The Impact of Negative Reviews on Online Search and Purchase Decisions*

Marton Varga[1] and Paulo Albuquerque[2]

*The authors express their gratitude to the participants of the 8$^{\text{th}}$ Workshop on Consumer Search and Switching Cost and the 39$^{\text{th}}$ Annual ISMS Marketing Science Conference for their valuable feedback. Additionally, they would like to extend a special thanks to the Wharton Customer Analytics Initiative and its partners for providing the data.

[1]Marton Varga is Assistant Professor of Marketing, Bocconi University, Italy. email: marton.varga@unibocconi.it (corresponding author)

[2]Paulo Albuquerque is Professor of Marketing, INSEAD, France. email: paulo.albuquerque@insead.edu

Abstract

Despite evidence indicating the significant influence of online reviews on purchase decisions, even after taking into account a product's average rating (Vana and Lambrecht 2021), the underlying factors responsible for this effect and the broader impact of reviews on consumer decision-making remain uncertain. This study uses click-stream data from a major online retailer to explore how negative reviews affect consumer search and purchase decisions. Leveraging exogenous variation created by the display of online reviews sorted by recency, the authors find that negative reviews significantly reduce a product's purchase probability because they (1) contrast with the often-high average product rating, (2) decrease the probability that consumers continue browsing for information about the focal product, (3) increase the probability of visiting the page of substitute products, and (4) increase the probability of viewing reviews about substitute products. Importantly, these effects apply to utilitarian products but not hedonic products and when reviews pertain to product functionality or customer service but not to taste-related factors. The authors estimate a product's vulnerability to negative reviews along two dimensions—purchase and search probability for substitutes—and display these effects on a two-dimensional map.

Keywords: Customer reviews, Online shopping, Information search, Elasticity of demand, Topic modeling.

Online retailers commonly post customer ratings and reviews on their websites. Online shoppers frequently refer to reviews for guidance at multiple stages of their online journey to obtain information and to evaluate, compare, and decide among various purchase options (Mudambi and Schuff 2010). Several studies have shown that reviews and ratings are regarded by consumers as highly credible inputs in purchase decisions (Floyd et al. 2014; Chen and Xie 2008; Brown et al. 2005; Kozinets et al. 2010; Liu 2006) and that consumers often trust reviews more than advertising (Cheong and Morrison 2008; Hung and Li 2007).

When consumers shop online and browse through product listings, they are typically presented with only a limited amount of information, such as the brand name, price, number of reviews, and average product rating. However, this limited information is often not enough for consumers to make an informed purchasing decision. Therefore, they seek out more detailed information by clicking through to individual product pages (Fung and Lee 1999). On these product pages, consumers find a wealth of additional information that can help them make a more informed decision. For example, they can read individual customer reviews, see detailed product descriptions, view high-quality product images, and even watch product videos.

While online shoppers clearly take product page information into account, previous research has focused on the impact on choice of aggregate measures, for example the average product rating, rather than on the role of single reviews (see e.g., Babić Rosario et al. 2016). One exception to this, a recent paper by Vana and Lambrecht (2021), shows that an individual online review can have a significant impact on purchase probabilities, even after accounting for the average rating of the product. This effect is moderated by the position in which the review is displayed, how much it contrasts with the average rating, and how much variance exists in ratings.

Our study builds on this line of research by focusing on the impact of negative reviews on online consumer decisions. We do so for several reasons. In psychology, the concept of *nega-*

tivity bias has been widely demonstrated. Bad emotions, bad experiences, and bad feedback have greater impact than good ones (e.g., Baumeister et al. 2001). Negative information tends to influence evaluations more strongly than positive information (Ito et al. 1998), and distinguishable motivational systems have been shown to underlie the assessment of negative and positive stimuli (Cacioppo, Gardner, and Berntson 1997). When subjects were asked to process multiple pieces of information with different valence about the same item or person (e.g., Fiske 1980; Klein 1996; Skowronski and Carlston 1989), negative information had a more significant impact on their final impression. Indeed, negativity bias has been found to exist not only in humans but even in animals (Rozin and Royzman 2001). In the choice domain, negative reviews significantly impact a consumer's propensity to choose the item due to an increased perception of risk (Angie et al. 2011), especially if it is accompanied by pictures (Xie et al. 2011). Finally, recent work has documented an increase in the frequency of fake reviews, particularly of positive reviews that are used to increase the product's average rating and number of reviews (He, Hollenbeck, and Proserpio 2022; Salminen et al. 2022). Facing this situation, consumers are likely to place more weight on the frequently genuine negative reviews.

We study the effects of negative reviews at three stages of the buying process. First, at the *initial information* stage, consumers collect information such as the average product rating and the number of available reviews before visiting the product page. From this prior information, usually displayed on a category page, consumers form an initial impression that likely moderates the relevance of negative reviews in the next decision steps. Second, at the *product information stage*, consumers collect more detailed information about products by visiting their pages. We test whether the presence of one or more negative reviews—in place of positive reviews—affects these information search decisions. Third, we evaluate whether these mechanisms jointly translate to stronger or weaker effects of negative reviews at the *purchase stage,* and if so, whether this effect is moderated by the type of product and the content of negative reviews.

To empirically study the effects of negative reviews, we use a dataset on consumer decisions at a large online retailer offering a broad selection of products in the technology and home-and-garden categories in the United Kingdom. The particularly rich dataset tracks multiple steps of the buying process, allowing us to study how the replacement of one or more positive reviews with the respective number of negative reviews affects search and purchase decisions.

Our identification strategy relies on the retailer's "newest first" policy to display customer reviews. When a new review is posted, the oldest review on the product page is relegated to the second page of reviews that most consumers do not visit. As different consumers see the product page at different times, i.e., before versus after the arrival of the newest review and respective relegation of the oldest to the second page, this variation in reviews creates a quasi-natural experimental setting that can be leveraged for our research purposes. Specifically, we compare decisions by consumers who arrived at the product page, scrolled down to the bottom of the product page, where up to 5 reviews are displayed, and found n negative reviews and $5 - n$ positive reviews, with decisions by consumers who did likewise but instead saw $n' \neq n$ negative reviews and $5 - n'$ positive reviews. A negative review is defined as a review with three stars or less, out of five stars. A review is considered positive when it has four or five stars.

Our results show that, for consumers who scrolled down to browse product reviews, the probability of purchasing a product decreases by 41.8% when one negative review is present compared to when no negative review is displayed. As these consumers represent about one fifth of all consumers visiting the product page (the remaining four fifths do not scroll down to read reviews), this effect approximately represents an 8.4% drop in the overall demand. Moreover, a single negative review on the product page increases the probability of further browsing for substitute products by 9.7%. Consumers who view more than one negative review, in place of positive ones, are even more likely to search for substitutes.

We also discover insights into the circumstances under which reviews have the most impact. Specifically, we find that negative reviews have a greater (lesser) effect on the purchase decisions of utilitarian (hedonic) products, which is consistent with the findings of Sen and Lerman's (2007) experimental study. This difference may be explained by the fact that preferences for hedonic products are more diverse, as personal taste plays a more significant role, while preferences for utilitarian products are more uniform. Hence, consumers can learn more from previous user experiences in utilitarian products categories (Feick and Higie 1992). Motivated by these findings, we take a closer look at the text of the reviews. Using Latent Dirichlet Allocation—a topic modeling approach—we find that negative reviews have a stronger effect on the purchase decision when they describe problems related to the product's functionality or to customer service. In contrast, negative reviews do not have a significant impact when they relate to matters of taste, such as product design and colors.

The closest work to ours is the one by Vana and Lambrecht (2021), who quantify the impact of individual online reviews on purchase decisions. Similarly to our work, the authors focus on the decisions of consumers who scroll down to the review section on the product page. Their study also benefits from the same variation in the display of reviews. However, their paper diverges from ours in that they primarily focus on positive reviews and on measuring how the impact of individual reviews on purchase decisions changes with the review's position (first to fifth) on the product page. In contrast, we study the effect of negative reviews and further explore the factors that lead consumers to react strongly to these reviews. We do so by looking at information search behavior alongside purchase decisions, and by analyzing the text of the reviews using a topic modeling approach to identify the conditions where reviews matter the most.

Our research contributes to the literature on online word-of-mouth (WOM) and customer reviews. The relationship between online WOM and sales has been the subject of various papers and three meta-analyses (Floyd et al. 2014; You, Vadakkepatt, and Joshi 2015; Babić Rosario et al. 2016). Research has demonstrated that reviews are considered by consumers to be compelling sources of information. For example, Chevalier and Mayzlin (2006) found that an additional one-star review negatively impacted the sales rank of a listed book using data from Amazon and Barnes & Noble. Sun (2012) showed that a higher standard deviation of historical ratings improved a book's sales rank on Amazon when the product's average rating was low. Wu et al. (2015) analyzed how consumers updated their restaurant preferences based on online reviews depending on their similarity to the review writer, and Liu et al. (Liu, Lee, and Srinivasan 2017) used a supervised deep learning algorithm to investigate how the content of reviews affected sales.

We add to this literature by quantifying the impact of negative customer reviews on product demand, formation of consideration sets, and information search decisions. To the best of our knowledge, our paper is the first to document the effect of individual reviews on both information search and purchase decisions. Our findings advance the debate on whether consumers use a simultaneous search framework when browsing for products—setting a predefined number of options to look at before starting search decisions (De los Santos, Hortacsu, and Wildenbeest 2012)—or carry out search in a sequential way—when learned information about searched products can affect the consideration set (Kim, Albuquerque, and Bronnenberg 2010; Honka 2014; Honka and Chintagunta 2016). By demonstrating that negative reviews prompt consumers to intensify their search for alternative products, our study suggests that consumers tend to search sequentially. We also extend the growing body of work on how consideration sets are formed and how marketing efforts influence the range of products that consumers are willing to consider (Demuynck and Seel 2018; Manzini and Mariotti 2018; Eliaz and Spiegler 2011; Van Nierop et al. 2010; Allenby and Ginter 1995) by quantifying the effect of negative reviews on the size of the search set.

Our research has substantial implications for practitioners. To illustrate the managerial usefulness of our findings, we provide a visualization for the product's exposure to negative reviews on a *vulnerability map*. For the set of consumers who scroll down to the review area, a two-dimensional map displays the effect of one additional negative review (replacing a positive review) on (1) the product's purchase probability and (2) on the probability that consumers will visit the page of substitutes after viewing the reviews of the focal product. With the growing importance of customer reviews in online shopping, we propose the concept of *elasticity of category sales to negative reviews*: the percentage change in sales of a product category due to a 1% increase in negative reviews accessible on the product pages of that category. We apply this metric to a diverse set of product categories and find that this elasticity is greater for categories where consumers browse for more products before purchase.

The rest of the paper is organized as follows. In the next section, we present the conceptual framework on the role of negative reviews over the purchase process. We subsequently introduce our research design and identification strategy. We then discuss results and managerial implications. The last section summarizes and provides future research directions.

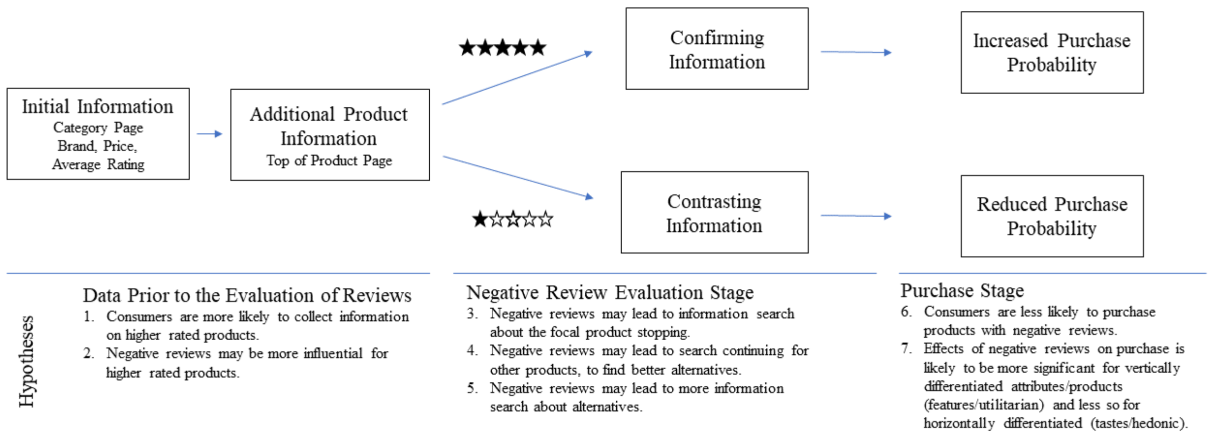
CONCEPTUAL SETTING

In this section, we introduce the proposed framework for understanding the role played by negative reviews on search and purchase decisions in an online environment. Figure 1 displays the sequence of stages of the online consumer journey studied in this paper.

Prior Information

Before evaluating individual reviews, potential buyers usually consider the product's average customer rating. This value is typically defined on a continuous scale of one to five and is an important piece of prior information that impacts a consumer's impression of the product (e.g., Floyd et al. 2014; Chen and Xie 2008). When consumers encounter a high average customer rating for a product, it creates a positive impression and leads to more clicks to the product page with an optimistic expectation about the remaining information. Once the consumer clicks on the product page and scrolls down to the review area, positive reviews will confirm the initial belief of quality, enhancing the overall impression of the product and increasing the probability of purchase. Conversely, encountering negative reviews for a

Figure 1: Framework for Evaluating the Impact of Negative Reviews on Consumer Choice



product with a high average customer rating will contradict the consumer's initial favorable belief, leading to a worse overall impression of the product. The contrast between the prior perception of quality (average rating) and the individual review scores is likely to increase the salience of the reviews and affect the decisions to collect further information (e.g., reading more reviews) and to purchase the focal product. Therefore, we expect that negative reviews will have a greater impact on search and choice decisions when the prior information is more favorable.

Effects of Negative Reviews on Search Decisions

By updating the perception of a product's quality, online reviews influence consumer decisions to gather information before the purchase decision (Zhang et al. 2014; Bigne, Chatzipanagiotou, and Ruiz 2020). Negative reviews, in particular, have the potential to degrade the consumers' perception of the focal product and prompt them to seek out substitutes. In our context, we expect the presence of a negative review to extend the duration of search because the additional information on the product page reduces the perceived utility of the product being evaluated. The finding would be consistent with the sequential search frameworks (e.g., Kim et al. 2010; 2016), which suggest that consumers tend to explore products one at a time, gathering information as they go. In contrast, simultaneous search models assume that consumers sample a fixed number of products and do not add further options based on information collected during the search process (e.g., De los Santos, Hortacsu, and Wildenbeest 2012).

Moreover, when faced with negative reviews that do not agree with a high average rating, consumers have evidence that it is useful to update prior beliefs by collecting more information. This is likely to make consumers proceed with the search and peruse reviews of alternative products. In essence, viewing a negative review is expected to prompt consumers to expand their consideration set, indirectly lowering the likelihood of the focal product being the final choice.

Moderating Effect of Product Type and Review Topic

Previous literature has established that consumers apply different information-processing strategies depending on product type, and that there are distinct hedonic and utilitarian shopping value dimensions (Babin, Darden, and Griffin 1994; Voss, Spangenberg, and Grohmann 2003). Hence, it is likely that the relevance of negative reviews differs by product type. For example, when searching for hedonic products, Lopez and de Maya (2012) observed that consumers in a negative mood showed attitudes and purchase intentions consistent with the valence of the review. However, consumers in a positive mood found arguments (and counter-arguments) to compensate for negative information and remained unaffected by negative reviews. Consumers often attribute the negative content of the review to the reviewer's own personal taste in a hedonic context. In contrast, for utilitarian products, consumers attribute the reviewer's opinion to product features, rendering negative reviews more instructive. In a related vein, a study by Sen and Lerman (2007) that investigated the influence of negative information on consumer choice detected a negativity bias solely in utilitarian product categories. Hence, we expect to find that negative reviews matter more for utilitarian products than for hedonic ones.

The topic of the negative review may also influence its effect on choice. While consumers could potentially disagree on taste-related attributes (Markopoulos and Clemons 2013) of a product (e.g., aspects of design, color, shape, material or size), they might easily agree on attributes related to the product's functionality (statements such as the product "does not turn on", "breaks apart", "is not powerful") or service quality (e.g., delivery, customer support, guarantee). Learning about the product's mismatch with a previous consumer suggested by a low rating on taste-related attributes does not eliminate the possibility that it could still be a good match for the consumer contemplating the purchase. Indeed, Sun (2012) showed that a greater variance in historical product ratings could increase sales for products with a low average rating. Conversely, learning that a product fell short on a functional attribute is likely to prompt doubts about whether it will fulfill its core purpose. Accordingly, we posit that negative reviews are of greater consequence when their topic relates to product functionality but are less relevant when they describe taste-related aspects.

Effects of the Evaluation of Negative Reviews on Purchase Decisions

In summary, we anticipate that negative reviews will reduce the likelihood of a purchase. This effect is expected to be stronger when prior information—in the form of average consumer rating—is in contrast to the negative review. It is also expected to be more pronounced as the number of alternative products considered increases due to the presence of negative reviews. Furthermore, the impact of negative reviews is expected to be greater when the review content and product type relate to vertically differentiated components.

DATA AND IDENTIFICATION

In this section, we first provide details about the dataset used in our empirical setting. We then give an intuitive explanation of our identification strategy, which includes model-free evidence on how negative reviews impact consumer decisions.

Data

We use data on consumer search and purchase decisions at a large online retailer based in the United Kingdom between February 1 and March 31, 2015. The dataset includes click-stream information on individual visits to web-pages of home-and-garden and technology products. A product is defined as a stock keeping unit (SKU). The retailer sells both utilitarian products—such as microwaves and juicers in the technology department, and mops and bath mats in the home-and-garden department—and hedonic products—such as televisions and consoles, as well as pet toys and beach huts. The products are classified by the online retailer into 627 categories. We observe 31,283 product pages that were visited by at least one consumer during the analysis period, of which 14,890 product pages had at least one review. For each product, we observe reviews, prices, product category and brand, and page visits and purchases during the two months.[3]

For our analysis, we define a *search session* as a visit that starts when a consumer first opens a product page and ends either with a purchase or with the last observed click in that category. Altogether, in our dataset, 121,391 unique consumers carried out 217,177 search sessions. They visited a total of 400,126 product pages during these sessions, viewed product reviews 74,383 times, and made a purchase in 9,502 of the cases. 64% of the consumers clicked on exactly one product page, 18% clicked on two product pages, while the rest of the consumers clicked on three or more product pages.[4] Assuming that clicking on a product

[3] For products with no sales, we did not obtain the price directly from the data provider but scraped archived data from `https://archive.org/`. For 1,693 products where price information was unavailable in the archives we used the category average price.

[4] In the remaining of the paper, we use the terms "click on product page" and "search for a product"

page implies consideration, the average consideration set size in the data is 1.85.

In terms of reviews, there were 575,084 product reviews with a respective rating available on the website. Products were rated on a 1-to-5 discrete scale represented by stars, with the following distribution: 57% of the reviews had a 5-star rating, 26% had a 4-star rating, 6% had a 3-star rating, 3% had a 2-star rating, and 8% had a 1-star rating. We classify a review as negative if it has three stars or fewer.

While browsing a category page, the product's price, average rating, and number of reviews are readily visible without the need to access the specific product page. Upon clicking on the product page, consumers are presented with two distinct sections (Figure 2). The *product description* section is immediately visible when the page is opened. It repeats the information from the category page and provides further details about the item, such as technical specifications, size, and color. In contrast, individual reviews are not immediately visible upon arrival at the product page; consumers must scroll down to view them. We denote the area allocated to reviews on the product page as the *review area.* The number of reviews displayed on the product page is either equal to the total available reviews or capped at five if the product has at least five reviews. By default, the reviews on the product page are ordered by recency.

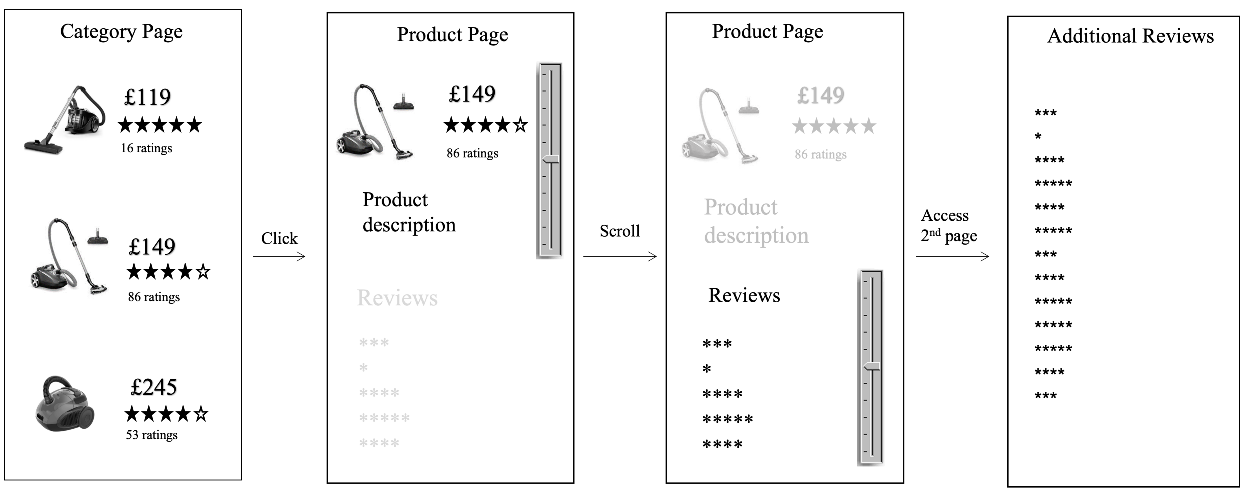
After scrolling down and accessing the review area, the consumer has the option to continue browsing for additional reviews on separate pages, each containing up to twenty reviews. Consumers can also choose to sort reviews by rating and helpfulness after they have scrolled down to the review area on the product page. Overall, consumers decided to browse the second page of reviews a total of 13,614 times and sorted the reviews on 1,106 occasions. In other words, about four out of five consumers chose to only view the reviews on the product page.

We highlight that this setup—reviews ordered by recency and accessed by scrolling down—is typical among retailers. For example, in the United States, reviews on Ama-

interchangeably.

zon.com are displayed at the bottom of the product page and sorted according to either recency or relevance. On Target.com, reviews are sorted by recency as the default option. Walmart.com shows the most helpful reviews from verified purchases by default but consumers can sort the reviews by recency. In France, Fnac.fr requires consumers to click on a link to obtain the customer reviews for a product, after which reviews are sorted by recency as well. At Tesco.com in the United Kingdom, consumers see the newest reviews first, close to the bottom of the product page.

Figure 2: Illustration of the Category Page, Product Page, the Review Area on the Product Page, and Additional Pages of Reviews



The unit of analysis is a consumer-product pair, i.e., a product page visit done at a given time by a unique consumer. For our analysis, we filter the data on two conditions. First, we only include observations where consumers scrolled down to the review section, which is measured by an indicator variable in our data. Second, we retain observations only when the product had at least one customer review at the time of the visit. The resulting dataset contains 50,211 search sessions, 68,960 product-consumer observations from 37,032 consumers.[5] In this dataset, 1,775 purchases were made.

Table 1 presents the descriptive statistics. In our data, half of the consumers viewed

[5]There were 5,423 observations with missing values for either the dependent or independent variables. We dropped these observations from our analysis.

at least one negative review on the product page. On average, consumers encountered .84 negative reviews after scrolling to the review area. The mean number of reviews during the consumer visits in our analysis is 122.5, while the average rating is 4.21. The average price is £88.5, with £118.9 for the technology products and £74.7 for home-and-garden products.

Table 1: Descriptive Statistics

	Mean	**SD**	**Min**	**Max**
Percentage of product pages with negative reviews[a]	51	50	0	100
# Negative reviews on product page[a]	.84	1.04	0	5
# Reviews per product[b]	122.5	248.9	1	4761
Average rating per product[b]	4.21	.63	1	5
Price per product (£)[b]	88.5	115.5	.01	4199.9
Number of products browsed per consumer[b]	2.21	2.17	1	43

a: Products for which at least one consumer scrolled to the reviews on the product page.
b: Consumer-product pairs where the consumer scrolled to the reviews on the product page.

Identification Strategy

During the two months of our data collection period, the online retailer updated the product pages daily with the publication of reviews. New reviews took the top spot on the review area, while older reviews moved further down or were relegated to subsequent review pages. This sequential arrival of reviews, without managerial influence, creates exogenous variation in the location of reviews on the retailer's website at the time of arrival of a consumer.[6]

Our identification strategy centers on comparing consumers who arrived at the product page, scrolled down to the review area and, from the total of N reviews there, found n negative reviews (and $N - n$ positive reviews) with consumers who also scrolled down and saw $n' \neq n$ negative reviews (and $N - n'$ positive reviews).[7] We also measure the impact of seeing *at least one* negative review by comparing consumers who scrolled down to the review area when there was at least one negative review with consumers who arrived when there was no negative review, either because none was there or because the negative review

[6] From conversations with company representatives, we believe that the firm does not manipulate the publication of submitted reviews beyond the standard censoring of offensive wording.

[7] N=5 if the product has at least 5 reviews, otherwise equals to the total available reviews.

got relegated to the second review page that the consumer did not browse. This method has the nature of a regression discontinuity approach, allowing us to compare outcomes between periods with and without the presence of additional negative reviews.. For an additional discussion of the identification using this exogenous shock, readers are referred to Vana and Lambrecht (2021).

A key assumption underlying our analysis is that consumers are not aware of the individual ratings until they scroll down to the review area. As a test of this assumption, we ran a regression explaining the scrolling decision as a function of the mean of the individual review ratings and found that the decision to scroll down is not significantly explained by the individual review ratings in the review area (Web Appendix, Table W1).

Using the exogenous variation in the number of negative reviews seen by consumers, we examine how their decisions are affected during a search session, such as whether they decide to buy the product and whether they visit pages of alternative products. Additionally, we explore other related consumer decisions, including whether they choose to visit the second page of reviews, sort the reviews of the focal product, and scroll down to the review area of substitute products.[8]

Table 2: Model-Free Evidence of the Effect of Negative Reviews on Consumer Choices

		# Consumers Deciding to				
# Negative Reviews	**# Arriving Consumers**	**Product Purchase**	**Open Second Review Page**	**Sort the Reviews**	**Search for Substitute**	**Scroll to Reviews of Substitute**
0	33,761	998	5,959	549	13,777	5,883
		(3.0%)	(17.7%)	(1.6%)	(40.8%)	(42.7%)
1	20,472	518	3,976	288	8,929	4,062
		(2.5%)	(19.4%)	(1.4%)	(43.6%)	(45.5%)
2	8,909	181	1,683	113	4,018	1,998
		(2.0%)	(18.9%)	(1.3%)	(45.1%)	(49.6%)
3	4,190	63	754	48	1,951	982
		(1.5%)	(18.0%)	(1.1%)	(46.6%)	(50.2%)
4	1,229	15	161	11	609	309
		(1.2%)	(13.1%)	(.9%)	(49.6%)	(50.7%)
5	399	0	40	2	189	77
		(.0%)	(10.0%)	(.5%)	(47.4%)	(40.7%)

Note: The table displays the number of consumers who decided to take a certain action after viewing the review section, using the sample of consumers who scrolled down on the product page. Percentages of consumers taking the respective decision conditional on the number of negative reviews are in parentheses.

[8]The substitute or competitor product is defined as the product that is searched immediately after the focal product.

Table 2 displays model-free evidence, conditional on the number of negative reviews observed by consumers arriving at the product page. Each column shows the number and the percentage of consumers who decided to take the respective action. Without additional control variables and with the caveat that the number of observations is limited when considering the presence of four or five negative reviews, we observe that from five positive reviews (i.e., zero negative reviews) to five negative reviews (i.e., zero positive reviews), there is a significant decrease in the rate of purchase. There is also a decline in the willingness to visit the second page of reviews about the focal product, a decrease in the rate of sorting reviews, an increase in the likelihood of visiting the page of a competing product, and a rise in the willingness to scroll to the review area on the substitute's page.

METHODOLOGY

We start our methodology description with the model on purchase outcomes. We then move backwards and address the different stages in the decision process, which can explain why negative reviews affect product purchase. We also describe the text-analysis method we use to study how the review topic moderates the effect of negative reviews.

We chose to evaluate the impact of negative reviews using separate regressions. Although this approach is at odds with recent work on search and choice modeling using integrated frameworks where the utility of search and choice depend on the same primitives (e.g., Ursu 2018; Ursu, Wang, and Chintagunta 2020), we opted for a reduced-form, staged approach for the following reasons. First, most previous approaches focused on studying only two consumer decisions, while our study considers several search decisions and the purchase outcome. We also take into account the moderating factor of prior information and product type. This complexity could make an integrated model lengthy to explore and difficult to estimate. Second, our objective is not to conduct counterfactual analyses but to provide inputs to managers about the sensitivity of each stage of the purchase process with respect to the presence of negative reviews. Consequently, we are first and foremost interested in the effect of negative reviews on the way consumers search for and react to information.

Effect of Negative Reviews on Product Purchase

To establish the effect of negative reviews on product purchase, we define the dependent variable as the yes/no decision of consumer i regarding the purchase of product j. For the purchase decision, we formulate the model:

$$Purchase_{ij} = \alpha_1 + \beta_1 R_{ij} + \mathbf{x}'_{ij}\boldsymbol{\gamma_1} + \mathbf{c}'_i\boldsymbol{\delta_1} + \mu_{1j} + \eta_{1i} + \epsilon_{1ij}, \tag{1}$$

where R_{ij} is the number of negative reviews on the page of product j that are displayed to consumer i at the time of visit. Note that the website shows the same reviews to different consumers, provided they browse the product reviews before the arrival of an additional review.[9] Therefore, the coefficient β_1 measures the effect of viewing one additional negative review, in place of a positive review, on the probability of purchase. Consequently, our empirical analysis starts by establishing the overall effect of an individual negative review on the purchase probability of the focal product.

In Equation 1, $\mathbf{x}_{ij}$ is a vector of product page characteristics at the time of the visit by consumer i including product j's average rating, log number of reviews, log price, and the position of the newest negative review on its product page. The term $\mathbf{c}_i$ is a vector of variables describing the consumer's search and purchase behavior until viewing the page of product j, including the view rank of j (i.e., how many products consumer i viewed previously) and a dummy variable indicating whether the consumer purchased a product from the same category during our data analysis window. Furthermore, μ_{1j} is a product-specific fixed effect, η_{1i} stands for a day-fixed effect corresponding to the time of the consumer visit, and ϵ_{1ij} is an independent and identically distributed (i.i.d.) error term. Standard errors are clustered

[9] We do not use time subscript in the equations for clarity of exposition but day-fixed effects are included in our formulation.

at the product level. For the purchase model, we restrict our attention to consumers who do not open the second review page or sort the reviews, allowing us to compare groups of consumers that are as homogeneous as possible.[10]

We use two alternative definitions for the term R_{ij} to gain comprehensive insights into the effect of negative reviews and, simultaneously, as a robustness check. In an alternative formulation, R_{ij} takes the value of 1 if the product page includes at least one negative review when visited by consumer i, and 0 otherwise. In a third formulation, we define four dummy variables, $R_{ij}^1, R_{ij}^2, R_{ij}^3, R_{ij}^4$, with $R_{ij}^n = 1$ if and only if the number of negative reviews is exactly n, for $n = 1, 2, 3$, and $R_{ij}^4 = 1$ if and only if the number of negative reviews is 4 or 5. The base category is zero negative reviews. Note that the interpretation of β_1 depends on which alternative version is used for R_{ij}. It can capture the marginal effect of the number of reviews, the overall effect of the presence of a negative review on the product page, or the effect of viewing an exact number of negative reviews.

Prior Information

To assess the impact of prior information on consumer decision-making, we examine whether products with higher average ratings and more reviews are more frequently selected for further consideration by consumers. Specifically, we model the decision to click on a product page as a function of the information provided on the category page and estimate the model

$$
\begin{aligned}
Visits_{jt} = a_0 + a_1 * Avgrating_{jt} + a_2 * Reviews_{jt} + a_3 Price_{jt} + \\
\sum_{k=1}^{K} b_k I^C\left(Category_j = k\right) + \sum_{l=1}^{L} b_l I^B\left(Brand_j = l\right) + \sum_{m=1}^{M} d_m I^D\left(Day_t = m\right) + e_{jt}, \quad (2)
\end{aligned}
$$

where $Visits_{jt}$ indicates the log number of clicks on product j on day t, $Avgrating_{jt}$ is the average consumer rating, $Reviews_{jt}$ is the log number of reviews, and $Price_{jt}$ is the log price

[10]We ran a robustness test including consumers who paginate or sort the reviews and the results are substantively similar.

of product j on day t. The terms I^C, I^B, I^D are indicator variables for category, brand and day, and e_{jt} is an error term. We estimate Equation 2 using product-day observations where at least one visit occurred.

In addition, we divide the sample into two sets using a median split based on the average rating and separately run the purchase decision regression proposed in Equation 1. This approach identifies whether the impact of a negative review on product purchase varies depending on the prior information about the product rating. On the one hand, pages of products with high ratings are likely to be visited more often. On the other hand, the presence of a negative review on the page of a higher-rated product creates a stark contrast between the average rating and the star rating of the negative review. This could lead to a stronger impact of the negative review on the purchase of products with higher average ratings.

Effect of Negative Reviews on Information Search

In order to examine the impact of negative reviews on outcomes beyond purchase, we employ the regression framework of Equation 1, modifying the dependent variable as follows, while keeping the same independent variables.

We explain whether the consumer decided to visit a second page of reviews about the focal product j using:

$$Open_{ij} = \alpha_2 + \beta_2 R_{ij} + \mathbf{x'}_{ij}\boldsymbol{\gamma_2} + \mathbf{c'}_i\boldsymbol{\delta_2} + \mu_{2j} + \eta_{2i} + \epsilon_{2ij}, \tag{3}$$

where $Open_{ij}$ is a dummy variable indicating whether the consumer opened the second review page of the focal product.

The decision to sort reviews about the focal product is modeled as

$$Sort_{ij} = \alpha_3 + \beta_3 R_{ij} + \mathbf{x'}_{ij}\boldsymbol{\gamma_3} + \mathbf{c'}_i\boldsymbol{\delta_3} + \mu_{3j} + \eta_{3i} + \epsilon_{3ij}, \tag{4}$$

where $Sort_{ij}$ is an indicator for the sorting decision of individual reviews by any possible criterion. In Equations 3 and 4, β_2 and β_3 measure whether negative reviews increase (or decrease) the consumer's willingness to view further information about the focal product.

Next, we regress a dummy variable $Search_{ij}$, indicating whether consumer i visited the page of at least one competitor after visiting the product page of j, on the same set of covariates as before:

$$Search_{ij} = \alpha_4 + \beta_4 R_{ij} + \mathbf{x'}_{ij}\boldsymbol{\gamma_4} + \mathbf{c'}_i\boldsymbol{\delta_4} + \mu_{4j} + \eta_{4i} + \epsilon_{4ij}, \tag{5}$$

where β_4 measures the effect of negative reviews on further visits to competing product pages. We exclude consumers who opened the second page of reviews or sorted the reviews to keep the sample as homogeneous as possible.[11]

Lastly, we seek to understand whether the consumer who, immediately after viewing the reviews of product j, visited the page of a competitor product $j' \neq j$ and scrolled down to the review area of that competitor. Hence, our dependent variable is the yes/no decision of scrolling down to the review area on the page of j'. We then run a regression where the dependent variable is $Scroll_{ij'}$, which is an indicator variable taking the value 1 if the consumer scrolled down to the reviews of j', conditional on visiting j' after viewing the reviews of j, and 0 otherwise:

$$Scroll_{ij'} = \alpha_5 + \beta_5 R_{ij} + \mathbf{x'}_{ij}\boldsymbol{\gamma_5} + \mathbf{c'}_i\boldsymbol{\delta_5} + \mu_{5j} + \eta_{5i} + \epsilon_{5ij}, \tag{6}$$

where β_5 measures the effect of negative reviews on the willingness to view reviews of the competitor product. These approaches consider all products. However, we will also conduct separate analyses for hedonic and utilitarian products to evaluate if the type of product has a notable influence on the effect of negative reviews.

[11] We run a robustness test including consumers who paginate or sort the reviews and found substantively similar results.

Moderating Effect of Review Topic

We used the Latent Dirichlet Allocation (LDA, Blei, Ng, and Jordan 2003) approach to categorize the negative reviews based on their content in order to study their heterogeneous effects on purchase decisions. The LDA method has previously been successfully applied in marketing research (Tirunillai and Tellis 2014).

The natural language processing toolkit *NLTK* was used to clean the data, Pythons's *gensim* library was used to create the dictionary of features through which the model is trained, and *Google Colab* was used to execute our LDA topic modeling task. To train the algorithm, we used the 25 most-recent reviews of all products in our dataset, regardless of how often consumers scrolled to or viewed these reviews. Although this choice was primarily driven by the computational intensity of the task, by doing so we also ensured that we did not bias the training sample by giving excessive weight to the reviews of popular products and that we utilize newer reviews that are more relevant and more often displayed to the consumers. We applied the standard filtering process in which irrelevant terms such as newline characters, single quotes, and stop words were removed from the text of the reviews.[12] Positive reviews (i.e., reviews with ratings higher than 3-stars) were excluded from the training set. The optimal number of topics was selected using the highest coherence score.

We used the trained algorithm to classify the topic of every negative review. For each review, contributing topics are identified and among these topics the dominant one is selected. The dominant topic is defined as the one with the highest percentage contribution to the overall text of the review. We were able to identify four distinct topics among the 11,452 unique negative reviews consumers viewed in our data.[13] We found that the three major topics jointly represent 97.1% of the content in negative reviews. To derive our labels, we took the common approach (Zhong and Schweidel 2020) to look at the most relevant words in

[12] In each review text, to improve accuracy, bi-words and tri-words (bi-grams and tri-grams) were also created as features to train the model.

[13] In the training set the optimal number of topics was six. However, considering the distribution of topics across all reviews consumers saw, we decided to merge the other three topics with the lowest frequency into a single category.

the unlabeled categories. In the first category, *functionality*, we encountered frequent words like “make”, “well”, “work”, ”value”, ”smell”. In the second category, *customer service*, we noticed frequent words such as “order”, ”line”, ”phone”, ”delivery”, ”wait”, ”deliver”, ”arrive”, ”service”, ”available”. In the third category, *matters of taste*, we observed the words “look”, ”cheap”, ”poor”, ”light”, ”nice”, ”wood”, ”plastic”, ”hard”, ”feel”, ”flimsy”, ”thin” among the most frequently appearing ones.

Overall, we found 19.8% of the negative reviews to have functionality as dominant topic, 56.7% to have customer service as dominant topic, and 20.6% to have matters of taste as dominant topic. The rest of the reviews (2.9%) predominantly addressed other topics and concerns. Negative reviews in the functionality topic typically note that the product failed to meet expectations of one or more of its core purposes. An example of this topic is a review that included: “product (...) producing a piece of toast that isn't actually fully toasted”. Negative reviews in the customer service category talk about the product being shipped late, poor customer service, or unexpected stock-outs. An example is a review with the text: “... could not find my order number, no delivery, reordered, phone call, no delivery, phone call to head office (...) no delivery, they canceled my order”. The matters of taste topic consists of negative reviews that point to the disappointing design or look of the product. An example text of this group contains the following: “... the oak effect frame is quite disappointing. It looks cheap because of the laminated covering ...”. As an illustration, we present the full text of the ten most typical reviews by topic in Tables W2, W3, and W4 of the Web Appendix.

Our strategy for estimating the effect of negative reviews according to topic is similar to the main approach with Equation 1, but, instead of having one variable indicating the number of negative reviews, we include one such variable for each of the four categories of reviews. By doing so, we measure the marginal effect of an additional negative review on a particular topic. We also run a second specification in which we use a dummy for the existence of a negative review of a certain topic on the product page.

RESULTS

Product Purchase

Table 3 displays the results related to the purchase probability of product j, for consumers who scrolled down to the review area. As previously described, we present the results for three different models. First, we consider the impact of the number of negative reviews on the product page. Next, we define the effect of a negative review by looking at the presence of at least one negative review on the product page. Third, we quantify the separate effects of having one, two, three, and four or five negative reviews on the product page.For ease of interpretation, the scale of the choice probability is defined between 0 and 100 and the coefficients can be interpreted as percentage point changes. We use the same re-scaling for all other binary dependent variables.

Looking at the results concerning the main effects of a negative review on purchase behavior, the estimates in Column I indicate that an additional negative review—in place of a positive review—on the product page lowers the probability of purchase by .63 percentage points on average, which translates to an average decrease of 26.87% in purchase probabilities. The results in Column II suggest that consumers who visited the product page when a negative review was present have a 1.16 percentage point lower probability of buying the product compared to those who visited the product page at a time when no negative review was present, which translates to a 41.80% decrease in purchase probability in our data. The findings presented in Column III indicate that the presence of one, two, three, or more negative reviews on the product page results in a significant decrease in the purchase probability, by .93, 1.36, 2.03, or 2.35 percentage points, respectively, when compared to a product page without any negative reviews. These results are in line with the model-free evidence section (Table 2).

In terms of session characteristics, we see that later searched products are less likely to be purchased and that customers who purchased in the category before are more likely to buy again. Although the average rating, number of reviews, and price have the expected signs, they are not statistically significant, which can be explained by the inclusion of product fixed effects. We find no evidence for position effect of the most recent negative review once the number of negative reviews is accounted for.

Table 3: Estimates for the Product Purchase Decision

	Model I		Model II		Model III	
# Negative reviews	-.627***	(.162)				
# Negative review > 0			-1.165***	(.358)		
# Negative reviews = 1					-.933**	(.371)
# Negative reviews = 2					-1.362***	(.461)
# Negative reviews = 3					-2.037***	(.542)
# Negative reviews = 4 or 5					-2.354***	(.781)
Product page information						
# Reviews (log)	.145	(.293)	.085	(.292)	.136	(.293)
Average rating	.383	(.411)	.301	(.428)	.278	(.427)
Price (log £)	-.588	(1.039)	-.609	(1.040)	-.598	(1.040)
Newest negative review at position 2	.119	(.279)	.335	(.302)	.252	(.300)
Newest negative review at position 3	.409	(.366)	.745*	(.399)	.576	(.404)
Newest negative review at position 4	-.144	(.372)	.285	(.420)	.052	(.440)
Newest negative review at position 5	.580	(.524)	.858	(.346)	.482	(.544)
Search and previous choice						
Second viewed product	-.381*	(.203)	-.388*	(.203)	-.382*	(.203)
Third viewed product	-.435*	(.260)	-.442*	(.260)	-.436*	(.260)
Fourth or later viewed product	-.545**	(.216)	-.547**	(.216)	-.545**	(.216)
Previous purchase in category	24.626***	(6.544)	24.691***	(6.547)	24.648***	(6.546)
Product fixed effects	Yes		Yes		Yes	
Day fixed effects	Yes		Yes		Yes	
R^2	.160		.160		.160	
# Observations	55,860		55,860		55,860	

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$

Note: The models present the estimated coefficients from Equation 1 using alternative definitions for the number of negative reviews on the product page at the time of consumer visit (R_{ijt}). The dependent variable is an indicator of product purchase. In Column I, the effect of negative reviews is measured with the total number of negative reviews appearing on the product page. In Column II, the negative review effect is measured with a dummy variable indicating whether there is at least one negative review review on the product page. In Column III, dummy variables indicate the number of reviews on the product page. Robust standard errors clustered at the product level in parentheses.

Prior Information

To evaluate the impact of prior information on the effect of negative reviews, we first evaluate if consumers are more likely to click on product pages with a higher average customer rating and with more reviews. Table 4 shows the results. We find that the effects of these two variables are positive and significant. We obtain similar estimates with and without brand fixed effects (models I and II respectively), suggesting that the effects of the reviews on clicks hold both within and across brands. Hence, as consumers come to the product page and arrive at the negative review section, they are more likely to have seen positive prior information about the product: better average rating and higher popularity as signaled by the number of reviews.

Table 4: Estimates for the Number of Daily Visitors as a Function of Information on the Category Page

	(I) Log Daily Visits		(II) Log Daily Visits	
Average rating	.013***	(.003)	.014***	(.004)
# Reviews (log)	.046***	(.002)	.042***	(.002)
Price (log £)	-.027***	(.005)	-.036***	(.005)
Brand fixed effects	No		Yes	
Category fixed effects	Yes		Yes	
Day fixed effects	Yes		Yes	
R^2	.142		.172	
# Observations	198,537		198,537	

***$p < .01$, **$p < .05$, *$p < .1$

Note: The dependent variable is the logarithm of the product's daily number of visits. Only products that have at least one review are included. Robust standard errors clustered at the product level in parentheses.

Next, to see whether prior information has a moderating role in the effect of negative reviews on product purchase, we re-estimate Equation 1 separately using products below and above the median average rating. In light of our proposed framework, we expect that negative reviews are more impactful when the average rating is higher. Compared to the coefficient of negative reviews in the previous section (of -.627), we observe that the effect of negative reviews is about half (-.341) and less significant for below-median rating products, while it is almost 50% higher for the above-median rating products (-.967).[14]

[14] We experimented with models incorporating interaction terms between the average customer rating and the number of negative reviews on the product page. Our findings support the existence of a "surprise effect," consistent with the results of Vana and Lambrecht (2021).

Table 5: Estimates for the Product Purchase Decision for Products with Low vs. High Average Rating

	Products with Low Average Rating	**Products with High Average Rating**
# Negative reviews	-.341* (.190)	-.967*** (.368)
Product page controls	Yes	Yes
Consumer controls	Yes	Yes
Product fixed effects	Yes	Yes
Day fixed effects	Yes	Yes
R^2	.169	.166
# Observations	27,534	28,326

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$

Note: Estimated coefficients from Equation 1, using products with either lower-than-median (first column) or higher-than-median (second column) average rating. The dependent variable is an indicator of product purchase. Product and consumer control variables included in both models. Consumers who open the second review page or sort the reviews excluded from both models. Robust standard errors clustered at the product level in parentheses.

Information Search Decisions

In this section, we discuss the results regarding different search process decisions, using the estimates from the model with separate effects of one, two, three, or more negative reviews.[15] In the first column of Table 6, we see that the greater the number of negative reviews, the less likely it is that the consumer opens the second review page. The second column shows that, compared to a product page with only positive reviews, a page with one or more negative reviews is more likely to deter the consumer from sorting the reviews (by rating, helpfulness, or other criteria). These findings suggest that the consumer's motivation to seek further information about the product decreases as the number of negative reviews increases.

The third column of Table 6 reveals that the more negative reviews about the focal product a consumer observes, the more willing the consumer becomes to visit the page of a competing product. In fact, even a single negative review on the product page is enough to have a positive and significant impact on the decision to explore alternative items. Consumers

[15] The conclusions drawn from the models with alternative formulations for the effect of negative reviews are substantially similar and are available upon request.

who viewed a product page with one negative review (with the rest of reviews being positive) have a 3.87 percentage-point larger probability of visiting pages of competitor products than consumers who did not face any negative reviews, which translates to a 9.65% increase in search probability. Finding additional negative reviews instead of positive reviews has a cumulative effect on the decision to search for an alternative, with the effect almost doubling in the case of three negative reviews out of five.

As an additional analysis of search behavior, we consider how the search decision depends on the stage of search, in this case the number of product pages visited previously to arriving at the page of the focal product. After splitting our sample into early (when no more than two product pages were previously clicked on) versus late (when at least three products were previously clicked on) arrivals to the product page and re-estimating Equation 5, we find that negative reviews significantly motivated consumers to search for competitors when the product was visited early on but did not significantly impact consumer search decisions for late visits. The respective coefficients are shown in the Web Appendix, Table W5.

The results in the last column of the Table 6 suggest that the presence of negative reviews on the focal product's page influences the decision to scroll to the review area of the competitor product. As the number of negative reviews increases on the focal product's page, consumers become more willing explore the reviews of substitutes. This suggests that they become more cautious and spend more time evaluating suitable alternatives when they encounter negative information, in order to avoid choosing a product that might disappoint.

Table 6: Effects of Negative Reviews on Product and Information Search Decisions

	(I) Open Second Review Page	(II) Sort the Reviews	(III) Search for Substitute	(IV) Scroll to Reviews of Substitute
# Negative reviews = 1	-.652	-1.070***	3.873***	.012
	(.883)	(.293)	(1.090)	(.019)
# Negative reviews = 2	-2.369**	-.977***	5.243***	.038*
	(1.022)	(.349)	(1.378)	(.022)
# Negative reviews = 3	-4.155***	-1.415***	7.128***	.059**
	(1.474)	(.446)	(1.789)	(.027)
# Negative reviews = 4 or 5	-9.466***	-1.439**	11.101***	.172***
	(1.989)	(.722)	(2.897)	(.044)
Product page controls	Yes	Yes	Yes	Yes
Consumer controls	Yes	Yes	Yes	Yes
Product fixed effects	Yes	Yes	Yes	Yes
Day fixed effects	Yes	Yes	Yes	Yes
R^2	.191	.102	.191	.292
# Observations	68,960	68,960	55,860	23,565

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$

Note: Columns I and II show estimates using the sample of consumers who scrolled down to the reviews. Consumers who open the second review page or sort the reviews are excluded from the substitute search model (Column III). The coefficients for review browsing of competitor products are estimated on a sub-sample of consumers who searched for a substitute following visiting the page of the focal product (Column IV). The dependent variables are indicators of the respective consumer decisions. Robust standard errors clustered at the product level in parentheses.

Moderating Effects of Product Type and Review Topic

Product type as moderator

To examine the potential differential effects of negative reviews across utilitarian and hedonic product categories, we conduct separate analyses for each category. Product categories were classified as a function of how they are perceived by consumers (Hirschman and Holbrook 1982). Hedonic products are those that provide a sensory experience of aesthetic or sensual pleasure, such as headphones or home decoration, while utilitarian products are primarily practical in nature, like washing machines or printer ink. As a result of their differing natures, hedonic products tend to be more distinct in their horizontally differentiated dimensions, while utilitarian products have a stronger vertically differentiated dimension.

The results in Table 7 show that negative reviews have a significant impact on reducing the likelihood of purchase for utilitarian products but not for hedonic products.[16] This

[16] This difference cannot be explained by the fact that our data includes more hedonic products in the technology than in the home-and-garden department, as we find significant effects in both departments (Web

finding suggests that consumers of utilitarian products, which are characterized by vertically differentiated attributes and experiential qualities such as durability and product/service quality benefit from access to other users' experiences with the product. This is not the case for hedonic products, where horizontal differentiation and taste differences are more prevalent.

Table 7: Estimates for the Product Purchase Decision for Utilitarian and Hedonic Product Categories

	Utilitarian Products			**Hedonic Products**		
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
# Negative reviews	-.742***			-.319		
	(.190)			(.316)		
# Negative review > 0		-1.506***			-.349	
		(.424)			(.708)	
# Negative reviews = 1			-1.252***			-.168
			(.440)			(.735)
# Negative reviews = 2			-1.734***			-.464
			(.542)			(.903)
# Negative reviews = 3			-2.371***			-1.340
			(.633)			(1.054)
# Negative reviews = 4, 5			-2.811***			-1.056
			(.955)			(1.006)
R^2	.168	.168	.168	.141	.141	.141
# Observations	40,724	40,724	40,724	15,136	15,136	15,136

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$

Note: Estimated coefficients from Equation 1 using alternative definitions for the number of negative reviews on the product page at the time of consumer visit (R_{ij}). The dependent variable is an indicator of product purchase. Product, consumer, and day fixed effects are included in all models. Robust standard errors clustered at the product level in parentheses.

Review topic as moderator

In this section, we present the results of the previously described LDA approach applied for categorizing the negative reviews according to their text. By doing so, we seek to deepen our understanding of what type of review content is particularly responsible for the decrease in purchase likelihood.

Our results for the product purchase decision (Table 8) indicate that negative reviews describing functionality issues or problems related to customer service have a significant

Appendix, Table W6).

negative impact on purchase probability. In contrast, reviews concerning taste-based issues have no influence on purchasing decisions. This finding is consistent with the discussion of product type, as taste-based attributes are more generally considered.

Table 8: Estimates for the Product Purchase Decision as a Function of the Dominant Topic of the Review

	(I)		**(II)**	
	Product Purchase		**Product Purchase**	
# Negative reviews				
on functionality	-.815***	(.257)		
on customer service	-.687***	(.213)		
on matters of taste	-.078	(.331)		
on other topics	-.957*	(.532)		
# Negative review > 0				
on functionality			-.819**	(.337)
on customer service			-1.039***	(.324)
on matters of taste			.029	(.429)
on other topics			-1.073*	(.589)
Product page controls	Yes		Yes	
Consumer controls	Yes		Yes	
Product fixed effects	Yes		Yes	
Day fixed effects	Yes		Yes	
R^2	.160.	.160.		
# Observations	55,860		55,860	

***$p < .01$, **$p < .05$, *$p < .1$

Note: The dependent variable is an indicator of product purchase. Column I shows results from a model in which the independent variables include the total number of negative reviews seen on the product page from a given review topic. Column II shows results from a model in which dummy variables indicate whether on the product page there is at least one negative review from a given topic. Robust standard errors clustered at the product level in parentheses.

To ensure that our findings are not solely driven by differences in star ratings between reviews in different topic categories, we examine whether reviews on taste-related topics have significantly higher star ratings than reviews in the other three topic categories. We find that this is not the case. Negative reviews concerning matters of taste have, on average, similar ratings (1.76) to negative reviews on functionality (1.65), while their frequency is almost the same (Web Appendix, Table W7). Since negative reviews related to functionality were found to significantly decrease purchase probability, we can conclude that our estimates are not solely a result of correlation between star rating and review topic. Instead, our findings suggest that the topic of the review does indeed have an impact on purchase behavior.[17]

Combining moderators: prior information and product type

Finally, we look at the interaction effect between prior information and type of product. To do so, we divide the products into above- and below-median groups with respect to (i) the number of reviews and (ii) average customer rating, resulting in four quadrants of products. Next, we run the purchase regression in each quadrant, separately for hedonic and utilitarian products.

We find that, for utilitarian products, negative reviews have a significantly negative impact on purchase probability, particularly when the product has a high average rating based on a relatively low number of reviews. In this scenario, we observe a decrease of 2.9 percentage points in purchase probability. However, when the product has a high average rating and a high number of reviews, the negative impact of negative reviews is halved, resulting in a decrease of 1.4 percentage points in purchase probability. In contrast, for hedonic products where taste and horizontal differentiation are more important, the effect of negative reviews becomes insignificant.

[17] We also conducted separate regressions on samples of utilitarian and hedonic products. In the case of utilitarian products, reviews on functionality and customer service exhibited significant negative effects. However, for hedonic products, reviews on functionality alone displayed marginal significance.

Table 9: Estimates for the Product Purchase Decision by Prior Information and Product Type

Variable: # Negative reviews	**High Number of Reviews & High Average Rating**		**Low Number of Reviews & High Average Rating**		**High Number of Reviews & Low Average Rating**		**Low Number of Reviews & Low Average Rating**	
Hedonic products	1.46	(1.00)	1.63	(2.04)	-.51	(.56)	-.17	(.46)
Utilitarian products	-1.40***	(.48)	-2.91**	(1.29)	-.45*	(.28)	.20	(.50)

***$p < .01$, **$p < .05$, *$p < .1$

Note: Estimated coefficients from Equation 1 for hedonic vs. utilitarian products using product sub-samples according to the four quadrants based on number of reviews and average rating. The dependent variable is an indicator of product purchase. Product and consumer control variables included in each regression. Consumers who open the second review page or sort the reviews are excluded from each regression. Robust standard errors clustered at the product level in parentheses.

Robustness Tests

To ensure that our results are robust and are not driven by how we have selected our sample or chosen the model specifications, we carried out a variety of robustness checks. Supporting tables are shown in the Web Appendix.

We have derived our main results using linear probability models. It could be that the results are dependent on using this specific *choice of functional form*. Therefore, we re-estimated Equations 1 and 5 using a logistic regression. We observe that negative reviews continue to have a significant impact on purchase and search behavior in the logit specifications (Web Appendix, Table W8). The respective coefficients measuring the impact of negative reviews on consumer behavior are all negative and significant at the 1% level.

It could be that there is an unobserved (to the researcher) relation between the consumer's purchase and product search decisions. In other words, the purchase and search equations could have a contemporaneous *error correlation*. If the correlation of errors is significant but assumed to be zero, the OLS estimation with each equation separately estimated might be biased (Rao 1974). To investigate whether there is significant correlation between the errors of the purchase and search equations, we conducted a seemingly unrelated regression (SUR) analysis that allows for correlated errors (Zellner 1962). We find that the correlation between the residuals is not significant (-.046) and that we obtain substantively identical coefficients using the SUR approach as with estimating the equations independently (Web Appendix, Table W9). We conclude that our findings are not biased due to contemporaneous correlation in the error terms.

One of the concerns when identifying the effect of online ratings on demand is the possibility that parameter estimates may be biased due to *spurious correlation*, that is, when critical reviews are simply evidence of another event occurring at the same time that the review is written (e.g., delivery delays, stock-outs). In these cases, the decline in sales may have occurred even without the existence of the negative review, as consumers may have learned about the issue through other sources. Previous research has encountered similar challenges when measuring the effects of reviews on sales. Reinstein and Snyder (2005) warn of spurious correlation between expert reviews and movie ticket sales induced by an underlying correlation in unobservable quality signals. Chintagunta et al. (2010) address the same issue concerning movie sales and reviews and Anderson and Magruder (2012) account for such correlation in the relationship between average Yelp ratings and sales. We test for spurious correlation by looking at whether a forthcoming negative review, i.e., one that is not yet on the product page but will be there within a week, has an impact on product purchase. If the drop in demand can be attributed to other reasons than the discovery of the negative customer review, a forthcoming negative review would correlate negatively with purchase, and positively with searching for substitutes. We find no evidence for any of these effects (Web Appendix, Table W10), suggesting that spurious correlation is not a concern in our dataset.

Finally, it could be that the time of creation of reviews is related to the total number of consumers who visit the product page during similar periods. An example of this would be a new advertising campaign on national television or through another medium. We run a regression that explains the log daily (and weekly) number of visitors to the product page at the time of review creation as a function of the rating of submitted reviews alongside our standard set of control variables. Our results indicate that there is no significant correlation between the valence of the created review and the number of visitors to the product page around the time of review creation (Web Appendix, Table W11). In other words, we do not find evidence that other *unobserved events* that could drive both the sentiment of the review and the number of product page visitors affected our results. This further strengthens our confidence in the causal relationship between individual negative reviews and product demand.

MANAGERIAL IMPLICATIONS

Mapping Product and Category Vulnerability to Negative Reviews

Using the estimates from the main model (Column III in Table 3), we compute the predicted probability of purchase for each consumer-product pair in our data. We then make a change by adding a negative review to the page of each product and compute the new predicted probability of purchase. With these two values, we estimate the impact of the submission of a negative review among consumers who view the reviews.

Formally, we obtain the purchase probabilities with and without posting an additional negative review using the estimated coefficients through

$$p_{ij} = \mathbb{E}(Purchase_{ij}|Data_{ij}, \hat{\alpha}_1^P, \hat{\beta}_1^P, \hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{\boldsymbol{P}}, \hat{\boldsymbol{\delta}}_{\mathbf{1}}^{\boldsymbol{P}}, \hat{\mu}_{1j}^P, \hat{\eta}_{1i}^P\})$$

and

$$\widetilde{p}_{ij} = \mathbb{E}(Purchase_{ij}|\widetilde{Data}_{ij}, \hat{\alpha}_1^P, \hat{\beta}_1^P, \hat{\boldsymbol{\gamma}}_{\mathbf{1}}^{\boldsymbol{P}}, \hat{\boldsymbol{\delta}}_{\mathbf{1}}^{\boldsymbol{P}}, \hat{\mu}_{1j}^P, \hat{\eta}_{1i}^P\}),$$

where $Data_{ij}$ refers to product characteristics at the time when it was visited by consumer i, while $\widetilde{Data}_{ij}$ contains the same characteristics as the original data except for the addition of a negative review.[18] We consider that the addition of a negative review modifies the product's

[18]We re-ran the analysis using the model where the dependent variable is the number of negative reviews

average rating and total number of reviews. For each product in our data, we calculate the within-product change in purchase probability due to the submission of a negative review by averaging the percentage changes across the N consumers:

$$\Delta_{\rho j} = \frac{\sum_i \left(\frac{\widetilde{p}_{ij}}{p_{ij}} - 1\right) \times 100}{N}.$$

In addition, we calculate the effect of a negative review on the probability of searching for further substitutes using

$$\Delta_{\varrho j} = \frac{\sum_i \left(\frac{\widetilde{\varrho}_{ij}}{\varrho_{ij}} - 1\right) \times 100}{N},$$

where ϱ_{ij} is the probability that consumer i searches for further substitutes without the additional negative review and $\widetilde{\varrho}_{ij}$ is the probability that consumer i searches for further substitutes with the additional negative review present on the page of the focal product. We denote the $(\Delta_{\rho j}, \Delta_{\varrho j})$ values as the product's *purchase and search vulnerability to a negative review.*

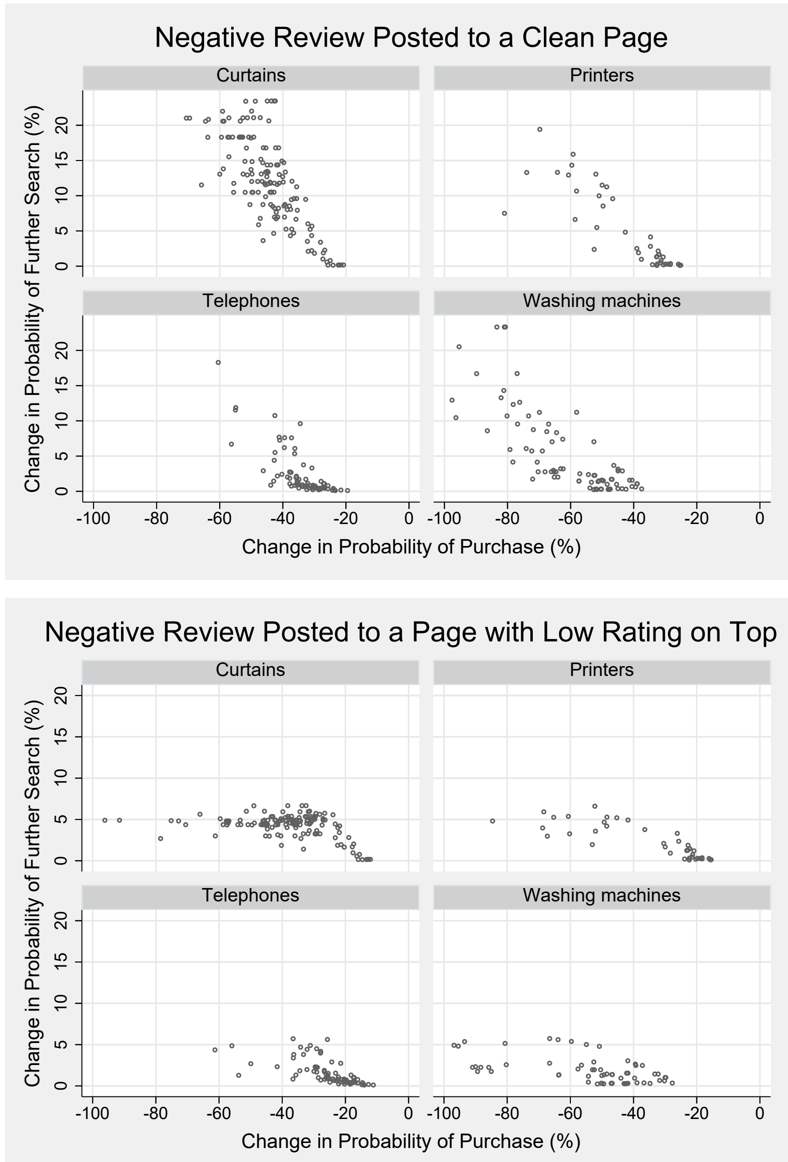
The top panel of Figure 3 shows the vulnerability to a negative review for four product categories as an example, using a sample that consists of products that have no previous negative reviews on the product page at the time of consumer visit. Hence, the negative review is the first and only negative review present on the product's page. The consequences are striking, with the probabilities of purchase declining significantly and the probabilities of searching for substitutes increasing across products in all four categories. Most products are doubly hit by the negative review, with both a reduction of the purchase probability and an increase in the probability of consumers considering alternative products.

We contrast these findings with the impact of the arrival of a negative review when there was already one negative review in the top position in the review area of the product page, using the same set of products as before. The impact of the arrival of this second negative

instead of respective dummies and found substantially similar results.

review is displayed in the bottom panels of Figure 3. The arrival of a second negative review still decreases the probability of purchase, albeit with some variability across product categories. However, the impact on product search is significantly reduced both in terms of average magnitude and variability across products. These findings suggest that the primary and strongest impact of negative reviews on product search is due to the first negative review appearing on the product page, with additional negative reviews having a lesser impact on the likelihood of consumers to browse more.

Figure 3: Impact of a Single Negative Review on Purchase and Further Search for Substitutes with (a) No Negative Review on Product Page and (b) One Negative Review Present in the Top Position



Note: Estimated percentage changes due to the submission of a single negative review. Dots represent mean effects by products in the respective categories. Sample of consumers who scrolled to the review area excluding those who open the second review page or sort the reviews.

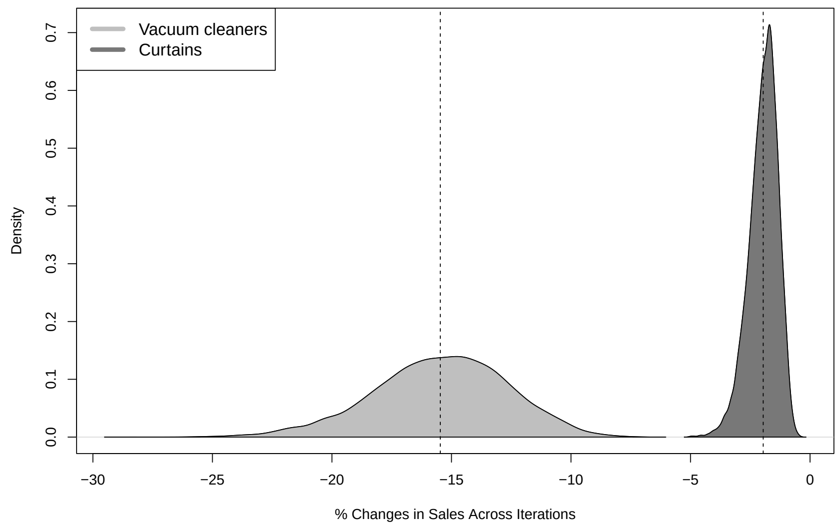
We note that these changes are calculated for consumers who scroll down to the part of the product page that contains the reviews but who do not browse for further review pages. These consumers generated 18.3% of the revenue during the two months in the analysis in the home-and-garden department, and 21.7% in the technology department. Assuming that for the remaining consumers the negative reviews have no effect (as they remain unseen), the overall impact would be about one-fifth of the results presented here. Thus, for most products, the drop in the overall demand, i.e., among all consumers regardless whether they view individual reviews or not, due to a single negative review posted in a product page (previously without negative reviews) would be in the range of 5% to 15%.

Elasticity of Category Sales to a Negative Review

In this section, we propose the metric of *elasticity of category sales to negative reviews*. To obtain this elasticity, we calculate the percentage change in sales in a product category as a function of a 1% increase in the total number of negative reviews on the product pages within that category, using the coefficients from Column III in Table 3.

In other words, we increase the negative reviews available on product pages by 1%, added randomly to product pages within the respective category. We record the implied change in sales at the category level due to the addition of these negative reviews, and repeat the entire process 10,000 times for a given product category. Figure 4 presents an example, displaying the resulting distribution of percentage changes in two product categories: vacuum cleaners and curtains. We obtain an elasticity of -15.4 for vacuum cleaners and of -1.9 for curtains.

Figure 4: Simulated Sales Changes due to a 1% Increase in the Number of Negative Reviews on Product Pages



Note: Sales changes estimated across 10,000 iterations, using the sample of consumers who scrolled to the reviews excluding those who open the second review page or sort the reviews.

To better understand how the negative review elasticity of sales is related to the product category, we regress the elasticity metrics on product and consumer variables (Web Appendix Table W12). The independent variables are: mean consideration set size of consumers who visit the category, mean average product rating in the category, mean log price in the category, mean log number of reviews for a product in the category, mean number of negative reviews on product pages in the category, a dummy for hedonic product, a dummy for technology product, the length of the product description, and an interaction term between hedonic and consideration set size.[19] Categories with higher prices and longer product descriptions tend to be more elastic, suggesting that consumers are particularly sensitive to negative reviews when shopping for expensive products and in categories where extensive information is available on the product pages. Additionally, we find that the negative review elasticity of sales has a strong correlation with the average consideration

[19] As all elasticities are negative, we use the absolute value as dependent variable for an easier interpretation. To avoid bias due niche products, we drop product categories that were visited by fewer than 50 consumers.

set size in the category, defined as the number of products consumers look at during their online session. This suggests that negative reviews have a greater impact on demand as competition increases (i.e., when consumers consider a larger set of products).

CONCLUSION

The popularity of online shopping and the significance of user-generated content have been increasing over the years. There are various instances where online sellers request customers not to leave negative reviews and offer their services to resolve any issues in exchange. Our study shows that sellers have a valid reason for doing so as even a single negative review can significantly reduce the likelihood of purchase. The impact on demand of multiple negative reviews is even more substantial.

Our study has shown that negative reviews not only affect product purchase, but they also lead consumers to search for substitute products. By increasing the consideration set, this extended search decreases the likelihood of the original product being selected at the end of the consumer journey. When consumers find negative reviews, consumers are more likely to stop browsing for further reviews about the focal product, suggesting that they are unwilling to spend more time and effort collecting information on that particular item.

We have also found that negative reviews have varying impacts on consumer behavior depending on the product category and the content of the reviews. In utilitarian product categories, negative reviews have a much greater effect on purchase likelihood compared to hedonic categories. Additionally, negative reviews that focus on product functionality or customer service are more likely to deter consumers from purchasing, while those that discuss taste-based preferences such as design, color, or material have little to no effect. These findings can provide valuable insights for online sellers looking to improve their product offerings and manage their online reputation.

Using our approach, the effect of a single negative review on purchase probability and competitor search probability can be established at the product level and represented on a two-dimensional vulnerability map. We have introduced a metric for the percentage change in sales caused by a 1% increase in the number of negative reviews in the category. This metric could serve as a reference point for those interested in understanding how online WOM affects sales in their relevant category.

We believe that our findings may motivate future research such as modeling how the design of product pages might change what consumers decide to browse for and eventually purchase. Another potentially fruitful direction could be to study whether and how low ratings influence the arrival frequency of new reviews through decreased purchase rates. Furthermore, the literature might benefit from a better understanding of why and under what circumstances consumers decide to view reviews.

References

Allenby, Greg M and James L Ginter (1995). "The effects of in-store displays and feature advertising on consideration sets". In: *International Journal of Research in Marketing* 12.1, pp. 67–80.

Anderson, Michael and Jeremy Magruder (2012). "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database". In: *The Economic Journal* 122.563, pp. 957–989.

Angie, Amanda D et al. (2011). "The influence of discrete emotions on judgement and decision-making: A meta-analytic review". In: *Cognition & Emotion* 25.8, pp. 1393–1422.

Babić Rosario, Ana et al. (2016). "The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors". In: *Journal of Marketing Research* 53.3, pp. 297–318.

Babin, Barry J, William R Darden, and Mitch Griffin (1994). "Work and/or fun: measuring hedonic and utilitarian shopping value". In: *Journal of consumer research* 20.4, pp. 644–656.

Baumeister, Roy F et al. (2001). "Bad is stronger than good". In: *Review of general psychology* 5.4, pp. 323–370.

Bigne, Enrique, Kalliopi Chatzipanagiotou, and Carla Ruiz (2020). "Pictorial content, sequence of conflicting online reviews and consumer decision-making: The stimulus-organism-response model revisited". In: *Journal of Business Research* 115, pp. 403–416.

Blei, David M, Andrew Y Ng, and Michael I Jordan (2003). "Latent dirichlet allocation". In: *the Journal of machine Learning research* 3, pp. 993–1022.

Brown, Tom J et al. (2005). "Spreading the word: Investigating antecedents of consumersâ positive word-of-mouth intentions and behaviors in a retailing context". In: *Journal of the Academy of Marketing Science* 33.2, pp. 123–138.

Cacioppo, John T., Wendi L. Gardner, and Gary G. Berntson (1997). "Beyond bipolar conceptualizations and measures: The case of aattitude and evaluative space". In: *Personality and Social Psychology Review* 1(1), pp. 3–25.

Chen, Yubo and Jinhng Xie (2008). "Online consumer review: Word-of-mouth as a new element of marketing communication mix". In: *Management Science* 54(3), pp. 477–491.

Cheong, Hyuk Jun and Margaret A Morrison (2008). "Consumersâ reliance on product information and recommendations found in UGC". In: *Journal of Interactive Advertising* 8.2, pp. 38–49.

Chevalier, Judith A. and Dina Mayzlin (2006). "The effects of word of mouth on sales: Online book reviews". In: *Journal of Marketing Research* 43(3), pp. 345–354.

Chintagunta, Pradeep K., Shyam Gopinath, and Sriram Venkataraman (2010). "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets". In: *Marketing Science* 29(5), pp. 944–957.

De los Santos, Babur, Ali Hortacsu, and Matthijs R. Wildenbeest (2012). "Search model of consumer search using data on web browsing and purchasing behavior". In: *The American Economic Review* 102(6), pp. 2955–2980.

Demuynck, Thomas and Christian Seel (2018). "Revealed preference with limited consideration". In: *American Economic Journal: Microeconomics* 10.1, pp. 102–31.

Eliaz, Kfir and Ran Spiegler (2011). "Consideration sets and competitive marketing". In: *The Review of Economic Studies* 78.1, pp. 235–262.

Feick, Lawrence and Robin A Higie (1992). "The effects of preference heterogeneity and source characteristics on ad processing and judgements about endorsers". In: *Journal of Advertising* 21.2, pp. 9–24.

Fiske, Susan T (1980). "Attention and weight in person perception: The impact of negative and extreme behavior." In: *Journal of personality and Social Psychology* 38.6, p. 889.

Floyd, Kristopher et al. (2014). "How online product reviews affect retail sales: A meta-analysis". In: *Journal of Retailing* 90.2, pp. 217–232.

Fung, Raymond and Matthew Lee (1999). "EC-trust (trust in electronic commerce): exploring the antecedent factors". In.

He, Sherry, Brett Hollenbeck, and Davide Proserpio (2022). "The market for fake reviews". In: *Marketing Science* 41.5, pp. 896–921.

Hirschman, Elizabeth C and Morris B Holbrook (1982). "Hedonic consumption: emerging concepts, methods and propositions". In: *Journal of marketing* 46.3, pp. 92–101.

Honka, Elisabeth (2014). "Quantifying search and switching costs in the US auto insurance industry". In: *The RAND Journal of Economics* 45.4, pp. 847–884.

Honka, Elisabeth and Pradeep Chintagunta (2016). "Simultaneous or sequential? search strategies in the us auto insurance industry". In: *Marketing Science.*

Hung, Kineta H and Stella Yiyan Li (2007). "The influence of eWOM on virtual consumer communities: Social capital, consumer learning, and behavioral outcomes". In: *Journal of advertising research* 47.4, pp. 485–495.

Ito, Tiffany A. et al. (1998). “Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations”. In: *Journal of Personality and Social Psychology* 75(4), pp. 887–900.

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

— (2016). “The probit choice model under sequential search with an application to online retailing”. In: *Management Science*.

Klein, Jill Gabrielle (1996). “Negativity in impressions of presidential candidates revisited: The 1992 election”. In: *Personality and Social Psychology Bulletin* 22.3, pp. 288–295.

Kozinets, Robert V. et al. (2010). “Networked narratives: Understanding word-of-mouth marketing in online communities”. In: *Journal of Marketing* 74(2), pp. 71–89.

Liu, Xiao, Dokyun Lee, and Kannan Srinivasan (2017). “Large Scale Cross Category Analysis of Consumer Review Content on Sales Conversion Leveraging Deep Learning”. In: *NET Institute Working Paper No. 16-09.*

Liu, Yong (2006). “Word of mouth for movies: Its dynamics and impact on box office revenue”. In: *Journal of marketing* 70.3, pp. 74–89.

López, Inés López and Salvador Ruiz De Maya (2012). “When hedonic products help regulate my mood”. In: *Marketing Letters* 23.3, pp. 701–717.

Manzini, Paola and Marco Mariotti (2018). “Competing for attention: is the showiest also the best?” In: *The Economic Journal* 128.609, pp. 827–844.

Markopoulos, Panos M and Eric K Clemons (2013). “Reducing buyers’ uncertainty about taste-related product attributes”. In: *Journal of Management Information Systems* 30.2, pp. 269–299.

Mudambi, Susan M. and David Schuff (2010). “What makes a helpful online review? A study of customer reviews on Amazon.com”. In: *MIS Quarterly* 34(1), pp. 185–200.

Rao, Potluri (1974). “Specification bias in seemingly unrelated regressions”. In: *Econometrics and Economic Theory*. Springer, pp. 101–113.

Reinstein, David A and Christopher M Snyder (2005). “The influence of expert reviews on consumer demand for experience goods: A case study of movie critics”. In: *The journal of industrial economics* 53.1, pp. 27–51.

Rozin, Paul and Edward B Royzman (2001). “Negativity bias, negativity dominance, and contagion”. In: *Personality and social psychology review* 5.4, pp. 296–320.

Salminen, Joni et al. (2022). “Creating and detecting fake reviews of online products”. In: *Journal of Retailing and Consumer Services* 64, p. 102771.

Sen, Shahana and Dawn Lerman (2007). “Why are you telling me this? An examination into negative consumer reviews on the web”. In: *Journal of interactive marketing* 21.4, pp. 76–94.

Skowronski, John J and Donal E Carlston (1989). “Negativity and extremity biases in impression formation: A review of explanations.” In: *Psychological bulletin* 105.1, p. 131.

Sun, Monic (2012). “How does the variance of product ratings matter?” In: *Management Science* 58.4, pp. 696–707.

Tirunillai, Seshadri and Gerard J Tellis (2014). “Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation”. In: *Journal of marketing research* 51.4, pp. 463–479.

Ursu, Raluca M (2018). “The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions”. In: *Marketing Science* 37.4, pp. 530–552.

Ursu, Raluca M, Qingliang Wang, and Pradeep K Chintagunta (2020). “Search duration”. In: *Marketing Science* 39.5, pp. 849–871.

Van Nierop, Erjen et al. (2010). “Retrieving unobserved consideration sets from household panel data”. In: *Journal of Marketing Research* 47.1, pp. 63–74.

Vana, Prasad and Anja Lambrecht (2021). “The effect of individual online reviews on purchase likelihood”. In: *Marketing Science*.

Voss, Kevin E, Eric R Spangenberg, and Bianca Grohmann (2003). "Measuring the hedonic and utilitarian dimensions of consumer attitude". In: *Journal of marketing research* 40.3, pp. 310–320.

Wu, Chunhua et al. (2015). "The Economic Value of Online Reviews". In: *Marketing Science* 5.34, pp. 739–754.

Xie, Xiao-Fei et al. (2011). "The role of emotions in risk communication". In: *Risk Analysis: An International Journal* 31.3, pp. 450–465.

You, Ya, Gautham G Vadakkepatt, and Amit M Joshi (2015). "A meta-analysis of electronic word-of-mouth elasticity". In: *Journal of Marketing* 79.2, pp. 19–39.

Zellner, Arnold (1962). "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias". In: *Journal of the American statistical Association* 57.298, pp. 348–368.

Zhang, Kem ZK et al. (2014). "Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model". In: *Decision Support Systems* 67, pp. 78–89.

Zhong, Ning and David A Schweidel (2020). "Capturing changes in social media content: A multiple latent changepoint topic model". In: *Marketing Science* 39.4, pp. 827–846.