ratio of links that meet the criteria we specified), we determined a goal of obtaining 1,786 URLs in Stage 1.

Our current sample is quite diverse in terms of age, which ranges from 18 to 75 (m=36.31, sd=11.44) and state of residence. We have participants from 39 states, and 41% percent of the participants are female and 59% male. The sample is highly educated: 91% of our sample has a degree and 9% only completed High School. 56% of the participants are full time employees, 15% are students and the rest have various situations.

Regarding use of Privacy Enhancing Technologies, 91 of our participants use ad blockers or anti trackers. 22 report usage of TOR or VPN, and 70 report having used opt outs. Usage of these tools is therefore relatively widespread. Overall, 115 of our 181 participants used at least one of these technologies.

In Stage 1, each participant provided 3 links, for a total of 543 links. However, 107 (20%) of links did not meet our specified criteria (physical products) and were not considered for Stage 2, producing a set of 436 usable advertised products, to which our experimental procedure added 436 competitor products and 436 random products, for a total of 1,308 data points.

Table 1 shows summary results for our variables of interest. It also shows repeated measures Anova tests for differences across conditions and the t values of contrasts between the ad condition and the other conditions.[6]

[6] As the results are preliminary and the data collection is ongoing, the current version of the manuscript (June 14, 2021) omits a detailed analysis of the results. Please contact the authors for a current version of the manuscript with updated results.

Variable/Type	Ad		Competitor		Random		RMANOVA	Contrast T	Contrast T
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	F	Ad vs Competitor	Ad vs Random
PI1 (interested in making a purchase)	3.35	1.99	3.22	1.91	2.95	1.85	7.23**	-0.12 (0.11)	-0.39** (0.11)
PI2 (willing to purchase this product)	3.35	2.00	3.29	1.97	3.04	1.92	4.52*	-0.05 (0.11)	-0.31* (0.11)
PI3 (will probably purchase this product in the next month)	2.77	1.83	2.65	1.76	2.36	1.68	11.55**	-0.12 (0.09)	-0.41** (0.09)
PI4 (will probably purchase this product from this website in the next month)	2.66	1.75	2.54	1.69	2.40	1.68	4.85**	-0.12 (0.09)	-0.26** (0.09)
Average of PI1, PI2, PI3	3.16	1.86	3.05	1.78	2.79	1.72	15.84**	-0.10 (0.09)	-0.37** (0.09)
Price Fairness	4.45	1.85	4.66	1.77	4.29	1.81	5.31**	0.20 (0.11)	-0.16 (0.11)
Perceived Quality	4.81	1.32	4.70	1.24	4.68	1.21	2.06	-0.11 (0.07)	-0.13 (0.07)
Relevance	4.01	2.05	3.94	2.08	3.58	1.98	23.76**	-0.06 (0.11)	-0.43** (0.11)
Familiarity with product type	5.00	1.85	4.93	1.82	4.51	2.06	13.27**	-0.05 (0.1)	-0.48** (0.1)
Familiarity with brand	3.11	2.34	2.78	2.23	2.83	2.21	3.26*	-0.33* (0.14)	-0.27* (0.14)
Familiarity with vendor	3.42	2.48	4.19	2.56	3.00	2.34	29.27**	0.78** (0.16)	-0.40* (0.16)

Table 1: Summary stats for all our variables of interest. 1 = Strongly Disagree/Extremely Unfamiliar/Very Low Quality/Very Unreasonable, 7 = Strongly Agree/Extremely Familiar/Very High Quality/Very Reasonable. N=759 for cases when “clothes group” is included, N=606 when they are excluded. The last three columns show the Repeated measures Anova for differences across conditions and differences in means for the contrasts. Standard errors in parenthesis. * $p<0.05$, ** $p<0.01$

6. Limitations

An important limitation in our experiment is that we use purchase intentions, which may not necessarily translate always into actual purchases. However, studies have shown that purchase intentions are in fact a good proxy (Morwitz et al., 2007; Pavlou & Fygenson, 2006). We plan to address this issue with a future study that will incorporate actual behavior.

Another limitation is that answers on the use of Privacy Enhancing Technologies are self reported, and therefore may be inaccurate. However, we were able to capture whether some of them were active during the study. Specifically, we were able to detect VPN, TOR and users being signed in to Google. We are unable to confirm automatically the use of incognito mode, opt outs or changes in the user’s browser settings. We tried to obtain automatic measures for the presence of ad blockers, but they failed in the middle of the study and we were not able to incorporate them. For those that we detect, it was still necessary to ask participants since failure on our side to detect it would not mean that they have never used it, they could have disabled it temporarily before accessing the study. We therefore incorporate both self reported and automatically obtained measures into our dataset.

In addition, it is important to explain that in the advertisement condition, there are two types of advertisements that the participants may see: behavioral and contextual. Behavioral ads are based on the user’s past behavior, while contextual ads are based on the content of a webpage. For example, let’s assume that a person suddenly gets an interest in superhero comics, and visits for the first time the website of an important comics seller. Also, the person has recently visited the website of a robot vacuum manufacturer. We assume that there is no correlation between being a comics fan and buying a robot vacuum. If the person, who is visiting the comics website for the first time, sees an ad for a robot vacuum, the ad would be behavioral, If they see an ad for a comics store, the ad would be contextual. Because we are interested in the impact of behavioral ads, this could pose a problem. However, we solve this problem by randomizing the websites that each participant visits from the ads. This randomization causes the influence of contextual ads to be the same across conditions. First, because the contextual ads seen in the random and the behavioral condition are going to be the same. Second, because the competitor condition comes from a search based on what is seen in behavioral ads, and therefore is affected in the same way. Our differences are therefore well identified. However, the presence of contextual ads still may introduce (although small) some noise. In addition, the fact that we observed differences between the ad and the random condition suggests that the ads observed were behavioral, as if they were not, random and ad conditions would have observed similar values.

Also, in order to reduce heterogeneity, our experiment was done with display advertisements that appear on desktop computers when visiting text content websites where log-in is not required. Our results may not extend to Social Networking websites or other types of websites where the user is required to be logged in, since those websites can collect more precise data about the consumer. They may also not extend to websites whose content is mainly audio or video. Since mobile platforms may use geolocation and other additional forms of data collection not available on Desktop, as well as using a different format, our results also may not extend to mobile platforms. These issues can be addressed in future experiments by focusing on those platforms.

References

Acquisti, A. (2014). Inducing customers to try new goods. Review of Industrial Organization, 44(2), 131-146.

Acquisti, A., & Spiekermann, S. (2011). Do interruptions pay off? Effects of interruptive ads on consumers' willingness to pay. Journal of Interactive Marketing, 25(4), 226-240.

Aguirre, E., Mahr, D., Grewal, D., De Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. Journal of retailing, 91(1), 34-49.

Alalwan, A. A. (2018). Investigating the impact of social media advertising features on customer purchase intention. International Journal of Information Management, 42, 65-77.

Amaldoss, W., & He, C. (2010). Product variety, informative advertising, and price competition. Journal of Marketing Research, 47(1), 146-156.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. The Quarterly Journal of Economics, 135(2), 645-709.

Balebako, R., Leon, P., Shay, R., Ur, B., Wang, Y., & Cranor, L. (2012, May). Measuring the effectiveness of privacy tools for limiting behavioral advertising. Web.

BBB. (n.d.). Overview of BBB Ratings: Better Business Bureau®. Retrieved May 05, 2021, from https://www.bbb.org/overview-of-bbb-ratings

Bart, Y., Stephen, A. T., & Sárváry, M. (2012). A Field Study of the Determinants of Mobile Advertising Effectiveness. INSEAD.

Bleier, A., & Eisenbeiss, M. (2015). Personalized online advertising effectiveness: The interplay of what, when, and where. Marketing Science, 34(5), 669-688.

Campbell, M. C. (1999). Perceptions of price unfairness: antecedents and consequences. Journal of marketing research, 36(2), 187-199.

Chen, J., & Stallaert, J. (2014). An economic analysis of online advertising using behavioral targeting. Mis Quarterly, 38(2), 429-A7.

Chiasson, S., Abdelaziz, Y., & Chanchary, F. (2018). Privacy concerns amidst OBA and the need for alternative models. IEEE Internet Computing, 22(2), 52-61.

Comanor, W. S., & Wilson, T. A. (1974). Advertising and market power (No. 144). Harvard University Press.

Cramer-Flood, E. (2020, July 06). Global digital ad Spending Update Q2 2020. Retrieved February 09, 2021, from https://www.emarketer.com/content/global-digital-ad-spending-update-q2-2020

Dehling, T., Zhang, Y., & Sunyaev, A. (2019, July). Consumer Perceptions of Online Behavioral Advertising. In 2019 IEEE 21st Conference on Business Informatics (CBI) (Vol. 1, pp. 345-354). IEEE.

Darley, W. K., & Smith, R. E. (1995). Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. Journal of advertising, 24(1), 41-56.

Digital Advertising Alliance (DAA). (2020, September 28). Americans value free Ad-supported online services AT $1,400/Year; annual value Jumps more than $200 since 2016. Retrieve March 31st, 2021 from https://digitaladvertisingalliance.org/press-release/americans-value-free-ad-supported-online-services-1400year-annual-value-jumps-more-200

Duff, B. R., & Faber, R. J. (2011). Missing the mark. Journal of Advertising, 40(2), 51-62.

Dursun, I., Kabadayı, E. T., Alan, A. K., & Sezen, B. (2011). Store brand purchase intention: Effects of risk, quality, familiarity and store brand shelf space. Procedia-Social and Behavioral Sciences, 24, 1190-1200.

Esteban, L., & Hernandez, J. M. (2007). Strategic targeted advertising and market fragmentation. Economics Bulletin, 12(10), 1-12.

Fisher, L. (2019, April 25). US programmatic ad SPENDING Forecast 2019. Retrieved April 01, 2021, from https://www.emarketer.com/content/us-programmatic-ad-spending-forecast-2019

Frik, A., Haviland, A., & Acquisti, A. (2020). The impact of ad-blockers on product search and purchase behavior: A lab experiment. In 29th {USENIX} Security Symposium ({USENIX} Security 20) (pp. 163-179).

Gal-Or, E., & Gal-Or, M. (2005). Customized advertising via a common media distributor. Marketing Science, 24(2), 241-253.

Gal-Or, E., Gal-Or, R., & Penmetsa, N. (2018). The role of user privacy concerns in shaping competition among platforms. Information Systems Research, 29(3), 698-722.

Goldfarb, A., & Tucker, C. E. (2011a). Online advertising, behavioral targeting, and privacy. Communications of the ACM, 54(5), 25-27.

Goldfarb, A., & Tucker, C. (2011b). Online display advertising: Targeting and obtrusiveness. Marketing Science, 30(3), 389-404.

Google. (n.d.). About seller ratings ads extensions. Retrieved May 08, 2021, from https://support.google.com/google-ads/answer/2375474?hl=en

Hagiu A, Jullien B (2011) Why do intermediaries divert search?RAND Journal of Economics42(2):337–362.

Hoban, P. R., & Bucklin, R. E. (2015). Effects of internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment. Journal of Marketing Research, 52(3), 375-393.

Iyer, G., Soberman, D., & Villas-Boas, J. M. (2005). The targeting of advertising. Marketing Science, 24(3), 461-476.

Jeziorski, P., & Segal, I. (2015). What makes them click: Empirical analysis of consumer demand for search advertising. American Economic Journal: Microeconomics, 7(3), 24-53.

Jiang, M. (2014). Search concentration, bias, and parochialism: A comparative study of Google, Baidu, and Jike's search results from China. Journal of communication, 64(6), 1088-1110.

Johnson, J. P. (2013). Targeted advertising and advertising avoidance. The RAND Journal of Economics, 44(1), 128-144.

Kirmani, A. (1990). The effect of perceived advertising costs on brand perceptions. Journal of consumer research, 17(2), 160-171.

Köcher, S., & Köcher, S. (2018). Should we reach for the stars? Examining the convergence between online product ratings and objective product quality and their impacts on sales performance. Journal of Marketing Behavior, 3(2), 167-183.

Kshetri, N. (2014). Big data′ s impact on privacy, security and consumer welfare. Telecommunications Policy, 38(11), 1134-1145.

Kunst, A. (2021, January 25). Online purchases by category in the United STATES 2020. Retrieved May 03, 2021, from https://www.statista.com/forecasts/997093/online-purchases-by-category-in-the-us

Kwak, H., Puzakova, M., & Rocereto, J. F. (2015). Better not smile at the price: The differential role of brand anthropomorphization on perceived price fairness. Journal of Marketing, 79(4), 56-76.

Laczniak, R. N., & Muehling, D. D. (1993). The relationship between experimental manipulations and tests of theory in an advertising message involvement context. Journal of Advertising, 22(3), 59-74.

Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. Journal of Marketing research, 50(5), 561-576.

Laroche, M., Kim, C., & Zhou, L. (1996). Brand familiarity and confidence as determinants of purchase intention: An empirical test in a multiple brand context. Journal of business Research, 37(2), 115-120.

Lewis, R. A., & Reiley, D. H. (2014). Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on Yahoo!. Quantitative Marketing and Economics, 12(3), 235-266.

Manchanda, P., Dubé, J. P., Goh, K. Y., & Chintagunta, P. K. (2006). The effect of banner advertising on internet purchasing. Journal of Marketing Research, 43(1), 98-108.

Marotta, V., Zhang, K., & Acquisti, A. (2021). The welfare impact of targeted advertising. Information Systems Research. Forthcoming.

McCue, T. (2019, March 20). 47 percent of consumers are blocking ads. Retrieved February 09, 2021, from https://www.forbes.com/sites/tjmccue/2019/03/19/47-percent-of-consumers-are-blocking-ads/?sh=81b00c42037e

Morwitz, V. G., Steckel, J. H., & Gupta, A. (2007). When do purchase intentions predict sales?. International Journal of Forecasting, 23(3), 347-364.

Pavlou, P. A., & Fygenson, M. (2006). Understanding and predicting electronic commerce adoption: An extension of the theory of planned behavior. MIS quarterly, 115-143.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology, 70, 153-163.

Picker, R. C. (2009). Online advertising, identity and privacy. U of Chicago Law & Economics, Olin Working Paper, (475).

Roscitt, R., & Parket, I. R. (1988). Direct marketing to consumers. Journal of Consumer Marketing.

Samuel, A., White, G. R., Thomas, R., & Jones, P. (2021). Programmatic advertising: An exegesis of consumer concerns. Computers in Human Behavior, 116, 106657.

Shaouf, A., Lü, K., & Li, X. (2016). The effect of web advertising visual design on online purchase intention: An examination across gender. Computers in Human Behavior, 60, 622-634.

Statcounter. (2021, February). Browser market share worldwide. Retrieved March 31, 2021, from https://gs.statcounter.com/browser-market-share

Todri, V. (2020). The Impact of Ad-blockers on Online Consumer Behavior. Marketing Science (Forthcoming).

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. Journal of marketing research, 51(5), 546-562.

Varian HR (1996) Economic aspects of personal privacy. Privacy and Self-Regulation in the Information Age (National Telecommunications and Information Administration, Washington, DC)

Van Doorn, J., & Hoekstra, J. C. (2013). Customization of online advertising: The role of intrusiveness. Marketing Letters, 24(4), 339-351.

Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., & Chen, Z. (2009, April). How much can behavioral targeting help online advertising?. In Proceedings of the 18th International Conference on the World wide Web (pp. 261-270).

Yao, S., & Mela, C. F. (2011). A dynamic model of sponsored search advertising. Marketing Science, 30(3), 447-468.

Zhang K, Katona Z (2012) Contextual advertising.Marketing Science. 31(6):980–994

Appendix 1. List of websites used for the experiment.

abc7.com
abcnews.go.com
agoodmovietowatch.com
alarmdj.com
aljazeera.com
alphabetize.org
apnews.com
bbc.com
bet.com
biography.com
bossip.com
bostonglobe.com
bravotv.com
breitbart.com
britannica.com
businessinsider.com
buzzfeed.com
buzzfeednews.com
cbslocal.com
cbsnews.com
celebritynetworth.com
cheatography.com
chicagotribune.com
chron.com
cnn.com
consequenceofsound.net
cosmopolitan.com
dailymail.co.uk
deadline.com
deadspin.com
dictation.io
drudgereport.com
dw.com
elle.com
etonline.com
ew.com
express.co.uk
extratv.com
famousbirthdays.com
faxzero.com
flap.tv
forvo.com

foxbusiness.com
gate2home.com
gizmodo.com
globalnews.ca
goodreads.com
healthline.com
heavy.com
hellomagazine.com
hollywoodreporter.com
hotnewhiphop.com
huffpost.com
ibtimes.com
independent.co.uk
insideedition.com
insider.com
intouchweekly.com
jezebel.com
justjared.com
keepmeout.com/en
latimes.com
legacy.com
lifeandstylemag.com
livescience.com
marketwatch.com
mercurynews.com
metro.co.uk
mirror.co.uk
msn.com
mtv.com
myfridgefood.com
mymeetingtime.com
nbcnews.com
news.com.au
nme.com
npr.org
nydailynews.com
nypost.com
nytimes.com
pagesix.com
people.com
perezhilton.com
politifact.com
popculture.com
radio.com

realclearpolitics.com
reuters.com
runpee.com
sbnation.com
scmp.com
screenshot.guru
sfgate.com
si.com
signature-maker.net/email-signature
slate.com
snopes.com
sortmylist.com/
spin.com
sportingnews.com
the-sun.com
theatlantic.com
thedailybeast.com
thegrio.com
theguardian.com
thehill.com
thehollywoodgossip.com
theverge.com
thewrap.com
tmz.com
toofab.com
topdocumentaryfilms.com
tvguide.com
upi.com
usatoday.com
usmagazine.com
vulture.com
weather.com
willrobotstakemyjob.com
wmagazine.com
worldometers.info
www.accountkiller.com/en
www.brianfolts.com/driver
www.copypastecharacter.com
www.flightradar24.com
www.handspeak.com/word
www.marinetraffic.com
www.mathway.com
www.midomi.com
www.pic2map.com

www.printfriendly.com
www.ted.com
www.tunefind.com/browse/tv
xxlmag.com
yahoo.com
zimbio.com

Appendix 2. Algorithm for finding the competitor.

1. Do a search for the full description. If there are a variety of offers (different models) and vendors in organic and sponsored results, skip to step 2.
1.A. If results are not varied enough (they repeat the same model), try to make the keywords less specific, one step at a time, by doing the following:
1.A.A. First eliminate any repetitive or redundant terms. Only these can be eliminated all at once since they do not affect the search results. All the others must be done one at a time
For example:
Bluetooth Speaker,MusiBaby M71 Speaker,Outdoor,Portable,Waterproof,Wireless Speaker,Bluetooth 5.0,Dual Pairing,Loud Stereo Booming Bass, 24H Playtime for Home,Party(Black) → Bluetooth 5.0 Speaker,MusiBaby M71 ,Outdoor,Portable,Waterproof,Wireless ,Dual Pairing,Loud Stereo Booming Bass, 24H Playtime for Home,Party(Black)
1.A.B. When possible, eliminate terms one at a time, going from the least related to function of the object to the most related. Usually brands are the least related to function, but not always.
For example: the term "2020 HP Laptop 15.6" FHD Micro-Edge WLED Computer, 10th Gen Intel Quad-Core i5-1035G1 (up to 3.6 GHz, Beat i7-7260U), 16GB RAM, 1TB SSD, Webcam, HDMI, WiFi 6, Bluetooth, Win10, w/GM Accessories" includes plenty of keywords that can be eliminated one at a time, and clearly "HP" and "Intel" are the least related to function since they are the just names of brands for which there are similar alternatives.
Another example: in "Pokemon trading cards, pack of 12", Pokemon is both the brand but descriptive of the object. So we would start with "pack of 12" instead since that is the term less related to function, since Pokemon is also the subject matter of the cards, and a Pokemon fan would probably be more interested in a 36 pack of cards instead than in a 12 pack of baseball cards.
1.A.C. If eliminating any term would make the search too generic, senseless, or eliminate the string completely, instead substitute for the most specific subtype. When this step is invoked, the RA informs that a review is required.
1.A.C.1. A different RA sees the product and the search term and tries to come up with a substitution independently.
1.A.C.2. If the two terms are similar (as measured by word vector similarity of 0.95 or greater), then the original substitution is used. If they are not, then they need to discuss and come up with consensus.
Note: If a search seems to show no offers, but only review sites, you can try adding the word "buy" before the term. For example: "buy graphic novels".
"Making the search too generic" means jumping from a search that produces highly specific results to a search that produces very generic results, when there are clearly intermediate steps that can be used to

specify a more similar range of products in the results. For example, with the term "Join or die" T-shirt, if we eliminate the term "Join or die", we jump from a specific T-shirt design right into generic t-shirts, when there are clearly other t-shirts that share the same topic. We therefore need to take a smaller step and instead go for similar topics. If we assume the "join or die" T-shirt is related to libertarian political causes (in this case it appeared on a libertarian website), we can use the word "libertarian" instead of "join or die" as an intermediate step and go for "libertarian t-shirts"

1.A.D. Every time you eliminate or substitute an item, Search again. If the variety is good enough, continue. Otherwise, go to step 1.A

2. The RA is given a randomly generated list with numbers 0 to 10 to determine which link in the search result to click first. 0 indicates the first (top-left) sponsored result. 1-10 indicate the respective organic results positions.

3. If the link is the same model as the ad, the page is not in english, does not sell to US customers in US dollars, or does not lead to an offer, use the next element on the list. Otherwise the result is used.

4. Keep going down the list in (2) until a good result is found.

5. In the landing page, select the highlighted offer or the offer in the top left if none is highlighted

6. Capture the price, description, image, brand and vendor that offer