The Welfare Loss of Information Manipulation on Consumers in Online Retail: Evidence from Incentive-Compatible Experiments*

Hajime Oiso, Hirohide Sakaguchi, Takanori Ida, and Toshifumi Kuroda

March 18, 2024

Abstract

Fake reviews distort not only information about a product but also search behavior in online retail. We conducted an online experiment to analyze the effects of fake reviews on product information and product search for three distinct product categories: digital cameras, headphones, and smartwatches. We find that fake reviews consistently direct consumers towards lower-quality products within all investigated categories. Moreover, we found that the average consumer faces significant search costs, leading them to examine only a limited selection of products in detail. Because fake reviews distort the perception of product quality and facilitate the inclusion of low-quality products in product searches, the negative impact of fake reviews on digital cameras is exacerbated when the products are displayed in order of the review ratings. Furthermore, when consumers are recommended to read the reviews carefully, it reduces the demand for goods, while it has no positive effect on the product quality.

JEL code: M30, D18, L81, D91, C91

Keywords: Fake reviews, search cost, platform, online retail

*Kuroda: Department of Economics, Tokyo Keizai University, 1-7-34, Minami-cho, Kokubunji-shi, Tokyo 185-8502, Japan. kuroda@tku.ac.jp. Acknowledgements: We gratefully acknowledge the financial support from JSPS KAKENHI grant numbers 22K01485 and Fukuzawa Fund. We are also grateful for the comments from participants of the JEA, JEMIOW, ABEF, and SWIE, and Katsunori Yamada and Shuhei Kitamura for their helpful comments. At the time of the research, Kuroda served as an Economist at the METI from 2021 to 2023. Oiso has served as an official for the Ministry of Internal Affairs and Communications, Japan (MIC) and other public bodies since 2004 but this research has been conducted under his other titles of Associate Professor (non-tenured) in the Faculty of Environment and Information Studies of Keio University until March 2024 and Senior Researcher of Keio Research Institute at SFC followingly. The views and opinions in this paper do not represent those of public bodies such as METI or MIC. AEA trial registration: AEARCTR-0011020. IRBs: Global Survivability Studies Unit, Center for the Promotion of Interdisciplinary Education and Research, Kyoto University. IRB Approval Date: 2023-03-13. IRB Approval Number: 202307.

1 Introduction

Digital platform operators, such as Amazon, match consumers with products not only by setting prices and product attributes but also by providing consumers with review information. Marketing literature (e.g., Kim, Albuquerque and Bronnenberg (2010)) and economics literature (e.g., Jolivet and Turon (2019)) have shown that certain display order of products on the website reduces consumers' search costs. On the other hand, firms can abuse such mechanisms so that consumers are affected by the ways in which information are presented, by inflating review ratings and presenting fake reviews, which then may reduce consumer welfare by misleading consumers into purchasing inferior products.

In March 2022, the Consumer Affairs Agency of Japan prohibited the fake review by the Act against Unjustifiable Premiums and Misleading Representations. The purpose of its regulation is similar to that of the European Digital Services Act (DSA), which came into effect in November 2022 to ensure autonomous and rational informed decision-making by consumers. In the United States, section 5 of the Federal Trade Commission (FTC) Act has long outlawed unfair or deceptive acts or practices in a comprehensive manner. While there is no significant difference among Japan, Europe, and the U.S. in terms of fake reviews being illegal, the effectiveness of the regulations is still in question.

To study the impact of different ways of displaying product information and from existence of fake reviews on consumer behavior in online retail, we conducted an online experiment with 6,205 participants with various treatments such as two types of review content, three types of product display orders, three types of review display orders, two interventions by presenting two-sided information and a warning message about the presence of incorrect or misleading reviews.

One type of review content we used was based on the reviews that have been deleted from an online retailer's site, while the other type is based on the reviews that have not been deleted. The online retailer's terms of service prescribe that incentivized reviews are to be deleted; therefore, it is assumed that the deleted reviews may contain fraudulent information manipulated through incentivization.

To examine the effects of the orders by which products are presented, we adopted three different orders: i) ratings based on evaluations from the reviews provided by users (product ratings), ii) the number of reviews, and iii) random as a control group, to sort the products. Both the orders by product rating and the number of reviews can be manipulated by companies if they distort those numbers, while in the random case, they cannot control the order of products in the product list.

We also sorted user reviews on products in three different orders of review ratings (number of stars labeled to each review), newest first, and random. Since the experimental site only displayed the top five reviews on product individual pages that are shown first when participants chose to view details, sorting reviews by review ratings, which reflects the informativeness of user reviews, may encourage consumers to read more reviews. Conversely, when reviews are displayed in order of the newest first, there may be more reviews that should eventually be deleted from the webpage than the other orders because it takes time for online mall operators to detect fake reviews. Moreover, if consumers are aware that reviews may be fake, they may read fewer reviews because the accuracy of reviews has declined. Conversely, if consumers are unaware, they are less likely to reduce the number of reviews they read and thus likely to be deceived by fake reviews.

In addition, we randomly assign participants to be exposed to the two-sided presentation of reviews, in which both positive and negative information are displayed simultaneously on the top of the review page. A similar technique to present both sides of information has been empirically studied in the field of persuasive communication in social psychology and has the potential to help consumers update their subjective beliefs on the accuracy of reviews.

To quantify the welfare loss of the fake reviews, we estimate the discrete choice model for product demand derived from the simultaneous search model. Using this demand function, we estimate willingness to pay (WTP) for the most desired product, which consumers revealed in the experiments. To estimate the true WTP, we offered respondents a one-time choice between two alternative draws, each with an equal probability of winning, after the consumer had selected the most desired products. One draw offered the most desired products, while the other offered 'points', which is a frequently adopted format of financial rewards to consumers in Japan, equivalent to a cash prize that we randomly assign an amount. This mechanism ensures that consumer behavior in online choice experiments is incentive-compatible.

We also collected data from participants on their online shopping experience, level of trust in online commerce, risk attitudes, information-seeking attitudes, participation environment, impressions of our experimental websites, and sociodemographic characteristics. In addition, we recorded click data from our experimental sites. Using this information, we will estimate the effects of information structure on consumer behavior, as well as the effects on consumer welfare, search costs, information processing costs, and prior for the product to obtain a more comprehensive understanding of the information design in the context of online retail services.

2 Literature

This paper follows the literature on search costs, product reviews, and two-sided presentation.

2.1 User review

In theoretical research, the issues of fake reviews and stealth marketing have been addressed by extending signaling models such as those of Nelson (1970) and Milgrom and Roberts (1986). Here, fake reviews are treated similarly to advertising, not only influencing consumer demand but also serving as a signal to consumers. Mayzlin (2006) finds that, in equilibrium, firms with lower-quality products engage in more stealth marketing. Moreover, the welfare loss due to stealth marketing depends on the difference in product quality and the number of consumers who can distinguish the quality of the former. Conversely, Dellarocas (2006) points to the possibility that both cases in which suppliers of high-quality products engage in more fraud and cases in which suppliers of low-quality products engage in more fraud can coexist in equilibrium. Which equilibrium emerges depends on the demand function, which is determined by the consumer's utility function for valuing expected quality. More recently, Yasui (2020) has conducted an analysis using a theoretical model that better reflects the characteristics of verification systems. In this model, firms dynamically adjust the number of fake reviews according to the quality of their products and their current ratings. Consumers anticipate the quality of products by considering these fake reviews generated by firms. In equilibrium, firms with higher quality products engage in more fraud, which serves as a signal of their high quality, resulting in fake reviews that benefit consumers. On the other hand, the presence of easily fooled naive consumers is also considered, and the paper mentions that regulating fake reviews may benefit these consumers.

A persistent problem in empirical research is the difficulty in determining whether a review is fake or not. To identify such fakes, Mayzlin, Dover and Chevalier (2014) and Luca and Zervas (2016) adopt a proxy approach. On the other hand, He et al. (2022) successfully identified products that hire fake reviews through a study of Facebook's fake review request group. Mayzlin, Dover and Chevalier (2014) use the difference in score distributions between Expedia and Tripadvisor, the two main review sites for the US hotel industry. They show that hotels owned by non-chain and independent owners, who have a greater incentive to cheat, are more likely to engage in fraudulent activities. Luca and Zervas (2016) find that fake reviews increase as review ratings decrease. In addition, they show that restaurants with the same type of competitors in the same area are more susceptible to attacks from low-rated fake reviews, while restaurants with chain restaurants in the same area are less likely to be targeted by such attacks. He et al. (2022) target Amazon.com and identify products that solicit fake reviews on social media for analysis. Rather than using a proxy, they identify products that actually engage in fraud, providing valuable insights not found in existing research. They find that fake reviews increase sales and that sellers of low-quality products are more likely to engage in fraudulent activity. In addition, they show that after fake reviews temporarily inflate ratings, the ratings experience a sharp decline, indicating the potential for real harm.

In recent years, studies have used experiments to estimate the impact of fake reviews and evaluate the effectiveness of platform interventions. Ananthakrishnan, Li and Smith (2020) show that a system that flags reviews as suspicious, rather than removing them outright when they are deemed fake, can increase consumer trust in the platform. Akesson et al. (2023) conducted experiments on a simulated e-commerce site and estimated that the welfare loss due to fake reviews is about 10 percent of consumer spending. They further show that educational interventions by the platform can reduce this welfare loss by about half. Hollenbeck (2023) investigates the fake product reviews on the reputation systems. They find that the consumers' mistrust of ratings increases price competition and partially offsets the welfare loss from misinformation. In addition, it is beneficial for the platform to remove all fake reviews along with consumer belief shifts.

2.2 Search cost

The search literature originally focused on the analysis of the distribution of prices, as exemplified by George J. Stigler (1961). More recently, Gavazza and Lizzeri (2021) have classified price search models into two categories: simultaneous search and sequential search. The former assumes that consumers fix the number of offers by paying a search cost, while the latter assumes that consumers decide whether to stop searching by weighing the additional cost against the expected benefits of obtaining more offers. Since the number of products per category is fixed in our experiment, our setting falls within the domain of sequential search models.

Product characteristics may also be subject to market friction. In their review of search models with product differentiation, Gavazza and Lizzeri (2021) discusses the literature on this topic, citing the work of Martin L. Weitzman (1979). To estimate WTP for products, we exclude price as a product attribute and instead ask participants to choose the best product for them. This implies that the design of our experiment falls under the product differentiation model.

The literature also examines the role of intermediaries in search markets. Gavazza and Lizzeri (2021) and Jullien, Pavan and Rysman (2021) review the relevant research on intermediaries and platforms in search markets. Our experimental setting emulates the design of online commerce platforms, which utilize not only information provided by suppliers but also information provided by other consumers. However, we have maintained the information of products unchanged during the experiment to gauge the impact of product information, rather than the network effect of platforms.

The empirical literature on search markets examines different types of markets. Kim, Albuquerque and Bronnenberg (2010) estimate an empirical model of sequential search proposed by Martin L. Weitzman (1979) using view-rank data for camcorder products. They find that consumers searching for camcorders on Amazon.com typically limit their search to 10-15 options. Counterfactual simulations of the effect of search costs show that reducing search costs results in 99% of consumers purchasing products of the same quality. The study concludes that most households benefit from Amazon's product recommendations due to reduced search costs. Jolivet and Turon (2019) also estimates an empirical model of sequential search using transaction data for CDs. They find that search costs are heterogeneous. Similarly, Gu and Wang (2022) estimates a sequential search model in which search costs depend on information complexity using panel tracking of consumer search behavior at an online travel agency. They find that cognitive costs account for a significant fraction of search costs, while load time costs contribute a small fraction. Our paper differs in that it acknowledges that online ranking data can be manipulated by producers. This implies that reducing search costs may not always increase consumer surplus, as consumers can easily be misled into purchasing inferior products. Fang et al. (2024) investigate the impact of a new search tool, which utilizes both visual and textual cues to suggest refined queries for general search terms. It reveals that the benefits of the search tool arise not only from direct suggestions but also from indirectly teaching customers to conduct more efficient searches on their own.

3 Research question

In this section, we describe the experimental hypothesis of this study following the literature of search cost, review, and two-sided presentation.

3.1 Basic behavior

User review Consumers can receive a signal of product quality from user reviews that influence consumer behavior in the marketplace. The online marketplace provides user reviews in the form of product ratings, number of reviews, and written language text. Both product ratings and number of reviews are numerical values, with the higher value indicating that the product is better. On the other hand, written language texts can be read by humans, but they need to be quantified in some way for quantitative analysis. However, it is possible to observe differences in consumer behaviors without quantifying the information of reviews if the types of reviews given to consumers are separated by treatment groups in an experiment rather than using observational data. Akesson et al. (2023) suggests that consumers can detect fake reviews by carefully reading them. Therefore, product reviews are expected to work by knowing the quality of the product and detecting fake reviews.

Search cost The literature on search costs shows that consumers do not receive all available information related to their choice. Kim, Albuquerque and Bronnenberg (2010) find that consumers typically limit their search to 10-15 options. Thus, consumers are unlikely to search for all products. Furthermore, due to the design of online commerce sites, it is expected that information about products displayed first will be more easily collected, while information about products displayed later6 will be less easily collected. In addition, Gu and Wang (2022) shows that the search cost depends on the complexity of the information. Regardless of the complexity of the information, because the information available on websites differs by product category, the search cost is expected to vary by product.

Based on the above hypotheses, we assume that consumers should respond to information presented on online shopping sites as follows.

1. Consumers are likely to click on a product that is higher on the list.
2. Consumers are likely to click on a product that has a higher user rating or a higher number of reviews.

We first check whether these behaviors are observed.

3.2 How should products be sorted and listed to maximize consumer benefits?

The economics literature shows that platforms sort products by product utility, and this practice is beneficial to consumers because it saves search costs to find the best products. However, our data scraped from an online retailer shows that fake reviews inflate the score or number of reviews of inferior products. As a result, the inferior products are likely to be shown at the earlier order in the product list and consumers are more likely to click on or purchase inferior products. To avoid inflated inferior products, consumers should pay attention to the details of product reviews. If there are no fake reviews, it may be best to sort products by review ratings or number of reviews. However, if there are fake reviews, sorting products by review ratings or number of reviews may not be the best, because the higher position may contain the inferior product. Therefore, the best product order for consumers depends on whether there are fake reviews, whether the product search cost is high or low, and whether the accuracy of product reviews is low or high.

3.3 Hypothesis

To investigate the best design for internet retail, we test the following hypothesis in the experiments. The primary interest of our experiment is the effect of fake reviews on consumers' choice behavior. We assign the fake review as treatment in our experiments.

Hypothesis 1 (H1) *Treatment group consumers are more likely than control group consumers to select the lower quality product.*

In order to understand the mechanism that creates the effect of fake reviews, we test the following hypotheses about the order of products as a secondary interest in this study.

Hypothesis 2 (H2) *Consumers tend to click on the products arranged vertically on a Web site that are arranged at the top. In addition, consumers can discover the 'best' product for them in fewer clicks when products are arranged in rating order or in number of review order than when products are arranged randomly.*

In addition, we will examine the effectiveness of providing information that encourages consumers to make better choices. We test a warning message that encourages consumers to refer not only to the number of reviews or the ratings but also to the text posted since reviews include fake reviews which consumers may spot by reading those texts.

Hypothesis 3 (H3) *The warning message that encourages the reader to read not only the number of reviews and their ratings but also the text of the reviews in order to spot fake reviews leads consumers to spot the fake reviews and results to choose the better goods. In addition, the impact of fake reviews can be mitigated by the warning messages.*

Finally, we will investigate the intersectional effect of fake reviews on the order of products.

Hypothesis 4 (H4) *The effect of fake reviews is exacerbated due to the order of products also manipulated by the fake review.*

4 Data

4.1 Review Data and Collection Process

Our dataset for the experiment was developed by referring to actual Amazon reviews. Although it is not possible to determine whether an Amazon review is fake or not, as shown in He et al. (2022), deleted reviews can be used as a proxy variable for fake reviews.

On Amazon, deleted reviews become unobservable. Therefore, if the data collection interval is long, it may be impossible to observe review deletions. In particular, if a review is posted and deleted during the unobserved period, the deletion will be unobservable. Therefore, it is necessary to collect product data at a high frequency.

The collection process is as follows. First, on a daily basis, we review the Amazon bestseller rankings and the first 20 pages of the search rankings, adding all products that appear at least once to our data collection targets[1]. At the same time, we collect the review pages of products in the data collection targets when they become available[2].

In this paper, we collected review data for three product categories (digital cameras, headphones, and smartwatches) by repeating these processes from January 2021 to September 2022. In addition, we obtained individual review deletion statuses in December 2022[3]. In the end, we collected data for 17,798 unique product IDs and a total of 2,183,009 reviews. After removing duplicates[4], there were 714,140 unique reviews. We collected metadata such as review ratings, posting timestamps, and review attributes such as review text and titles. For product data, we collected information about whether the manufacturer has a dedicated store on Amazon, seller information, and more. Product-level descriptive statistics are shown in Table 1. It shows average review ratings, number of reviews, removal rate, marketplace share, and average review ratings by removal status for each product category.

The disadvantage of our approach which uses deleted reviews as a proxy of

[1]Managed by product IDs on Amazon

[2]While rankings and search rankings are collected daily, reviews are collected continuously. In concrete, we retrieve the reviews of all targeted products in sequence, repeating the process as new product data becomes available.

[3]In this paper, we collect reviews at the product page level. Sometimes, even though a review is not deleted, it disappears from the review page. Therefore, we separately obtain the deletion status. A deletion is considered if the individual review page returns a 404 status code

[4]On Amazon, products with different variations, such as color or sets, are handled with distinct IDs. Reviews are not linked to a single product ID but rather to multiple variation products. As a result, when collecting reviews at the product level, duplicate reviews may occur.

Table 1: Summary statistics of the consumer reviews from Amazon.co.jp

	Headphone		Smart watch		Digital camera	
	Mean	Std	Mean	Std	Mean	Std
Average review ratings	3.95	0.72	3.61	1.02	4.14	0.59
Average review ratings(deleted)	4.26	0.83	4.31	0.98	4.12	1.01
Average review ratings(undeleted)	3.90	0.77	3.51	1.04	4.11	0.62
Number of reviews	196.95	412.28	51.08	141.74	58.35	74.54
Deletion rate	0.11	0.20	0.14	0.26	0.06	0.14
Market price	5.68	10.11	12.93	19.29	58.71	87.70
Marketplace rate	0.60	0.49	0.47	0.50	0.35	0.48
Number of products	8536		5183		4073	

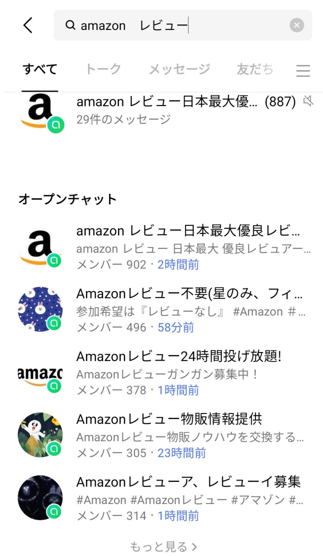
Note. The unit of Price is in 1,000 JPY which is about 6.62 UDS in November 2023.

fake reviews is that it cannot detect fake reviews without any review text. Such no-text fake reviews that increase the ratings and the number of reviews is not affected even if they are posted. The social networking sites commonly used[5] to solicit fake reviews in Japan not only solicit fake reviewers with review text but also those with rating only that our data could not include. Figure 1 shows a typical fake review community in Japan. In those communities, both rating only and rating with text are solicited.

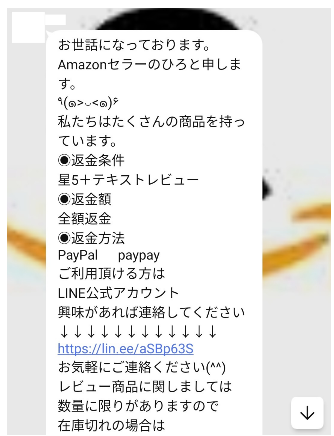
5 Experimental design

The experiments were conducted in March 2023. We intended to recruit 5,000 adult participants representing Internet users in Japan. To recruit these participants, we used the services of an Internet survey company. The company sent an email to potential participants asking about their willingness to participate in an online survey about online shopping. The email included details of the survey periods and also informed respondents that the survey was being conducted by academic researchers. In addition, respondents were informed that they would receive 250 JPY equivalent points for their participation and that a small number of respondents would receive the product they selected in the survey. Once potential participants agreed to participate, the survey company sent them an email containing the URL of the experimental website. The company continued to send these emails until

[5]https://www.nhk.or.jp/gendai/articles/4335/



(a) Examples of fake reviews hiring community in LINE



(b) Examples of hiring message in LINE

Figure 1: Examples of Japanese Fake Review Recruitment Communities

Note. The left figure is a screenshot of a search for the term ”amazon review” on the LINE mobile app. Various communities hiring fake reviews are listed as candidates, some of which require posting with review text and others require only rating. The right figure is the actual message for hiring fake review on line. They offer a cash refund of the full product price as a reward and solicit reviews with a 5 star rating and text.

the target number of responses, representative of Japanese Internet users by gender and age was reached.

The website first asked participants about their current situations such as their experience with online retail. Participants were then presented with pages that simulated an online retail store, where they were asked to select the most desired products or indicate that they did not want any products in a particular category. The product categories on these pages were digital cameras, headphones, and smartwatches, with 10 digital cameras, 30 headphones, and 30 smartwatches listed under each category. The order in which the categories are presented is randomized.

The simulated store pages consist of three levels: a product listing page, a product detail page, and a page displaying all reviews.

Figure 2 presents the first-level page that displays a list of products in a single category. The list included the product name, the average user rating (product rating), the total number of reviews, and a product image. In the case of a warning message assigned, the first-level page also shows the warning message which suggests consumers to carefully read the written review to avoid the lower quality products. The warning message is only shown in this page. The message is different from the warning of Akesson et al. (2023) in the content and frequency. Akesson et al. (2023) informs the manipulation that they made and the warnings are also displayed on the product selection page.

When respondents clicked the ”view product details” button, they were taken to the second level page for that product. Once respondents have viewed the second level page and returned to the first level page, a ”None of them are desired” button appears at the top of the product list. Clicking this button takes respondents to the confirmation page.

Figure 3 presents the second level page that provides detailed product specifications and five reviews for the product, each with a review rating. The page also contained three different types of buttons: the first was a ”Back to List” button that took respondents back to the first-level page. The second was a ”Choose It” button that took respondents to the confirmation page. The third button was a ”None of the products are desired” button. The fourth was a ”Review is helpful” button displayed below each review text. The fifth button was a ”See more reviews” button that took respondents to the third level page. The product detail pages can also present a two-sided view based on the respondent’s assigned treatment.

The third level page displayed ten reviews along with a two-sided presentation and has two types of buttons. The first was a ”back to the previous page” button that took respondents back to the second level page. The second type was the ”Review is helpful” button displayed below each review text.

When respondents reach the confirmation page, they are presented with a ”Confirm” button and a ”Back to Product Page” or ”Back to Product List Page” button,

Figure 2: First-level page with warnig message (In Japanese)

Note. The yellow framed area at the top of the page is the warning label. The warning informs the following messages.
On this site, you can select products by reading product evaluations (reviews) submitted by customers. As a rule, the reviews on this site are posted as they are by our customers. Therefore, they may include postings that do not necessarily conform to the facts or that may be misleading. To ***get more accurate information*** *about a product, it is effective to*

- *Referring not only to the "number of stars" or "number of ratings" but also to* ***the text of the review***
- *Refer to both* ***high*** *and* ***low*** *rated reviews.*
- *Be suspicious of reviews that contain* ***unnatural language****.*

Product photos and names are masked to avoid copyright infringement in this figure.

Figure 3: Second-level page (In Japanese)

Note. The second level page shows the photo, name, specs, rating, number of reviews, and three reviews of products. Photos, names, brands, model names, and styles are masked to avoid copyright infringement in this figure.

Figure 4: Third-level page with two-sided presentation (In Japanese)

Note. The third level page shows a photo, name, rating, number of reviews, and ten reviews of products. Photos and brands are masked to avoid copyright infringement in this figure.

depending on the previous page. Once respondents click the ”Confirm” button, they will not be able to navigate back to previous pages.

After respondents confirm their product selection, the site prompts them to enter a drawing for a chance to win a product or points equal to a randomly awarded cash prize. Respondents were informed that the number of points would be randomly assigned and that they would only have one chance to enter. This design will allow us to estimate respondents' WTP for the product in an incentive-compatible manner.

We randomly assigned participants to two information sources, three product orders, and three review orders, whether or not the second-level page included a two-sided presentation of reviews, and whether or not the first-level page included a warning about the presence of false or misleading reviews.

Following the three product choice experiments, we asked respondents to provide feedback on their overall impressions and experiences with the site, as well as their experiences and perceptions about online shopping and the Internet.

The survey included two traps to determine whether respondents had read the site description carefully. The first trap confirmed whether respondents had read a text explaining the experiments. Following Bottan and Perez-Truglia (2022), respondents were asked to select ”none of the above” from a list of 11 options. This trap was placed between the second and third categories of product choices.

The second trap confirmed whether respondents had ignored the price of products during the choice experiments on the most desired products. Respondents were asked to indicate which product attributes were important in their selection of products on the website, from a list of attributes that included ”price” which was not shown at all to the respondents in fact, as well as other attributes commonly described on product detail pages. This trap is presented after respondents have completed the three categories of product selection.

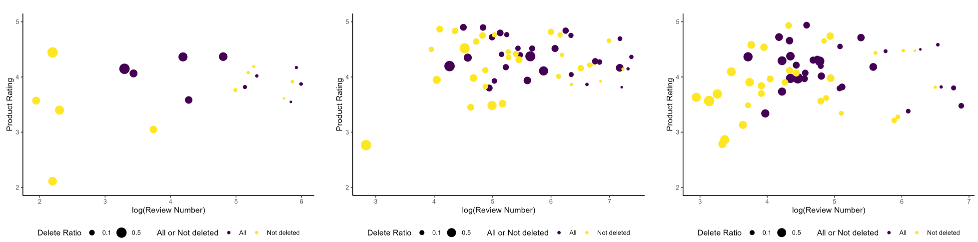
5.1 Product and review selection

In our experiments, we select 10 digital cameras, 30 headphones, and 30 smartwatches as the product list from our crawled data described in 4.1. We also create both 10 deleted and non-deleted base reviews based on only non-deleted or only deleted reviews for each product from the crawled data.

Product selection The products used in the experiment were selected to have a good distribution of ratings and review deletion rates, provided that they are currently available for product giveaways, are not too expensive (less than 50,000 JPY), and have enough reviews.

We first select products in each category that have both 15 deleted reviews and 15 not deleted reviews. Because some reviews are written in languages other than Japanese, some are difficult to rewrite based on them, and some are mixed with what appear to be reviews of other products, we limited the number of products with more reviews than the required minimum of 10. Then, the products were divided into 15 tiles based on five levels of consumer rating and three levels of review deletion rate by each category. Since there are not enough digital camera products that meet the above criteria, we randomly selected three headphones, three smartwatches, and only one digital camera product from each tile.

Figure 5 shows the correlation between the number of reviews and the consumer rating for all posted reviews and not deleted reviews for each product category. Excluding deleted reviews, the number of reviews must decrease, but the rating does not necessarily decrease. While there is no correlation between the number of reviews and ratings for all reviews, excluding deleted reviews creates a trend toward lower ratings for products with fewer reviews. On the other hand, for products with a high number of reviews, there appears to be no correlation between rating and the number of reviews, even when deleted reviews are excluded.



(a) Digital Camera (b) Headphone (c) Smartwatch

Figure 5: Ratings and number of reviews

Note. The size of the dots represents the deletion rate. The darker colors are for all reviews aggregated, and the lighter colors are for only those reviews that were not deleted.

Table 2 summarises the product ratings and number of reviews for each product category, with 'All' as the average of all reviews, including deleted reviews, and 'True' as the average excluding deleted reviews, market price as the price in the amazon.co.jp at the time of data construction, and the review deletion rate for each product.

Review selection Reviews for each product were divided into eight tiles based on two ratings, two review votes, and two review dates for the smartwatch and

Table 2: Product data

Category		Mean	SD	Min	Median	Max
Digital Camera	Rating (All)	3.996	0.289	3.548	4.042	4.369
	Rating (True)	3.613	0.665	2.111	3.687	4.444
	Review number (All)	180.900	144.130	27.000	146.000	401.000
	Review number (True)	125.800	130.884	7.000	94.500	351.000
	Market Price	18.212	14.924	4.290	13.160	44.579
	Delete ratio	0.186	0.176	0.029	0.135	0.519
Headphone	Rating (All)	4.361	0.332	3.800	4.357	4.900
	Rating (True)	4.224	0.492	2.765	4.279	4.867
	Review number (All)	527.467	475.013	71.000	287.500	1596.000
	Review number (True)	349.333	352.849	17.000	185.000	1381.000
	Market Price	4.986	5.599	0.980	3.749	30.113
	Delete ratio	0.212	0.171	0.036	0.177	0.732
Smartwatch	Rating (All)	4.162	0.405	3.340	4.206	4.939
	Rating (True)	3.880	0.557	2.786	3.871	4.933
	Review number (All)	236.567	260.107	41.000	117.000	977.000
	Review number (True)	141.200	161.776	19.000	73.000	663.000
	Market Price	4.066	1.040	1.999	4.189	5.999
	Delete ratio	0.276	0.175	0.036	0.255	0.733

Note. 'All' represents the average of all reviews, including deleted reviews, and 'True' represents the average excluding deleted reviews.

headphones. However, due to a limited number of reviews, the reviews for digital cameras were divided into three tiles based on three ratings. On the advice of legal experts, we refrained from quoting the original text of the reviews, which we obtained from Amazon. Instead, we used the GPT algorithm to convert each Amazon review into a separate sentence. We then reviewed the algorithm's output and eliminated any unnatural expressions or information that was not present in the input data. After this human refinement, we manually selected ten reviews for each product to encompass a wide range of ratings, review votes, review dates, and information contained therein.

Reducing the selection set Some of the products selected through the above process included products that were difficult to use in the experiment. For example, the headphones category included chargers and cables for headphones. In addition, some of the products that were in stock when the data was collected in January were not in stock when we selected products in February. Since there were still more products than the target number after excluding these products, we manually selected 10 products for digital cameras and 30 products for both headphones and smartwatches to include a wide range of ratings, deletion rates, and information from reviews.

5.2 Randomization

Randomization Method Randomization is performed in the site server using SHA1PRNG, a pseudorandom number generation (PRNG) algorithm. For the randomization unit 1) described below, adaptive randomization (biased coin design) based on Rosenberger (2002) is also used.

Unit Four randomizations were performed as follows:

1. Participants were randomly assigned to 36 groups for combinations of all treatments except the warning message.
2. Participants were randomly assigned to four groups to receive different amounts of money in the WTP question.
3. Participants were randomly assigned to receive a warning message or not.
4. Participants were randomly assigned to six groups to determine the order in which the three product categories would be displayed on the website.

For each participant, the first two randomizations were performed for each product category. This means that these two randomizations were performed three times for each participant.

On the other hand, the last two randomizations - which determined the warning message that would be effective across categories when displayed, and the order of the product categories - were performed only once for each participant.

5.3 Sample

We obtained 6,205 individual responses from above mentioned experimental website. However, there are some invalid responses that we can not analyze those data or should be excluded from our sample.

The website was designed in such a way that it was not possible to return to previous steps after moving to the next step, with setting the steps of the first question, each of the three categories of product selection, and the last question. However, some participants browsed the website in a particularly irregular way, such as opening multiple tabs of the website with one tab to move forward and the other tab to navigate back to the previous pages, in order to escape from our design to restrict move between steps. We detected respondents who acted in such an irregular way from the click logs as much as possible and excluded them from the analysis. In addition, 962 individuals did not complete the survey.

Furthermore, 1,267 individuals did not clear the first trap, which confirms whether respondents had read a text explaining the experiments. 996 individuals do not clear the second trap, which asks individuals to ignore the price. In total, 1,914 individuals failed to clear those two traps.

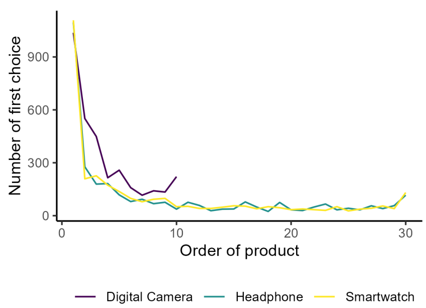
After removing those invalid responses, we obtained 3,167 individuals' responses from the website. We conducted the balance check which tests statistical differences between treatment and controls in those individuals' characteristics. We find there are no statistical differences between treatment and controls for all treatments.

6 Results

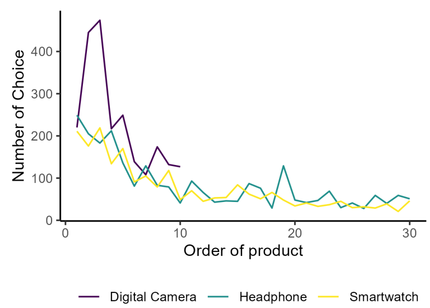
We first describe basic consumer behavior in the experimental sites, before describing willingness-to-pay amounts and treatment effects.

6.1 Consumer choice

Figure 6 shows the basic consumer behavior on the product list. The left side of the figure plots the order of the participant's first choice product on the horizontal axis and the number of views on the vertical axis. The right side of the figure shows, for each order on the product list page, how many participants selected the products in that order as their most desired product. The left side of the figure shows that the product order affects which product consumers tend to click the first. In addition, the right side of the figure shows that the last product on the list is more likely to be selected first, but the last product on the list is not more likely to be selected last. This suggests that consumers are scrolling to the bottom of the page to understand the choice set. These figures show that consumers look at the products they see immediately, but do not select the products they see immediately.



(a) First choice



(b) Final choice

Figure 6: Product order and choice

In order to see the impact of the fake review and the order on consumer behavior, figure 7 shows the number of final choices according to the order of product by whether the review contains fake reviews and the order of product over product categories.

The left line plots in the figure of each product category show how many participants selected the products in that order as their most desired product when the order of products was randomly assigned to each individual. It shows the tendency of consumers to click on the product at the top of the list, irrespective of the product attributes. In addition, this tendency is irrelevant to whether the review contains fake reviews or not.

Unlike when the order of products is randomly assigned to each individual, when products are ordered by the number of reviews or ratings, the position in which a particular product is placed is fixed depending on the order and whether the reviews contain fakes. Comparing the random order with the review number order or the rating order, consumers tend to choose the earlier-order products more. In addition, some products in the list are more attractive than other products regardless of whether the review contains fake reviews.

While consumers respond similarly to the order across product categories, there are differences across categories in the impact of the order on the choice. Among the three product categories, the relationship between the order and the choice is weakest for digital cameras, followed by headphones, and strongest for smartwatches. Although it is not easy to compare digital cameras with other categories because there are only 10 products rather than 30 products, the number of times that the product was selected in random order is nearly constant. On the other hand, in the cases of headphones and smartwatches, the product at the top of the list in random order was selected more than twice as often as those at the bottom of the list. The top-listed smartwatches were also more likely to be selected than headphones. This suggests that smartwatches have less heterogeneity among products than headphones, and there is less incentive to search for many products. It is also consistent with this speculation that consumers' choice of smartwatches is more concentrated in products that appear higher on the list than headphones when the sort order is in rating order.

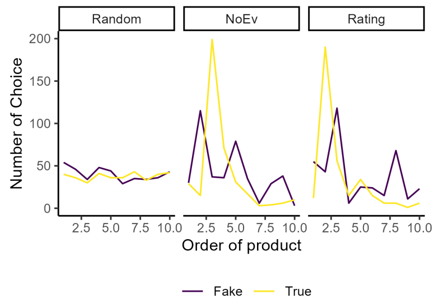
Figure 8 shows the basic behavior of consumers with respect to the product ratings and the number of reviews. The left side of the figure plots the product rating of the participant's first choice product on the horizontal axis and the number of participants on the vertical axis. The right side of the figure plots the product rating of the participant's last product choice on the horizontal axis and the number of participants on the vertical axis. The figure shows that consumers tend to choose products with higher ratings in both their initial and final choices. On the other hand, consumers show no clear trend in the relationship between the number of reviews and choices in both initial and final choices.

6.2 Reservation price

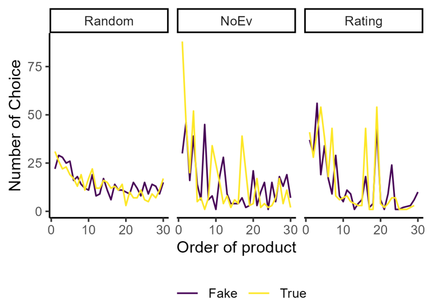
Figure 9 plots the amount of money offered to consumers on the vertical axis and the percentage of those who prefer the product on the horizontal axis for each category. In all product categories, consumers who are offered a higher amount of points tend to abandon the option of keeping the product as a random gift. The figure also shows that the OLS predicts that the majority of respondents have a positive reservation price for headphones and smartwatches, while only about 35% of respondents have a positive reservation price for digital cameras.

6.3 Effect of deleted review, product order, and warning message

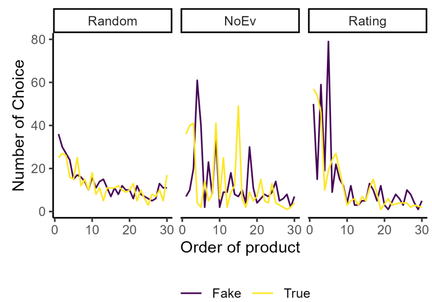
We first analyze the effect of deleted reviews, which represent the fake review treatment on their product choice. The effects are estimated OLS which regress a de-



(a) Digital Camera

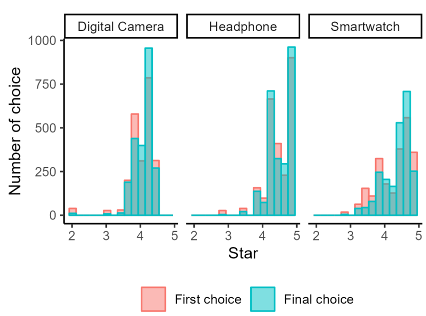


(b) Headphone

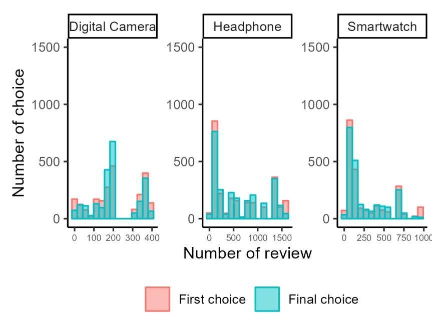


(c) Smartwatche

Figure 7: Product order and choice by orderings



(a) Product rating



(b) Number of review

Figure 8: Product rating and review

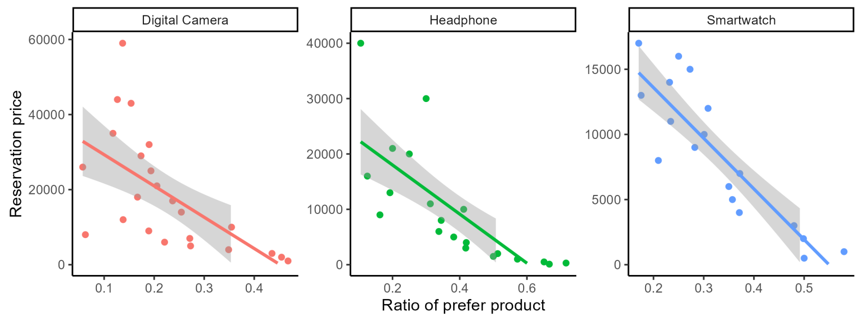


Figure 9: Choice between products or money

Note. Points are the ratio of respondents who choose a product for each amount of money. Straight lines are predicted by OLS with 95% confidential interval.

pendent variable on treatments, consumer characteristics, and intersections of fake review treatments with a warning treatment, ordering treatments, and two-sided presentation treatments. To focus on the effect of the fake review, the ordering, and the warning message, we omit the two-sided presentation treatments and consumer characteristics from the tables. The intersections of those treatments are described in section 6.4. As for the test statistics, we employ the Randomization inference p-values for random assignment to a treatment, which provides different results in the case of high-leverage conditions, such as intersections of treatments(Young, 2019). We employ the p-values less than 0.05 to be significant.

Digital Camera Table 3 shows the treatment effect of deleted review treatment, rating order treatment, and review number order treatment.

”Buy product” represents that consumers prefer to get some of the product, the order is the display order of the chosen product, with value increasing from top to bottom, the true star is the average product rating calculated by undeleted reviews for the product, ”True Num of Review” is the total number of undeleted reviews for the product, and the ”Market Price” which represents the price of the chosen product in amazon.co.jp at the time of data construction, which is not displayed in the experiments. Because the attributes of products are only observed when consumers choose any products they want, consumers who do not want any products are excluded from the regression of treatment effects other than ”Buy Product”.

Deleted reviews make consumers increase demand for products and choose a product that is positioned early in the order, a lower true star, a lower true number of reviews, and a lower price. The randomization inference shows a significant effect of the deleted review on the buy product, order, and true star, while the standard error of the regression coefficient does not necessarily reject the null hypothesis.

The warning message reduces the demand for digital cameras. However, it does not increase the true star.

Both the rating order and the review number order make consumers choose more highly ranked products, while this does not affect the true star or the true number of reviews.

The above results show that fake reviews increase consumer demand in the selection of digital cameras, causing them to select lower quality, lower priced products that are displayed higher in the list compared to the source of true information. Furthermore, the order did not have the effect of increasing the selection of better products.

Table 3: Treatment effect on the choice of digital cameras

	Buy Product	Order	True Star	True Num of Review	Market Price
Deleted review	0.084	-0.287	-0.315	-38.203	-9.337
	(0.035)	(0.236)	(0.046)	(8.287)	(1.284)
	{0.000}	{0.005}	{0.000}	{0.000}	{0.000}
Rating order	0.024	-2.526	0.001	-5.668	-0.062
	(0.028)	(0.189)	(0.037)	(6.641)	(1.029)
	{0.155}	{0.000}	{0.962}	{0.165}	{0.921}
Review Num order	0.027	-2.060	0.025	5.190	0.112
	(0.027)	(0.185)	(0.036)	(6.504)	(1.008)
	{0.103}	{0.000}	{0.238}	{0.191}	{0.860}
Warning message	-0.051	-0.103	0.005	1.413	0.116
	(0.023)	(0.154)	(0.030)	(5.401)	(0.837)
	{0.003}	{0.330}	{0.853}	{0.692}	{0.860}
N	3281	2285	2285	2285	2285
R squared	0.050	0.115	0.145	0.084	0.141
F Statistic	7.079	12.256	16.027	8.585	15.463

Note. Standard errors are in parentheses. Randomization inference p-values for random assignment to a treatment are in curly brackets. The unit of price is in 1000 JPY which is about 6.62 UDS in November 2023. Other treatment effects, individual characteristics, and product fixed effects are omitted.

Headphone Table 5 shows the treatment effect on the choice of headphones. Deleted reviews make consumers increase demand for products and choose a product that is positioned later in the order, a lower true star, a lower true number of reviews, and a lower price. The randomization inference shows a significant effect on all five measures, while the standard error of the regression coefficient does not always reject the null hypothesis.

The warning message also decreases the demand for products.

Rating orders increase the demand for products and choose a product that is positioned early in the order, higher true star, and lower prices. While review number order commonly increases product demand and choose a product that is positioned early in the order, true star decreases, the true number of reviews and market prices increase.

Other than the order, the effect of deleted reviews is qualitatively the same with the digital camera, while the size of effects on the demand and the true star is larger than the digital camera. The effect of the warning message on the headphones is the same as the effect on the digital camera. Unlike digital cameras, product order significantly affects product quality. In the rating order, better products with higher true ratings are purchased, while in the review count tour, the respondents are directed to higher products with lower true ratings. The difference between the rating order and the number of reviews needs to be further examined through the nature of the supply side.

Smartwatch Table 7 shows the treatment effect on the choice of the smartwatch. Deleted reviews make consumers increase demand for products and choose a product that is positioned early in the order, a lower true star, a lower true number of reviews, and a lower price. The randomization inference shows a significant effect on all five measures, while the standard error of the regression coefficient does not always reject the null hypothesis. The warning message decreases the demand for products. It is the same with other product categories. Rating orders decrease the demand for products and choose a product that is positioned early in the order, higher true star, lower true number of reviews, and lower prices. Review number order does not affect product demand and true star, but commonly choose a product that is positioned early in the order. The true number of reviews decreases, unlike the rating order.

The effect of deleted reviews and the warning message on the smartwatch choice is qualitatively the same as the digital camera choice. The size of the effects is mixed.

Summary of results Based on the above estimation results, the following conclusions can be drawn about the previous hypothesis. The first hypothesis which predicts that fake reviews lead consumers to lower-quality products is supported in all categories. The impact of fake reviews is strong enough not only to cause

Table 5: Treatment effect on the choice of headphones

	Buy Product	Order	True Star	True Num of Review	Market Price
Deleted review	0.026	1.026	-0.051	-96.424	-0.813
	(0.031)	(0.770)	(0.033)	(39.239)	(0.442)
	{0.042}	{0.001}	{0.000}	{0.000}	{0.001}
Rating order	0.030	-3.371	0.103	-33.779	-0.808
	(0.024)	(0.598)	(0.026)	(30.496)	(0.344)
	{0.042}	{0.000}	{0.000}	{0.077}	{0.000}
Review Num order	0.005	-2.891	-0.040	139.913	0.992
	(0.024)	(0.606)	(0.026)	(30.874)	(0.348)
	{0.739}	{0.000}	{0.009}	{0.000}	{0.000}
Warning message	-0.030	-0.534	0.014	-13.343	0.033
	(0.020)	(0.491)	(0.021)	(25.006)	(0.282)
	{0.039}	{0.136}	{0.339}	{0.446}	{0.867}
N	3223	2526	2526	2526	2526
R squared	0.087	0.037	0.046	0.090	0.051
F Statistic	12.636	4.010	5.025	10.305	5.638

Note. Standard errors are in parentheses. Randomization inference p-values for random assignment to a treatment are in curly brackets. The unit of Price is in 1,000 JPY which is about 6.62 UDS in November 2023. Other treatment effects, individual characteristics, and product fixed effects are omitted.

Table 7: Treatment effect on the choice of smartwatches

	Buy Product	Order	True Star	True Num of Review	Market Price
Deleted review	0.055	-0.887	-0.158	-16.474	-0.196
	(0.036)	(0.767)	(0.046)	(17.308)	(0.093)
	{0.000}	{0.012}	{0.000}	{0.026}	{0.000}
Rating order	-0.033	-5.369	0.099	-22.558	0.001
	(0.027)	(0.607)	(0.036)	(13.695)	(0.073)
	{0.044}	{0.000}	{0.000}	{0.003}	{0.985}
Review Num order	-0.007	-2.064	-0.014	24.585	-0.028
	(0.027)	(0.595)	(0.035)	(13.418)	(0.072)
	{0.643}	{0.000}	{0.524}	{0.005}	{0.521}
Warning message	-0.035	-0.571	0.034	-4.632	0.042
	(0.022)	(0.494)	(0.029)	(11.150)	(0.060)
	{0.031}	{0.109}	{0.102}	{0.523}	{0.319}
N	3219	2272	2272	2272	2272
R squared	0.046	0.064	0.053	0.068	0.037
F Statistic	6.479	6.399	5.257	6.830	3.638

Note. Standard errors are in parentheses. Randomization inference p-values for random assignment to a treatment are in curly brackets. The unit of Price is in 1,000 JPY which is about 6.62 UDS in November 2023. Other treatment effects, individual characteristics, and product fixed effects are omitted.

substitution between products but also to cause consumers to want categories of products that they would not want if given true information.

Regarding the secondary interest, both the rating order and the review number order make consumers choose early in the order of products in all three categories. Therefore, we confirm the second hypothesis for all product categories.

Regarding the third hypothesis, the effect of warning messages does not lead consumers to higher-quality products in all categories. Furthermore, the fact that the warning message reduces the demand for categories implies that the online platform may not want to provide the warning message to inform the consumers. The mitigation effect of the warning messages are investigated in section 6.4.

6.4 Intersections of treatment effects

The mitigation effect of the warning message in our hypotheses 3 and 4 is about the combination of fake reviews with product ordering and warning labels on the product quality investigated by the intersection of the treatments. Table 9 shows the intersections of treatment effects with the warning message and the product orders on the true star. The first three rows are the intersection of deleted reviews and warning messages. The second three rows and the third three rows are the intersection of deleted review and two types of ordering, respectively.

In the case of digital cameras and smartwatches, table 9 shows the intersection of warning signs and deleted reviews has negative significant impacts on true stars. This indicates that the effect of deleted reviews is exacerbated rather than mitigated by the warning message in the case of digital cameras and smartwatches. Furthermore, in the case of headphones, the effect of the intersection of warning messages and deleted reviews does not significant effect on the true star.

The results of this experiment are different from the findings of Akesson et al. (2023). The first difference is the content of the warning label. Akesson et al. (2023) note the characteristics they used to alter the reviews as characteristics of fake reviews, whereas we note not only the score and number of ratings of the reviews but also reading the text, reading both high and low scoring reviews, finding unnatural expressions, which apply to our data among their warnings. Another difference is the difference between reviews presented as fake reviews. Akesson et al. (2023) create fake reviews that have the characteristics they recognize as characteristic of fake reviews, whereas our review uses reviews that Amazon has removed as probable fake reviews. As far as we could see during the creation of the reviews, several had characteristics as fake reviews included in the warning messages, but most of the deleted reviews were naturalistic in their presentation. On the other hand, the deleted reviews often contained positive evaluations of the features included in the product name, and the analyst was able to recognize this fact by reading many of the reviews, knowing whether they were deleted or undeleted reviews. The difference between deleted and undeleted reviews would not have been recognized without knowing which was the review being read.

Regarding the hypotheses 4, the intersection of deleted review and rating order has negative significant coefficients in the case of the digital camera. Conversely, the intersection of deleted review and review number order has positive significant coefficients in the case of the smartwatch. Thus, the findings present a mixed result for Hypothesis 4.

Table 9: Intersectional effect on the product quality

Treatment	Digital Camera	Headphone	Smartwatch
Deleted * Rating order	-0.076	-0.013	0.049
	(0.051)	(0.036)	(0.050)
	{0.009}	{0.937}	{0.035}
Deleted * Review Num order	0.009	0.014	0.060
	(0.050)	(0.037)	(0.049)
	{0.008}	{0.500}	{0.062}
Deleted * Warning message	-0.063	0.002	-0.050
	(0.041)	(0.030)	(0.041)
	{0.707}	{0.493}	{0.026}
N	2285	2526	2272
R squared	0.145	0.046	0.053
F Statistic	16.027	5.025	5.257

Note. Standard errors are in parentheses. Randomization inference p-values for random assignment to a treatment are in curly brackets. Other treatment effects, individual characteristics, and product fixed effects are omitted.

6.5 Estimate willingness to pay

We estimated WTP for products by using an incentive-compatible questionnaire. The survey asked consumers whether they preferred the most desired product or a randomly selected amount of points equivalent to cash. In other words, it asks $u_j \succ u_p$ where u_j is the utility of the most preferred product and u_p is the utility from the points. It can be transformed to $u_j - u_p \succ 0$, which is equivalent to the consumer purchase behavior when the price is equal to the amount of money of points. Therefore, we can estimate the demand for products by a discrete choice model and calculate the WTP for each product by dividing the utility coefficient by point coefficients.

We use observed choice data as panel data for the first period in which consumers choose between products or "Not buy" as the outside option and the second period in which they choose between the product that they chose in the first period and the points that we randomly assigned. We also pooled three product category choices as a panel to maintain the individual coefficients between product categories.

In the first period, the consumer faces the indirect utility of products $v_j(T)$ and the search cost $c_i(T)$. Regarding $v_j(T)$, the consumer observes the characteristics of the goods, which depends on the products j and the assignment of treatments T. The consumer also faces the search cost $c_i(T)$. It's important to note that although consumers must pay search costs certainly, products are obtained by lottery. It makes the product search cost in this experiment the inverse of the subjective probability of winning p_i times higher than the search cost in the actual product search.

We assume that the simultaneous search behavior of consumers is as follows. At the beginning of period 1, consumers set the number of products to see the product detail page. Consumers click through the product pages in order, starting with the product displayed at the top. This assumption is consistent with the observation which was shown in the figure 6. Hence, the formula of search cost can be $c_i(T) = \beta_{c,i}\text{order}_j(T)$, where T is assigned treatment, $\beta_{c,i}$ is individual coefficients of search cost, and $\text{order}_j(T)$ is the order of product j under the treatment T. We assumed a normal distribution for the distribution of search costs because some consumers may prefer to search for many products. If consumers' search cost is too high to investigate any goods, they prefer to choose "Not buy". We assume the indirect utility of "Not buy" as reference level 0. Consumers who chose "Not buy" will either move on to select another product category after completing the first period or finish selecting the last product category and terminate the choice experiment. The consumer i's utility of choosing product j in period 1 as follows:

$$u(T)_{1ij} = p_i v_j(T) - \beta_{c,i}\text{order}_j(T) + e_{1ij} \tag{1}$$

where e_{1ij} is independently, identically distributed (iid) extreme value.

In the second period, the consumer faces the choice between the product chosen in the first period and randomly assigned points which can be used as equivalent amounts of Japanese yen at various merchants. In this period, consumers' indirect utility of goods is $v_j(T)$ and of points are $\beta_p\text{point}_{i,j}$, where β_p is almost equivalent with the indirect utility of money and $\text{point}_{i,j}$ is randomly assigned amount of points to consumer i who choose product j. In this period, search cost is irrelevant to the consumer choice. Hence, the consumer i's utility of choosing product j in period 2 as follows:

$$u(T)_{2ij} = v_j(T) + e_{2ij} \text{ for product } j \quad (2)$$

$$u_{2i,\text{point}} = \beta_p \text{point}_{i,j} + e_{2i,\text{point}} \text{ for point} \quad (3)$$

Hence, the choice probability $s_i j$ of consumer i choosing product j under treatment T is as follows:

$$s(T)_{ij} = s_{1ij} \times s_{2ij} \text{ where } \begin{cases} s_{1ij} = \dfrac{p_i v_j(T) - \beta_{c,i}\text{order}_j(T)}{\sum_l \left(p_i v_l(T) - \beta_{c,i}\text{order}_l(T)\right)} \\ s_{2ij} = \dfrac{v_j(T)}{v_l(T) - \beta_p \text{point}_{i,j}} \end{cases} \quad (4)$$

The estimated parameters of the random coefficient logit model are shown in table 11. Positive coefficients of 'Point' indicate that consumers are more likely to choose points when offered larger points. The negative coefficients of 'Order' indicate that consumers face significant search costs, meaning they are less likely to choose the product indicated below. The amount of search cost evaluated by points is about 721 JPY/Order(4.76 USD/Order). It indicates consumers face high search costs compared with the average market price of the products, which is 18,200 JPY for digital cameras, 4,990 JPY for headphones, and 4,070 JPY for smartwatches, respectively. Even though search costs are estimated to be higher than product utility, it is not orders of magnitude larger than in Kim, Albuquerque and Bronnenberg (2010) that finds about 30% consumers search less than 5 products in online commerce by using a sequential search model. In the case of the online travel agency, Ursu (2018) also finds the comparable search cost from the position is 1.85 USD to 3.73 USD by using a sequential search model.

Other than the point and the search cost, we estimate the effects of product attributes and the effects of treatments. Regarding the product attributes, Both the product rating and the number of reviews have positive effects on the consumers' choice.

Regarding the treatment effects, the fake review treatments increase the demand for products and the warning message decreases the demand for products, which is consistent with the result of OLS in section . Because the order and review number are controlled as the product characteristics, the coefficients of rating order and review number order are not significant. It indicates that the order, the rating, or the number of reviews itself is important for consumers, but what order the product is sorted does not affect product demand.

Regarding the distribution of the parameters, the inside goods dummy has significant distribution, which consumer choice does not maintain IIA property between the choice of goods and ”Not buy”. There is also significant distribution in the search cost.

Table 11: Demand parameters

	(1)
Point	0.068 (0.003)
Product order	−0.049 (0.002)
Rating	1.529 (0.097)
Review number	0.001 (0.000)
Deleted Review	1.569 (0.342)
Warning message	−0.095 (0.042)
Two-sided presentation	−0.030 (0.042)
Rating Order	0.050 (0.052)
Review Num Order	0.030 (0.052)
Review order: Date	−0.015 (0.052)
Review order: Useful	−0.001 (0.052)
SD: Inside Products	2.790 (0.115)
SD: Product order	0.053 (0.004)
Num.Obs.	1 184 904
AIC	54 192.7

Note. The unit of price is 1,000 JPY which is about 6.62 UDS in November 2023. Standard errors are in parentheses. Product fixed effects are omitted. The units of observation are individuals and alternatives (products, not buy, and points).

6.5.1 WTP estimates

Table 12 summarizes the estimates of WTP for products by category. As the figure 9 suggests, many individuals are predicted to have negative WTPs for products. However, after removing the search cost from the indirect utility of goods, the WTP for products is positive on average for all categories. The average WTP for digital cameras is lower than the average market price of digital cameras 18,200 JPY, while the average WTP for headphones and smartwatches is higher than the average market price 4,990 JPY and 4,070 JPY, respectively.

7 Welfare loss of the fake reviews

We evaluate the welfare loss of fake reviews from three sources. The first source of welfare loss is the loss due to the misperception that the product quality is inflated due to fake reviews. We estimate this loss as the monetary value of the difference

Table 12: Willingness to pay for product by category

Category	mean	SD	min	median	max
Digital Camera	13.104	19.883	−89.913	12.607	46.632
Headphone	9.992	20.704	−82.420	11.530	38.232
Smartwatch	7.081	17.095	−84.279	5.341	34.604

Note. The unit of observations is individuals. The unit of WTP is in 1,000 JPY which is about 6.62 UDS in November 2023.

between the true rating and the inflated rating of the product selected by the consumer to whom the false review was assigned. The second loss is the loss from the distortion of product choice which the fake reviews make consumers choose the wrong product. We estimate the choice under counterfactual situations when consumers assigned fake reviews are presented with true information and determine the difference in consumer surplus from the observed choice. The last source of the welfare loss is the increase in search costs to find the best product for individuals.

Table 13 shows the average and the standard deviation of estimates of three sources of welfare loss. The average welfare loss from the low quality is 5,800 JPY, 1,570 JPY, and 3,020 JPY (38.4 USD, 10.6 USD, and 19.9 USD in November 2023), respectively. The average welfare loss from the wrong products is 8,270 JPY, 3,060 JPY, and 4,600 JPY (54.7 USD, 20.2 USD, and 30.4 USD in November 2023), respectively. In contrast with the first two welfare losses, the loss from the search cost is negligible.

Comparing the source of welfare loss, the loss from choosing the wrong products is higher than the effect of low quality. The ability to estimate not only the intensive margin due to changes in product quality but also the extensive margin due to product substitution by estimating the demand function is a strength of this experiment, which directly measures willingness to pay. This result demonstrates the importance of including the extensive margin in the welfare evaluation of fake reviews.

Comparing the categories, despite the value of the star is the same between categories, the welfare loss from the fake review is highest in the case of digital cameras. It depends on the difference between the product star of deleted reviews and that of not-deleted reviews in the experiments. It implies that the external validity of this difference may not hold if the product is drawn in another way.

Table 13: Welfare loss of the fake review

		Digital Camera	Headphone	Smartwatch
Low quality	mean	5.803	1.566	3.015
	sd	0.698	0.262	0.313
Wrong products	mean	8.274	3.062	4.600
	sd	0.062	0.168	0.619
Search cost	mean	0.065	−0.096	−0.087
	sd	0.182	0.605	0.585

Note. The unit of welfare loss is in 1,000 JPY which is about 6.62 UDS in November 2023.

8 Conclusion

We conducted a study to examine the impact of deleted reviews and product orders in an online marketplace. We found that the average consumer faces significant search costs, leading them to examine only a limited selection of products in detail. Because fake reviews distort the perception of product quality and facilitate the inclusion of low-quality products in product searches, the negative impact of fake reviews on digital cameras is exacerbated when the products are displayed in order of the review ratings. Furthermore, when consumers are recommended to read the reviews carefully, it reduces the demand for goods, while it has no positive effect on the product quality.

References

Akesson, Jesper, W. Robert Hahn, D. Robert Metcalfe, and Manuel Monti-Nussbaum. 2023. "The Impact of Fake Reviews on Demand and Welfare."

Ananthakrishnan, Uttara M, Beibei Li, and Michael D Smith. 2020. "A Tangled Web: Should Online Review Portals Display Fraudulent Reviews?" Information systems research, 31(3): 950–971.

Bottan, Nicolas L., and Ricardo Perez-Truglia. 2022. "Choosing Your Pond: Location Choices and Relative Income." Review of Economics and Statistics, 104(5): 1010–1027.

Dellarocas, Chrysanthos. 2006. "Strategic manipulation of internet opinion forums: Implications for consumers and firms." Management Science, 52(10): 1577–1593.

Fang, Lu, Yanyou Chen, Chiara Farronato, Zhe Yuan, and Yitong Wang. 2024. "Platform Information Provision and Consumer Search: A Field Experiment." SSRN Electronic Journal.

Gavazza, Alessandro, and Alessandro Lizzeri. 2021. Frictions in product markets. Vol. 4, Elsevier B.V.

George J. Stigler. 1961. "The Economics of Information." Journal of Political Economy, 69(3): 213–225.

Gu, Chris, and Yike Wang. 2022. "Consumer Online Search with Partially Revealed Information." Management Science, 68(6): 4215–4235.

He, Sherry, Brett Hollenbeck, Gijs Overgoor, Davide Proserpio, and Ali Tosyali. 2022. "Detecting Fake Review Buyers Using Network Structure: Direct Evidence from Amazon." SSRN Electronic Journal, 1–26.

Hollenbeck, Brett. 2023. "Misinformation and Mistrust : The Equilibrium Effects of Fake Reviews on Amazon . com." 1–60.

Jolivet, Grégory, and Hélène Turon. 2019. "Consumer search costs and preferences on the internet." Review of Economic Studies, 86(3): 1258–1300.

Jullien, Bruno, Alessandro Pavan, and Marc Rysman. 2021. Chapter 7 - Two-sided markets, pricing, and network effects. Vol. 4, Elsevier B.V.

Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg. 2010. "Online demand under limited consumer search." Marketing Science, 29(6): 1001–1023.

Luca, Michael, and Georgios Zervas. 2016. "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud." https://doi-org.kras.lib.keio.ac.jp/10.1287/mnsc.2015.2304, 62(12): 3412–3427.

Martin L. Weitzman. 1979. "Optimal Search for the Best Alternative." Econometrica, 47(3): 641–654.

Mayzlin, Dina. 2006. "Promotional chat on the internet." Marketing Science, 25(2): 155–163.

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier. 2014. "Promotional Reviews: An Empirical Investigation of Online Review Manipulation." American Economic Review, 104(8): 2421–55.

Milgrom, Paul, and John Roberts. 1986. “Price and Advertising Signals of Product Quality.” https://doi.org/10.1086/261408, 94(4): 796–821.

Nelson, Phillip. 1970. “Information and Consumer Behavior.” https://doi.org/10.1086/259630, 78(2): 311–329.

Rosenberger, William F. 2002. Randomization in clinical trials theory and practice. Wiley series in probability and statistics, New York:Wiley.

Ursu, Raluca M. 2018. “The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions.” Marketing Science, 37(4): 530–552.

Yasui, Yuta. 2020. “Controlling Fake Reviews.” SSRN Electronic Journal, , (2018): 1–39.

Young, Alwyn. 2019. “Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results.” Quarterly Journal of Economics, 134(2): 557–598.

A Balance check

Table 14: TRfake

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRwarningmessage	0.508	0.500	0.511	0.500	0.003	0.010
TRproductorderRating	0.324	0.468	0.331	0.471	0.008	0.010
TRproductorderNoEv	0.334	0.472	0.334	0.472	0.000	0.010
TRrevieworderDate	0.327	0.469	0.332	0.471	0.006	0.010
TRrevieworderUseful	0.336	0.473	0.339	0.474	0.003	0.010
TRtwosided	0.499	0.500	0.500	0.500	0.001	0.010
character_female	0.524	0.499	0.523	0.500	-0.001	0.010
character_age	48.176	15.060	48.451	14.996	0.275	0.305
character_individualincome	2.454	2.747	2.523	2.710	0.069	0.055
character_education_year	5.758	1.923	5.767	1.911	0.009	0.039
D_character_experienced	0.158	0.365	0.154	0.361	-0.004	0.007
H_character_experienced	0.411	0.492	0.396	0.489	-0.015	0.010
S_character_experienced	0.114	0.318	0.114	0.318	0.000	0.006
character_amazonuser	0.808	0.394	0.796	0.403	-0.011	0.008
character_shopno	2.966	1.487	2.958	1.470	-0.008	0.030
D_character_categoryknowledge	0.078	0.269	0.074	0.262	-0.004	0.005
H_character_categoryknowledge	0.116	0.321	0.115	0.319	-0.001	0.006
S_character_categoryknowledge	0.046	0.208	0.047	0.212	0.002	0.004
D_character_categoryinterested	0.045	0.208	0.048	0.215	0.003	0.004
H_character_categoryinterested	0.116	0.321	0.109	0.312	-0.007	0.006
S_character_categoryinterested	0.091	0.288	0.096	0.295	0.005	0.006
character_aware_stealth	0.296	0.457	0.292	0.455	-0.004	0.009
character_aware_inflated	0.418	0.493	0.415	0.493	-0.003	0.010
character_aware_deflated	0.128	0.334	0.123	0.329	-0.005	0.007
character_aware_contaminated	0.236	0.425	0.220	0.414	-0.016	0.009

Table 15: TRproductorderRating

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRfake	0.497	0.500	0.509	0.500	0.011	0.012
TRwarningmessage	0.505	0.500	0.518	0.500	0.012	0.012
TRproductorderNoEv	0.000	0.000	0.000	0.000	0.000	0.000
TRrevieworderDate	0.338	0.473	0.325	0.469	-0.013	0.012
TRrevieworderUseful	0.325	0.468	0.346	0.476	0.021	0.012
TRtwosided	0.497	0.500	0.502	0.500	0.005	0.012
character_female	0.526	0.499	0.515	0.500	-0.011	0.012
character_age	48.362	15.132	48.060	14.805	-0.302	0.372
character_individualincome	2.462	2.715	2.517	2.708	0.055	0.067
character_education_year	5.767	1.923	5.774	1.923	0.007	0.048
D_character_experienced	0.147	0.354	0.161	0.368	0.014	0.009
H_character_experienced	0.399	0.490	0.411	0.492	0.012	0.012
S_character_experienced	0.118	0.323	0.112	0.316	-0.006	0.008
character_amazonuser	0.802	0.399	0.812	0.391	0.010	0.010
character_shopno	2.977	1.469	2.958	1.481	-0.019	0.037
D_character_categoryknowledge	0.077	0.266	0.075	0.263	-0.002	0.007
H_character_categoryknowledge	0.109	0.311	0.116	0.320	0.007	0.008
S_character_categoryknowledge	0.050	0.218	0.042	0.201	-0.007	0.005
D_character_categoryinterested	0.046	0.209	0.053	0.224	0.007	0.005
H_character_categoryinterested	0.111	0.314	0.118	0.323	0.007	0.008
S_character_categoryinterested	0.092	0.290	0.092	0.289	-0.001	0.007
character_aware_stealth	0.289	0.453	0.304	0.460	0.014	0.011
character_aware_inflated	0.423	0.494	0.423	0.494	0.001	0.012
character_aware_deflated	0.124	0.329	0.126	0.332	0.003	0.008
character_aware_contaminated	0.227	0.419	0.228	0.420	0.001	0.010

Table 16: TRproductorderNoEv

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRfake	0.497	0.500	0.503	0.500	0.005	0.012
TRwarningmessage	0.505	0.500	0.506	0.500	0.001	0.012
TRproductorderRating	0.000	0.000	0.000	0.000	0.000	0.000
TRrevieworderDate	0.338	0.473	0.325	0.468	-0.013	0.012
TRrevieworderUseful	0.325	0.468	0.343	0.475	0.018	0.012
TRtwosided	0.497	0.500	0.500	0.500	0.003	0.012
character_female	0.526	0.499	0.530	0.499	0.004	0.012
character_age	48.362	15.132	48.515	15.139	0.153	0.374
character_individualincome	2.462	2.715	2.488	2.761	0.026	0.068
character_education_year	5.767	1.923	5.747	1.905	-0.020	0.047
D_character_experienced	0.147	0.354	0.160	0.366	0.013	0.009
H_character_experienced	0.399	0.490	0.400	0.490	0.000	0.012
S_character_experienced	0.118	0.323	0.111	0.315	-0.007	0.008
character_amazonuser	0.802	0.399	0.792	0.406	-0.010	0.010
character_shopno	2.977	1.469	2.951	1.486	-0.026	0.037
D_character_categoryknowledge	0.077	0.266	0.077	0.267	0.001	0.007
H_character_categoryknowledge	0.109	0.311	0.123	0.328	0.014	0.008
S_character_categoryknowledge	0.050	0.218	0.047	0.211	-0.003	0.005
D_character_categoryinterested	0.046	0.209	0.042	0.200	-0.004	0.005
H_character_categoryinterested	0.111	0.314	0.110	0.313	-0.001	0.008
S_character_categoryinterested	0.092	0.290	0.098	0.297	0.005	0.007
character_aware_stealth	0.289	0.453	0.291	0.454	0.002	0.011
character_aware_inflated	0.423	0.494	0.404	0.491	-0.018	0.012
character_aware_deflated	0.124	0.329	0.127	0.333	0.003	0.008
character_aware_contaminated	0.227	0.419	0.228	0.420	0.002	0.010

Table 17: TRrevieworderDate

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRfake	0.496	0.500	0.507	0.500	0.011	0.012
TRwarningmessage	0.514	0.500	0.495	0.500	-0.019	0.012
TRproductorderRating	0.323	0.468	0.324	0.468	0.000	0.012
TRproductorderNoEv	0.333	0.471	0.329	0.470	-0.005	0.012
TRrevieworderUseful	0.000	0.000	0.000	0.000	0.000	0.000
TRtwosided	0.496	0.500	0.500	0.500	0.004	0.012
character_female	0.518	0.500	0.525	0.499	0.008	0.012
character_age	47.894	14.864	48.305	15.271	0.411	0.376
character_individualincome	2.528	2.689	2.449	2.726	-0.079	0.067
character_education_year	5.754	1.892	5.736	1.940	-0.018	0.048
D_character_experienced	0.154	0.361	0.150	0.357	-0.004	0.009
H_character_experienced	0.402	0.490	0.405	0.491	0.003	0.012
S_character_experienced	0.108	0.310	0.121	0.326	0.013	0.008
character_amazonuser	0.805	0.397	0.796	0.403	-0.009	0.010
character_shopno	2.940	1.469	2.983	1.487	0.043	0.037
D_character_categoryknowledge	0.083	0.276	0.069	0.254	-0.014	0.007
H_character_categoryknowledge	0.116	0.321	0.119	0.324	0.003	0.008
S_character_categoryknowledge	0.051	0.219	0.044	0.206	-0.006	0.005
D_character_categoryinterested	0.049	0.216	0.042	0.200	-0.007	0.005
H_character_categoryinterested	0.117	0.322	0.114	0.318	-0.003	0.008
S_character_categoryinterested	0.100	0.301	0.095	0.293	-0.006	0.007
character_aware_stealth	0.296	0.457	0.291	0.454	-0.006	0.011
character_aware_inflated	0.423	0.494	0.420	0.494	-0.003	0.012
character_aware_deflated	0.125	0.330	0.128	0.334	0.004	0.008
character_aware_contaminated	0.226	0.418	0.224	0.417	-0.001	0.010

Table 18: TRrevieworderUseful

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRfake	0.496	0.500	0.505	0.500	0.009	0.012
TRwarningmessage	0.514	0.500	0.519	0.500	0.005	0.012
TRproductorderRating	0.323	0.468	0.336	0.472	0.012	0.012
TRproductorderNoEv	0.333	0.471	0.339	0.473	0.005	0.012
TRrevieworderDate	0.000	0.000	0.000	0.000	0.000	0.000
TRtwosided	0.496	0.500	0.503	0.500	0.007	0.012
character_female	0.518	0.500	0.528	0.499	0.010	0.012
character_age	47.894	14.864	48.737	14.940	0.843	0.369
character_individualincome	2.528	2.689	2.488	2.768	-0.040	0.068
character_education_year	5.754	1.892	5.796	1.918	0.042	0.047
D_character_experienced	0.154	0.361	0.164	0.370	0.010	0.009
H_character_experienced	0.402	0.490	0.404	0.491	0.002	0.012
S_character_experienced	0.108	0.310	0.113	0.317	0.005	0.008
character_amazonuser	0.805	0.397	0.806	0.396	0.001	0.010
character_shopno	2.940	1.469	2.962	1.479	0.022	0.037
D_character_categoryknowledge	0.083	0.276	0.076	0.265	-0.007	0.007
H_character_categoryknowledge	0.116	0.321	0.112	0.315	-0.004	0.008
S_character_categoryknowledge	0.051	0.219	0.044	0.205	-0.007	0.005
D_character_categoryinterested	0.049	0.216	0.049	0.217	0.000	0.005
H_character_categoryinterested	0.117	0.322	0.108	0.310	-0.009	0.008
S_character_categoryinterested	0.100	0.301	0.087	0.281	-0.014	0.007
character_aware_stealth	0.296	0.457	0.296	0.457	0.000	0.011
character_aware_inflated	0.423	0.494	0.408	0.492	-0.015	0.012
character_aware_deflated	0.125	0.330	0.124	0.329	-0.001	0.008
character_aware_contaminated	0.226	0.418	0.233	0.423	0.007	0.010

Table 19: TRwarningmessage

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRfake	0.502	0.500	0.504	0.500	0.003	0.010
TRproductorderRating	0.322	0.467	0.333	0.471	0.011	0.010
TRproductorderNoEv	0.336	0.472	0.331	0.471	-0.005	0.010
TRrevieworderDate	0.339	0.473	0.320	0.467	-0.019	0.010
TRrevieworderUseful	0.331	0.471	0.344	0.475	0.013	0.010
TRtwosided	0.508	0.500	0.492	0.500	-0.016	0.010
character_female	0.532	0.499	0.516	0.500	-0.016	0.010
character_age	48.769	15.212	47.877	14.836	-0.892	0.305
character_individualincome	2.405	2.718	2.569	2.736	0.164	0.055
character_education_year	5.774	1.897	5.751	1.937	-0.023	0.039
D_character_experienced	0.152	0.359	0.160	0.367	0.008	0.007
H_character_experienced	0.401	0.490	0.405	0.491	0.004	0.010
S_character_experienced	0.105	0.307	0.123	0.328	0.017	0.006
character_amazonuser	0.798	0.401	0.805	0.396	0.007	0.008
character_shopno	3.002	1.503	2.923	1.453	-0.079	0.030
D_character_categoryknowledge	0.076	0.266	0.076	0.265	0.000	0.005
H_character_categoryknowledge	0.107	0.309	0.125	0.330	0.018	0.006
S_character_categoryknowledge	0.042	0.200	0.051	0.219	0.009	0.004
D_character_categoryinterested	0.048	0.215	0.045	0.208	-0.003	0.004
H_character_categoryinterested	0.120	0.325	0.106	0.308	-0.015	0.006
S_character_categoryinterested	0.092	0.289	0.096	0.295	0.004	0.006
character_aware_stealth	0.283	0.451	0.305	0.461	0.022	0.009
character_aware_inflated	0.393	0.489	0.439	0.496	0.046	0.010
character_aware_deflated	0.112	0.315	0.139	0.346	0.028	0.007
character_aware_contaminated	0.228	0.419	0.228	0.420	0.000	0.009

Table 20: TRtwosided

	0 / Mean	0 / Std. Dev.	1 / Mean	1 / Std. Dev.	Diff. in Means	Std. Error
TRfake	0.502	0.500	0.503	0.500	0.001	0.010
TRwarningmessage	0.517	0.500	0.502	0.500	-0.016	0.010
TRproductorderRating	0.326	0.469	0.329	0.470	0.003	0.010
TRproductorderNoEv	0.333	0.471	0.334	0.472	0.001	0.010
TRrevieworderDate	0.329	0.470	0.330	0.470	0.000	0.010
TRrevieworderUseful	0.336	0.472	0.340	0.474	0.004	0.010
character_female	0.518	0.500	0.530	0.499	0.012	0.010
character_age	48.270	15.060	48.358	14.996	0.088	0.305
character_individualincome	2.490	2.745	2.487	2.712	-0.003	0.055
character_education_year	5.766	1.916	5.759	1.919	-0.007	0.039
D_character_experienced	0.154	0.361	0.158	0.365	0.004	0.007
H_character_experienced	0.395	0.489	0.412	0.492	0.017	0.010
S_character_experienced	0.110	0.313	0.118	0.323	0.008	0.006
character_amazonuser	0.793	0.405	0.810	0.392	0.017	0.008
character_shopno	2.945	1.465	2.979	1.492	0.033	0.030
D_character_categoryknowledge	0.077	0.267	0.075	0.264	-0.002	0.005
H_character_categoryknowledge	0.113	0.317	0.118	0.323	0.005	0.006
S_character_categoryknowledge	0.046	0.210	0.047	0.211	0.000	0.004
D_character_categoryinterested	0.051	0.221	0.042	0.201	-0.009	0.004
H_character_categoryinterested	0.117	0.321	0.109	0.312	-0.008	0.006
S_character_categoryinterested	0.095	0.293	0.093	0.291	-0.001	0.006
character_aware_stealth	0.295	0.456	0.294	0.456	-0.001	0.009
character_aware_inflated	0.417	0.493	0.416	0.493	-0.001	0.010
character_aware_deflated	0.127	0.333	0.124	0.330	-0.003	0.007
character_aware_contaminated	0.228	0.420	0.227	0.419	-0.001	0.009