The Effect of Individual Online Reviews on Purchase Likelihood

Prasad Vana

Anja Lambrecht*

7th September 2020

* Prasad Vana (prasad.vana@tuck.dartmouth.edu) is Assistant Professor of Marketing, Tuck School of Business, 100 Tuck Hall, Hanover, NH 03755, United States of America. Anja Lambrecht (alambrecht@london.edu) is Professor of Marketing, London Business School, Regent's Park, London NW1 4SA, United Kingdom. The authors thank Kusum Ailawadi, Puneet Manchanda, Nicos Savva and Scott Neslin for comments on earlier versions of the paper.

The Effect of Individual Online Reviews on Purchase Likelihood

ABSTRACT

Online product reviews constitute a powerful source of information for consumers. Past research has studied the effect of aggregate measures of reviews (such as, average product rating and number of reviews) on consumer behaviour. In this study, we investigate how individual reviews displayed on a product webpage affect consumers' purchase likelihood. Identifying this effect is challenging because retailers are free to select which reviews to display on the product page and in what order, making the display of reviews in particular positions potentially endogenous. We address this challenge by utilizing an empirical context where the retailer displays reviews by recency and exploit the variation in review positions generated as newer reviews are added on top of older ones. We find that individual reviews have a strong effect on consumer purchase decisions. These effects are particularly pronounced when individual reviews help consumers resolve uncertainty about the product or contrast with the aggregate information that is instantly available on the product page. (160 words)

Keywords: Online product reviews, electronic commerce, endogeneity

1 Introduction

Online product reviews are a powerful source of information for consumers purchasing at online retailers. Consumers use reviews to learn about the quality of products (Zhao et al. 2013) or about their preferences (Wu et al. 2015). For firms, reviews are important because they may impact sales (Chintagunta, Gopinath, and Venkataraman 2010; Chevalier and Mayzlin 2006).

Online retailers typically display aggregate measures such as the average rating of the product and the number of reviews available for the product at the top of a product's webpage. Further down, consumers find individual reviews, often by a large number of consumers. Past research on the effect of reviews on consumer behavior has focused on how aggregate measures of reviews such as the average rating and the number of reviews affect outcomes such as sales (see, for example, Chintagunta, Gopinath, and Venkataraman 2010; Duan, Gu, and Whinston 2008; Liu 2006; Chevalier and Mayzlin 2006; Godes and Mayzlin 2004) but little is known about the effect of individual reviews displayed on a product page.

Understanding the role of individual reviews on purchasing is important: while up to 86% of consumers read online reviews,[1] consumers read less than 10 individual reviews before forming an opinion about a product (BrightLocal 2018) – despite many products having tens or even hundreds of reviews. Consumers' purchase decisions may hence be particularly influenced by the few reviews that consumers are exposed to – such as those that are clearly displayed on the product page. Here, the most prominently displayed feature is a review's star ratings which is easily visible and highly salient to even a cursory browser. If individual reviews should affect a consumer's purchase decision above and beyond the average rating of a product, this may potentially give the retailer, who selects which reviews to display on the product page, considerable power over consumer choice.

[1] The number refers to online reviews for local businesses.

In this research, we ask how the star rating of individual reviews affects consumers' purchase decisions above and beyond a review's aggregate rating and how this effect varies with the position in which a review is displayed. We identify the effect of individual reviews on purchases using a unique dataset from an online retailer situated in the United Kingdom.

A challenge in identifying the effect of individual reviews on purchases can potentially arise if the selection of reviews that are displayed and the order in which they are listed is endogenous: the retailer might employ an algorithm to determine the selection and order of reviews with the objective of influencing a consumer's purchase decision in a desired way.[2] In such a case, a researcher, who does not know the algorithm used, is unable to disentangle the effect of a review and its position from the star-rating and the content of a review on sales (omitted variable bias). By contrast, in our context the retailer does not influence which reviews are displayed to consumers, or in which position, but instead sorts them by recency. As a result, the addition of a new review serves as an exogenous shock that shifts all existing reviews down by one position, allowing us to identify the effect of a review's star rating and position on purchases. We confirm in various tests that the timing of this shift in position is unrelated to the rating and position of all other reviews as well as to the number of new customers arriving at the product's site.

We examine the effect of a review's star-rating and position on the product page on purchase likelihood using two alternative approaches. First, we leverage the full data that cover an individual review's multiple shifts across the different positions on the product page. Thus, the data cover a large number of online reviews for a large variety of products across many consumer visits to the website. One potential disadvantage, however, of this approach is that products that are more frequently viewed by consumers would be overrepresented

[2] For details on the algorithms used by Amazon, Yelp and Airbnb, see https://www.geekwire.com/2015/amazon-changes-its-influential-formula-for-calculating-product-ratings/, https://www.yelp-support.com/article/How-is-the-order-of-reviews-determined?l=en_US, and https://www.airbnb.com/help/article/1885/why-arent-my-reviews-shown-in-order respectively.

relative to those that are rarely viewed. Additionally, since the data span up to two months for each individual review, it may be difficult to control for unobserved variables that change during this time period.

Second, we therefore report the results of a regression discontinuity analysis. Here, we focus on a narrower time-window around the addition of a new review and compare the probability of purchase before the addition of a new review in first position to that after the new review's addition. While the first approach allows us to utilize more data points, the second means we can more cleanly identify the effect of the addition of a new review.

Throughout, conditional on a consumer viewing individual reviews, we find that individual reviews have a strong effect on consumer purchase decisions, above and beyond the average rating of a product and the number of ratings available. Specifically, both our approaches indicate that this effect is most pronounced for replacing a 1-star review in the first position with a 5-star review.

The fact that given the average rating of a product and the number of ratings it has obtained, individual reviews still affect consumers' purchases is surprising. We therefore ask when the impact of individual reviews is most pronounced. We propose that individual reviews affect purchase probability either when they resolve uncertainty about product quality or when these reviews contrast with information otherwise displayed. Our results support these patterns. First, we find that indeed reviews matter most when there is limited information available on the product page. Similarly, reviews have a greater impact on purchases when the product's characteristics are likely to appeal only to a subset of customers, measured by the variance in ratings. Both analyses indicate that consumers use individual reviews to resolve uncertainty. Second, our evidence supports that reviews are of value to consumers when they contrast with information that consumers may otherwise gather on the website. We find that individual reviews with a high star-rating are particularly effective when the average rating of the product is relatively low. Similarly, positive reviews increase purchase likelihood particularly for low-priced products, that is, instances when the price is likely to signal lower product quality.

Our research contributes primarily to a literature on how online reviews affect consumer decisions. Early research that related aggregate sales data to the average rating and the volume of reviews available for a product found that higher average ratings increase sales of movies (Chintagunta, Gopinath, and Venkataraman 2010), video games (Zhu and Zhang 2010) and books (Chevalier and Mayzlin 2006). More recently, Zhao et al. (2013) showed that consumers learn more from a book's average rating by reviewers than from their own experience with other books in the same genre. Wu et al. (2015) investigate the value of reviews for both consumers and restaurants. Ho-Dac, Carson, and Moore (2013) find that reviews matter less for strong brands than for weaker brands. Ludwig et al. (2013) find that changes in positive affective content of reviews have a smaller effect on subsequent conversion rates than changes in negative affective content. However, while the recent papers use data on individual reviews, they assume that consumers read all reviews that are available for a product – an assumption that does not necessarily reflect real-life behavior. Instead, we focus on only those reviews that were displayed to a consumer when they visited the retailer's website and the effect of an individual review's star rating and its position in the list of reviews on the consumer's purchase decision.

Our research also relates to a recent stream of literature on how the position of an item on a list affects choice. Ursu (2018) finds that rankings affect what hotels consumers search but not purchases, conditional on having searched.. De los Santos and Koulayev (2017) show that customized rankings can increase click-through rates for hotels. Jeziorski and Moorthy (2017) find that positions of search ads and advertiser prominence are substitutes. Relatedly, Narayanan and Kalyanam (2015) show that position effects exist in only some positions and are stronger when the advertiser is smaller and the consumer has less prior experience with the keyword for the advertiser. Wu et al. (2015) use observational data to document the relationship between review position and perceived accuracy of a review. We contribute to this literature that has largely abstracted from the role of reviews by demonstrating that the position in which reviews are displayed can significantly affect purchases. We also provide evidence that this may be most important when reviews either resolve uncertainty about quality or when information in reviews contrast with that displayed on the product page.

Our results have implications for retailers and policy makers. They suggest that a retailer's choice of which reviews to display on the first page influences purchase likelihood. This effect size is comparable with the effect sizes of typical online marketing actions such as online display ads, email and mobile coupons. Retailers should be aware that listing reviews by recency might inadvertently keep consumers from buying an objectively good product, based on high average ratings, simply because a recent customer disliked it. Such potentially unintended effects of individual reviews on consumer purchase decisions may be especially relevant for retailers who currently display reviews in chronological order rather than algorithmically. [3] However, the patterns we find also suggest that when retailers present in-depth product information on their website, the importance of individual reviews may be attenuated. If this is the case, displaying more granular information on a product's page may help consumers avoid putting disproportionate weight to individual reviews.

Our results are further relevant in the context of a broader debate on how algorithms shape consumer decisions (Lambrecht and Tucker 2019). A worry from a policy maker's viewpoint could be that firms could manipulate consumers' purchase decisions by strategically adjusting the display of reviews. We leave it to future research to explore

[3] Many retailers currently display reviews in chronological order. At the time of writing, this included Gap and Crate and Barrel in the US, John Lewis and Carphone Warehouse in the UK and MediaMarkt in Germany.

whether such concerns would be warranted or whether consumers would learn and adjust their behavior. At the same time, our results that consumers allocate disproportionate weight to individual reviews suggests that even in the absence of deliberate manipulation, a simple algorithm that is blind to the favorability of a review might still result in consumers making decisions that are not in their best interest: even if a consumer might objectively benefit from a product, an unfavourable review displayed prominently could dissuade them from purchasing. Thus, while policy makers increasingly turn their attention towards the effect of algorithms on consumers' choices, demanding, for example, that firms make algorithms public,[4] our results emphasize the importance of nuance in such debates.

2 Empirical Setting and Data

2.1 Empirical Setting

The data cover all products listed under the "Technology" (e.g. laptops/desktops, TVs, cameras, and mobile phones) and "Home and Garden" (e.g. furniture, kitchen/bathroom ware, electricals, and garden furniture/ware) sections of a major online retailer in the United Kingdom.[5] Other productst that the retailer sells are not part of the data set.

When a consumer searches for a product on the retailer's website, the retailer first returns a webpage that displays a list of relevant items from which the consumer can then click to a specific product page. On the product page, the retailer displays a picture of the product along with a product description and the product's price. On top of the page, and instantly visible to the consumer, the website displays the average rating of the product based on customer reviews as well as the number of reviews collected as of that date.

[4] See for example http://www.pcworld.com/article/2908372/the-ftc-is-worried-about-algorithmic-transparency-and-you-should-be-too.html, https://www.ftc.gov/news-events/blogs/techftc/2015/03/booting-new-research-office-ftc and https://www.thelocal.de/20161026/merkel-demands-transparency-from-internet-giants.

[5] The data was provided as part of a data grant by the Wharton Customer Analytics Initiative (WCAI). A non-disclosure agreement with WCAI prevents us from naming the retailer.

When a consumer scrolls down the product page, they see the the star rating of individual reviews (one to five stars) and, mostly, a brief qualitative product assessment (see Figure 1). A product page displays up to five reviews, sorted by recency. If a consumer clicks on a link to view the next page of reviews, up to twenty reviews are displayed per page.

All reviews are hosted by a third party firm. Only customers who purchased the product can review: after a purchase, customers receive an email inviting them to write a review. The review is submitted to the third-party firm that vets reviews to detect inappropriate content and is responsible for the display of reviews on the retailer's website.

2.2 Data

The retailer provided three data sets collected between February 1 and March 31, 2015. The first dataset consists of consumer browsing data. Two variables identify individual consumers. Consumers logged into the retailer's website are identified based on their login details. Consumers not logged in are identified based on the computer's IP address and tracked across browsing sessions using cookies.[6] The full data has 126,375 unique user IDs and 46,170 unique IP addresses. The data records consumer activity on the website based on "sessions". An hour of inactivity on an open page is noted as the end of the session.

For each product page in each session (i.e., a "product-session"), the firm records through a pixel tracker whether a consumer scrolled down to the part of the webpage that displays individual reviews. Since our research focuses on the effect of individual reviews on purchase, we limit our data to sessions where consumers scrolled down the product page to view individual reviews, resulting in 1,078,403 impressions of reviews.

Since we focus our analysis on the first page review impressions, we remove from the data impressions of the second or any further pages of reviews, resulting in 628,479 review impressions. Even though this reduces the total number of impressions by about a third, the

[6] If a consumer does not allow cookies, the retailer tracks them within but not across browsing sessions.

impressions removed stem from only 15.41% of sessions. To not bias the estimate of the effect of a review position when there is in fact no review, we remove 58,433 impressions where the product page displayed less than five reviews.

Additionally, we remove impressions where the average product rating was missing (56,412 impressions), where one or more of the five individual reviews displayed was missing a rating (31,536 impressions), or where the data had other inconsistencies, leaving us with a dataset of 446,260 impressions. Finally, we remove all products that did not have a new review added in the two-month data period and so had no shift in the position of reviews. This final dataset has 380,450 impressions. This dataset forms the basis of the first of our two sets of analyses which focuses on an individual review's multiple shifts (MS) across the different positions on the product page. In total, 82.05% of ratings have 4- or 5-stars, a pattern which is in line with previous research (see, for example, Chevalier and Mayzlin 2006); see Section A1 of the Web Appendix for details of the distribution of star ratings.

Our regression discontinuity approach (RD) focuses on the short time window around the addition of a new review in the first position on the product page, the only position where we observe a new review that previously was not present on the product page. To set up the data, we start with all the product-sessions in the MS dataset created above. We next identify the product-sessions that occurred immediately before the addition of a new review X (with say, reviews A, B, C, D, and E displayed in that order) and the sessions that occurred immediately after the addition of X (with reviews X, A, B, C, and D displayed in that order). Thus, the sessions before and after the addition of a new review have four reviews in common on the product page, though in different positions. In our main analyses, we use a 14-day period before and a 14-day period after the addition of the review and limit our data to only those cases where we have at least one product-session in both the two weeks before and in the two weeks after the addition of the review. If two successive reviews are added within the 14-day window, we limit our observations to the shorter time-window.[7]

To cleanly identify the effect of the addition of a new review, we discard instances where two or more new reviews were added between two successive sessions of that product. We also limit our analysis to changes in reviews for which we observe at least one browsing session before and one session after the addition of the new review. The resulting dataset has 45,363 sessions from the addition of 7,184 new reviews and covers 2,612 products. Table 1 shows the distribution of reviews before and after the addition of a new review in the first position. Again, many new reviews added are positive; 54% have 5-star ratings.

We estimate the discontinuity in purchases that arises from a change in star ratings. Such effects may vary with the type of change (e.g. a change from a 2-star to a 4-star rating may have a different effect than from a 2-star to a 5-star rating), requiring us to focus on changes with a sufficiently large number of observations. As 5-star ratings are most frequent, we focus on observations around instances when a 5-star rating in position 1 is replaced by a 1-, 2-, 3- or 4-star rating. This final dataset has 8,820 sessions. Note that it is not possible to focus instead on the addition of a 1, 2, 3 or 4-star review in position 1 pushing a 5-star review into position 2. In this case, we would have no variation in the rating in position 2, making the star-rating in position 2 simultaneous with the treatment. We would then be unable to control for it and it would be impossible to identify the effect.[8]

The second data set records orders placed. We match this data to the browsing data based on login information, user ID and timestamp. Thus, our data records for any impression

[7] See Section A2.1 of the Web Appendix for details on data preparation steps to account for sessions where we have time-windows shorter than 14-days due to successive reviews added within 14-days of the last review.
[8] For a similar reason, we are unable to measure in this framework the effect of a change in positions 3, 4 and 5. For example, if we were to try and measure the effect of a shift of a 1-star review in position 3 to position 4, there would be no variation in the star ratings in position 4.

of the review part of the website whether the consumer purchased the product. It gives us 2,833 purchases for the MS analysis and 2,000 purchases for the RD setup.[9]

Third, the retailer provided a review's star rating, content and the date it was first displayed. We map review details to browsing sessions. On average, a product has 4.24 stars, 146.53 reviews and costs £73.52[10] (Table 2). An average review has 30.63 words.

2.3 Empirical Strategy

Our identification, in both the MS and RD approach, relies on the retailer's policy to order reviews by recency, pushing existing reviews down the list as a new review is added. We argue that this policy creates an exogenous shift in a review's position that enables us to identify the causal effect of a review's position on purchase. To illustrate, consider Figure 2, which shows the same review page as Figure 1, but after a new review has been added and all existing reviews shifted down by one position. The key assumption is that the date when this review is added is exogenous (as it is unrelated) to the positions and star ratings of the reviews already displayed. This identification approach is broadly similar to the one used in Ghose, Goldfarb and Han (2012). The difference between our approach and theirs lies in the fact that in our case the order of reviews is determined chronologically and shifts as new reviews are added. By contrast, in Ghose, Goldfarb and Han (2012) a new post appears in different positions for different users, depending on traits that vary across those users.

In the MS approach, we identify the effect of the position of a review on a consumer's purchase decision by comparing the purchase probability when a particular review is in position one with the purchase probability when the same review is in position two, controlling for other covariates such as the ratings of the other four reviews on the page. A similar approach applies to all other positions. This allows us to use data from a large variety

[9] We do not include purchases by consumers who did not scroll down to the review part of the product as our focus is the effect of individual reviews on consumer behaviour.
[10] There is very little price variation in the data as price is only reported when customers purchased so that for some products we rarely observe prices. In our data, only 6.77% of products ever change price.

of products across a large number of online reviews covering multiple shifts in review positions throughout the duration of the data.

However, the MS approach has two shortcomings. First, there is heterogeneity in the frequency with which new reviews are added to product pages. For example, if a product receives a single new review in our data period, then all sessions from the entire time period are used to identify the effect of adding the new review. By contrast, for a product that receives a new review every two days, product sessions from only two days identify the effect of a particular review in a specific position since later (or earlier) sessions are associated with other shifts in reviews. This imbalance may lead to the estimation of effects for some products to be based on more sessions relative to others. Second, to the extent that the period before and after the change in reviews is long, it is possible that unobserved factors that we do not control for affect the results.

Our regression discontinuity design addresses these concerns. Here, the identifying assumption is that as we focus on a specific time window before and after the addition of a new review, all variables that may affect the purchase decision (other than the reviews displayed on the page) experience little variation during this time period. Any change observed in the purchase rate between the session before and after the addition of a review can hence be attributed to the change in the display of reviews.

As explained in Section 2.2, our identification relies on the variation in star ratings in position 1 introduced by the addition of a 5-star rating in this position that pushes a 1-, 2-, 3- or 4-star review into the next position. We thus control for the average star rating of reviews in positions two to four. Note that we are unable to alternatively control for each individual rating as those in position 2 are collinear with the change in position 1 (e.g., for replacing a 3-star review with a 5-star review, all reviews in position 2 would then have 3 stars).

While the RD approach in some ways improves identification, it significantly restricts our use of the data. We focus on sessions that fall into a 14-day window on either side of the addition of a new review. We also need to observe at least one session before and after the addition of the new review. Lastly, we can only focus on the discontinuity arising from adding a 5-star review to position one which pushes a 1, 2, 3 or 4-star review into position 2. We will therefore report results from both estimation approaches.

We discuss four concerns that may affect our identification. First, it could be that the selection of reviews that are displayed and the order in which they are listed is endogenous: the retailer might employ an algorithm to determine the selection and order of reviews with the objective of influencing a consumer's purchase decision. In such a case, a researcher, who does not know the algorithm, is unable to disentangle the effect of a review and its position from the star-rating and the content of a review on sales (omitted variable bias). This is not a concern in our setting since the retailer in our empirical context does not employ such an algorithm but sorts reviews by recency.

Second, there may be factors that are not product-specific that drive both the addition of new positive or negative reviews to a product's page and the purchase likelihood of that product. If that were the case, any relationship we might measure could simply be an outcome of an exogenous shock that influences both variables rather than a causal effect of a review's position on purchases. In our context, such a relationship seems unlikely: First, there are no obvious seasonalities during our data period that would drive both the arrival of many new customers and the writing of reviews for recently purchased items. Second, there is a time lag between purchasing an item and a review being posted simply because the item needs to be shipped, used, the review submitted and then vetted by the third party firm that manages the reviews.[11] Notwithstanding, we conduct a check to rule out that other factors drive both the addition of new reviews to a product's page and the arrival of customers on that product's page. We estimate a regression using as dependent variable the average score of newly submitted reviews. Column (I) of Table 3 illustrates that the average rating of reviews newly displayed for a product on a particular day is not directly related to the number of people browsing that day.

Third, if the timing of the addition of new reviews were related to the star rating of a previous review, this could affect the sample size to identify a particular effect and so the significance of that estimate. We check whether indeed the time taken for the addition of a new review depends on the star-rating of the reviews displayed on the product page. We estimate a Cox proportional hazard model with the time taken until the addition of a new review as the dependent variable and the star rating of the reviews on the product page as the independent variables of interest. We control for the average rating and the number of reviews of the product. We control for product-specific heterogeneity using a stratified baseline hazard approach (Prentice and Gloeckler 1978). Of the 7,184 new reviews added, we include the 6,302 reviews in this analysis for which the timing, average rating, and number of reviews were available. Column (II) of Table 3 shows that the time until a new review is added does not depend on the star rating of the reviews on the product page. [12] Column (III) displays similar results from an OLS estimation. We conclude that the star-rating displayed on the first page of reviews does not affect the time until a new review is added.

Fourth, there could be medium to long-term changes to product attributes over the product's lifecycle that might simultaneously affect the pattern of reviews and purchases. These could include technological advances or shifts in consumer tastes, which may lead to

[11] See Section A3 of the Web Appendix for a detailed discussion and analysis of the delay between a consumer buying and their review being displayed on the page and the number of reviews added within that duration.
[12] We also check quadratic and log specification of star ratings, the effect of the count of reviews of different star-ratings and star-rating of reviews separate by position. Results are available from authors upon request.

differences in the distribution of the star-rating of reviews added and at the same time to a change in sales. However, such medium to long-term changes are less of a concern since we focus on a relatively short two-months' time-window that makes it unlikely a significant number of product changes would occur during our data period. This is especially the case in our RD design where we focus on a 28-day time period.

2.4 Model-free evidence

To demonstrate patterns in the raw data before getting to an estimation, we report model-free evidence. First, we check whether the purchase probability varies with the star rating of the review displayed in position one as this review is likely the most salient one. Table 4 demonstrates that product-sessions that display a 5-star review in position one are associated with subsequent purchases of that product in 3.77% of cases whereas product-sessions that have a 4-star review in position one are followed by purchases in only 3.14% of cases. In instances where a 1-star review is in position one, consumers purchase in only 1.70% of cases. This provides some indication that the star rating of a review in position one affects a consumer's purchase decision. Note that we do not control for a product's average rating, the number of reviews available for a product or product-specific fixed effects which motivates our analysis in the next section.

Second, we focus on instances where a 1-, 2-, 3-, or 4-star review in the first position is replaced by a 5-star review. We limit our time window to two weeks before and after the addition of a new review, based on precise 24-hour windows from the timestamp of the new review. We divide the windows further into 7-day periods. The first two periods cover sessions that occurred 14-8 and 7-0 days before the discontinuity, the third and fourth period cover sessions in days 0 -7 and 8-14 following the discontinuity. Each point in Figure 3 shows the mean purchase rate among all sessions in the specific period, illustrating a discontinuity in purchase rates after the addition of a 5-star review. Figure 4 computes the mean purchase rate among all the sessions in the two periods before the discontinuity and similarly the mean purchase rate in the two periods after the discontinuity. The mean purchase rate in the sessions following the addition of the new 5-star review is significantly higher than in the sessions prior to the addition of that review ($p<0.05$).

3 Analysis and Results

3.1 Multiple Shifts in a Review's Position

3.1.1 Model

We initially utilitze the information arising from the multiple shifts reviews make across the product pages that occur as new reviews are added. This means that each product-session is captured in five observations corresponding to the five different reviews in five different positions on the product page of that session. We estimate the probability that a purchase occurs in a product-session as a function of the star rating and the position of each review displayed to the consumer:

$$Purchase_{rk} = \alpha + \sum_{s=1}^{5}\sum_{p=1}^{5}\beta_{sp}I_{spk} + \delta AvgRating_k + \theta AvgRat4Reviews_{rk} + \lambda NumberOfReviews_k + \rho Mobile_k + \tau Weekend_k + \sum_{p=1}^{5}\gamma_p ReviewWords_{pk} + \mu Product_k + \upsilon Week_k + \varepsilon_{rk} \qquad (1)$$

Here, index *r* represents a review, *k* represents the product-session, *s* represents the star rating of the review and *p* its position during session *k*. The variable $Purchase_{rk}$ represents the purchase decision in session *k* when review *r* was in position *p*. It is indexed by a review and not by a consumer because the data is at the level of a review impression rather than consumers. That is, across different sessions, the position of the same review shifts and the effect of its position on purchase in that session is modelled in (1).

I_{spk} is an indicator variable that has a value of 1 when review *r* has a star rating of *s* and position *p* during session *k*, and 0 otherwise. Thus, β_{sp} captures the effect of review *r* with star rating s in position p on the purchase likelihood of a consumer in session k. Additionally, we control for the average rating of the product displayed at the product page, $AvgRating_k$, the average rating of the other four reviews displayed on the page apart from review r, $AvgRat4Reviews_{rk}$, the number of reviews that are available for the product, $NumberOfReviews_k$, whether the website was accessed through a mobile device or a desktop, $Mobile_k$, and whether the session occurred on a weekday or a weekend, $Weekend_k$. Finally, $ReviewWords_{pk}$ controls for the number of words in a review in a given position to account for the possible effect of the length of a review[13] as well as fixed effects for the week and the product. Product fixed effects account for the fact that some products in the data could have better quality than others and so they also receive better reviews and obtain larger sales. We cluster standard errors at the product level.

3.1.2 Results

We estimate a linear probability model with product fixed effects as in Equation (1). Recall also that our results are conditional on a consumer having viewed individual reviews as our data consists only of such sessions. Column (I) of Table 5 demonstrates that compared to a 1-star review in position one, a 5-star review in the same position increases purchase probability by 0.88 percentage points.[14] This compares to an average purchase probability in the sample used for this estimation of 3.25%. 4 and 5-star reviews in other positions tend to have similar though less pronounced effects.

We calculate whether the effect of a review varies by position and check if coefficients β_{sp} differ across the positions of a review. For example, Column (I) of Table 5 shows that while displaying a 5-star review in position one increases purchase probability by

[13] We checked if the sentimentality of the textual content of the review affects consumer purchase. We used NLTK tool (http://www.nltk.org/) and calculated for each review subjectivity and polarity. Our results did not indicate a significant effect of these measures on purchase.
[14] The rest of the star-rating position effects described are also relative to a 1-star review in position one, which is our holdout category, but we avoid repeating the comparison for each result for the sake of brevity.

0.88 percentage points, the same review in position two only increases purchase probability by 0.50 percentage points. We check if this difference of 0.38 percentage points is significantly different from zero. Table 6 compares the effect of a review with a given star-rating in the first position to the effect of the same review in all other positions. It shows that the effect of a 5-star review is significantly more pronounced in the first position than in other positions. Reviews tend to be least effective when in positions three or four.

Note further that neither a product's average rating nor the number of reviews affect purchase probability (Column (I) of Table 5). This may be the case because of little variation in the two variables within the two-months time window and any variation across products would be captured by product fixed effects. It may also be a result of collinearity with the average rating of the other four reviews on the product page (r^2=0.6558, p<0.0001). To check for this possibility, we first drop $AvgRat4Reviews_{rk}$ from (1), as shown in Column (II) of Table 5, and then $AvgRating_k$ from (1), as shown in Column (III) of Table 5. In Column (II), average rating has a significant and positive effect. However, omitting this variable likely biases the results of the effects of star-ratings and position of individual reviews with big shifts in their magnitudes and statistical significances. This is also indicated in Column (III), which closely mirror the results of Column (I). Consequently, we use the model specified in Equation (1) and estimated in Column (I) as the main specification for all further analyses.[15]

[15] Note that due to multicollinearity, the average rating coefficient is negative and significant in some estimations (e.g. Column (III) and Column (IV) of Table 9). In all such cases, the coefficient on average rating is not significantly different from zero once the average rating of the other four reviews on the product page is dropped.

3.2 Regression discontinuity

3.2.1 Model

In the RD approach, each observation corresponds to a single product-session. This is because we are interested in comparing the difference between purchase rates before and after the addition of a review. We estimate the following model:

$$Purchase_k = \alpha + \beta After_k + \delta AvgRating_k + \theta AvgRat4Reviews_k + \lambda NumberOfReviews_k + \rho Mobile_k + \tau Weekend_k + \sum_{p=1}^{5} \gamma_p ReviewWords_{pk} + \mu Product_k + \varepsilon_k \qquad (2)$$

Here, index *k* represents the product-session. The variable $Purchase_k$ represents the purchase decision in session *k*. The main independent covariate of interest is $After_k$ which is a dummy variable that takes the value of 1 when session k corresponds to a session after the addition of the 5-star review in position one. It is 0 if session *k* occurs before the addition of the new review. All other independent variables are defined the same way as in Equation (1). We cluster standard errors at the product level.

3.2.2 Results

The results in Column (I) of Table 7 indicate that, on average, replacing a 1- to 4-star review with a 5-star review in the first position increases purchase probability by 1.58 percentage points. We break this down by each star-rating in position one. Column (II) indicates that replacing a 1-star review with a 5-star review in position one leads to an increase in purchase probability of 2.75 percentage points. This compares to an average probability to purchase in this sample of data of 3.38%.

Recall that if two successive reviews are added within the 14-day window, we limit our observations to the shorter time-window. We check if our specification is robust to such shortening of windows (or "bandwidths").[16] This would be of particular concern if the

[16] The checks are explained in more detail in Section A2 of the Web Appendix.

purchase probability of a session varied by how far away the session was within the bandwidth from the time of addition of the 5-star review. To check this, we add dummy variables to denote different durations within a bandwidth. An additional concern could be that a bandwidth is more likely to be shortened on one side of discontinuity. We add dummy variables to denote which side of discontinuity gets shortened. Our results in Column (III) and Column (IV) reproduce the specification of Column (I) and Column (II) with the bandwidth controls. As can be seen, our results continue to hold.

Three reasons lie behind the different magnitudes in effect sizes across our RD and MS estimation approaches. First, they are based on different selections of the data as detailed in Section 2.2. Thus, the estimates compare to different mean purchase probabilities, 3.25% and 3.38%, respectively. Second, the RD estimation captures the effect of the change in reviews only in the 14-day window around the point of discontinuity while the MS method captures the effect for as long as a review stays in a position in the data duration. Third, the MS approach includes cases with any review of any star-rating replacing any other review whereas in the RD setup, we only consider 5-star reviews replacing 1, 2, 3, or 4-star reviews. Importantly, though, they both illustrate a significant effect on purchase probabilities.

To further check the validity of the two sets of results, we use the data used for the RD analysis and restructure it at the level of a review impression to mirror the MS setup. Table 8 demonstrates that replacing a 1-star review in position one by a 5-star review increases the purchase probability by 2.47 percentage points closely matching the RD estimate of 2.72 percentage points (Column (IV) of Table 7).

We next check whether the RD results are sensitive to the choice of the time window of 14 days before and after the addition of a new review by varying the time windows by +/- 3 days. The results for 11-day time windows are shown in Column (I) and Column (II) of Table 9, those for 17-day time windows are spercentage pointshown in Column (III) and

Column (IV). The magnitude of the effect of adding a new 5-star review is similar across all time windows.

In sum, our results from both the MS and RD approaches indicate a strong and consistent effect of individual reviews on purchase.

3.3 Robustness

We report robustness checks to ensure our results are not an artefact of how we cleaned the data as well as checks for our model specification. We focus here on the MS analysis and report similar results for the RD analysis in the Web Appendix, Section A4.

We initially report robustness checks related to the data selection. First, some consumers may view the part of the webpage containing reviews very briefly because they quickly examine star ratings and decide to move on, or purchase, or due to accidental clicks.[17] Column (I) of Table 10 excludes sessions where the review part of the website was displayed for less than five seconds. The results are consistent with the main MS estimation.

Second, our analysis focuses on impressions of the first page where reviews where displayed, dropping impressions of later pages. In Column (II) we limit the data to only those sessions where the consumers did not browse past the first page of reviews. While the pattern of results holds, the effect of individual reviews is generally stronger.

Third, recall that the MS analyses include reviews from 5,433 of the 7,275 products in the data which have at least one review added during the data period. Column (III) demonstrates that our results are robust to using data from all products in the data.

We turn to robustness checks regarding our specification. First, Column (IV) demonstrates that the results hold with a quadratic specification for three control variables – average rating, average rating of the other four reviews on the page and number of reviews.

[17] Accidental clicks are common, see for example http://www.adweek.com/digital/do-you-accidentally-thumb-over-ads-all-time-googles-working-171314/.

Second, our data includes a count of the number of “thumbs up” and “thumbs down” votes given to each review by other consumers. Our main analysis does not include these variables as we do not observe them as time-varying but only at the end of the data period. In Column (V) control for the two variables. The results continue to hold.

Third, recall that our estimation is based on a panel of products. The concern could be that certain consumers view specific products and this type of consumer displays a specific type of behaviour. We re-arrange the data to estimate the regression for consumer sessions. That is, each observation of the data now represents a consumer’s product-session. In each product-session, this consumer views five individual reviews. Table 11 demonstrates that our findings are robust to this alternative specification. Note that this specification accounts for all five reviews that a consumer is exposed to and therefore captures the full set of effects rather than isolating the effect of individual reviews as the earlier analyses do.

3.4 When Do Individual Reviews Matter?

It is surprising that given the average rating of a product and the number of ratings it has obtained, individual reviews still affect consumers’ purchases. We ask when individual reviews are most likely to impact purchase decisions. We suggest two types of settings when this would be the case. First, individual reviews are likely to provide valuable information to consumers, and impact their purchase decision when these reviews resolve uncertainty around product quality that may affect consumers’ purchase probability (Erdem Keane and Sun 2008). As uncertainty resolves, given a mean expectation about product quality, purchase likelihood increases. In our setting, uncertainty may arise from limited product information being available or it may be specific to the type of product a consumer considers purchasing.

We suggest that consumer uncertainty is higher when the description on the website provides little product information. In such settings, reviews could resolve uncertainty and increase purchase probability. When product descriptions are exhaustive, we expect reviews to provide little incremental information and have little effect on purchase decisions. We focus on the MS data setup. Results in Section A5 of the Web Appendix demonstrate that the findings broadly hold in a RD design though the smaller sample size reduces significance.

We stratify our estimation by how extensive the product description is. Table 12, Column (I) displays the results for products where the description has a below-median number of words. Column (II) displays results for instances where the product description has an above-median number of words. While we find a strong effect of 4- and 5-star reviews when products have only a short description, individual reviews appear to have little effect when product descriptions are exhaustive. Likewise, the average rating of the four other reviews on a page has a stronger effect when product reviews are short.

We acknowledge that this analysis is not based on random assignment of products to product description. It is possible that unobserved product characteristics are correlated with the length of the description: descriptions might be shorter when purchase decisions are particularly straight-forward. However, if that were the case we would expect a lower effect of reviews on purchases when descriptions are short which is not what we find.

We turn to how uncertainty about product quality varies across products. There are many reasons why consumers could be uncertain about product quality. We focus on uncertainty related to horizontal differentiation such as if a particular type or color of bedsheet may be a good fit in one climate or for one style but not another. These dimensions are difficult to capture in our data directly, especially across a large number of heterogeneous products. Instead, we use the variance of the ratings a product has previously received as an indicator for whether a product is likely to be a good fit for some consumers but not for others. A high variance in prior ratings indicates a product which may appeal to a subset of consumers but not to everyone and for which consequently consumers are likely to have higher uncertainty. A low variance indicates a product which similarly appeals to consumers.

Table 12 presents the results when we stratify products based on the variance of past ratings. Columns (III) and (IV) clearly indicate that relative to a 1-star rating in position one, 4- and 5-star ratings increase a consumer's likelihood to purchase a product when there is a large variance in reviews. For products with a low variance in past reviews, individual reviews rarely contribute to consumer decisions, suggesting again that individual reviews mostly matter when they resolve consumer uncertainty about product quality.

The second setting where individual reviews are likely to influence purchase decisions is when they contrast with inferences consumers make about products based on information otherwise available. This would be the case when the aggregate rating of a product, which is typically used as a signal for product quality (e.g. Zhao et al. 2013; Wu et al. 2015) is low but an individual review instead displays a high star-rating. Similar, in instances of high aggregate ratings, the low star-rating of an individual review could provide additional information. Column (I) of Table 13 indicates that reviews with a high star rating are particularly effective in increasing purchase likelihood for products with a low average rating. By contrast, according to Column (II) only 1- and 2-star ratings affect purchase probability for products with high average ratings.

Additionally, consumers often rely on price as an indicator for product quality (Erdem, Keane and Sun 2008; Guo and Jiang 2016; Wang and Van der Lans 2018); a higher price is seen as a signal of higher product quality. Thus, for products priced low relative to others in the category, reviews with 4- or 5- star ratings should impact purchases. Column (III) confirms this. By contrast, Column (IV) suggests that for highly priced products, only 1-star ratings influence purchases though in general there is little effect of individual reviews.

In sum, our results suggest that the effect of star ratings of individual reviews on purchase is most pronounced when individual reviews either help resolve uncertainty about product quality or when they contrast with information displayed on the product page.

3.5 Managerial Implications

We discuss four managerial implications. First, our results show that individual reviews displayed on the product page have a strong effect on purchase. We compare them to the effects found for other forms of online marketing. Table 14 shows that research on banner ads finds purchase probability increases between 0.33 to 0.74 percentage points (Goldfarb and Tucker 2011a and 2011b; Urban et al 2014; Goldfarb and Tucker 2014). Research on email coupons has noted purchase incidence increases by 0.2 to 1.54 percentage points (Lewis 2014; Fong 2016) and a mobile coupon increased purchase rate by 0.22 percentage points (Danaher et al. 2015). Thus, our result that compared to a 1-star review in position one, a 5-star review in the same position increases purchase probability by 0.88 to 2.72 percentage points is close to such estimates. This suggests that our insights about the effect sizes of individual reviews are not negligible relative to other marketing actions that a firm can take, especially since the order of reviews entails little variable costs..

Second, the discussion of the impact of online reviews has largely focused on aggregate measures such as the average valence and volume. However, these studies demonstrate varying effect sizes, including positive and null effects, possibly because of the large heterogeneity of reviews aggregated in average ratings or because individual reviews were not controlled for. Our results point to the importance of individual reviews. They suggest firms should not be complacent if products have high average ratings or a large number of reviews as individual reviews on the first page can still impact sales.

Third, firms increasingly use algorithms to determine the order in which reviews are displayed. Our results suggest that firms may benefit from being aware of how individual reviews impact purchase when designing such algorithms. Recall, however, that our results are based on a context where reviews are displayed by recency. If a firm were to strategically sort the order of reviews, consumers might learn and adjust their behavior.

Fourth, though our results on a mechanism are not based on random assignment of products to descriptions, they suggest that the importance of individual reviews may be attenuated when websites present in-depth product information. In doing so, retailers may be able to help consumers avoid allocating perhaps disproportionate weight to individual reviews. At the same time, our findings provide evidence that consumers value reviews particularly when they contrast with other quality signals they may receive from the website.

4 Conclusion

As consumers increasingly shift spending online, online reviews have become a powerful source of information. Here, we investigate how individual reviews affect purchase decisions, a question that has as of yet received little attention. However, it matters since consumers typically focus on few individual reviews and retailers have latitude in deciding which reviews to display prominently, potentially shifting consumer purchase behaviour.

Our analysis is based on data on consumer browsing and purchase behavior of a multi-product online retailer. The key feature of our data is that the retailer displays reviews chronologically. This policy creates exogenous variation in the position in which any particular review is displayed as the addition of a new review pushes down the existing reviews in the list. This variation allows us to identify the effect of individual reviews and the position in which they are displayed on purchase probabilities.

We implement two empirical approaches – first, relying on the multiple shifts of reviews across positions. Second, a regression discountinuity approach that focuses on a limited time window around the addition of a new review. Throughout, we find that replacing a 1-star review in position one by a 5-star review increases purchase probability. We then demonstrate two types of settings when the effect of reviews on purchases is particularly pronounced. We find this to be the case, first, when individual reviews can resolve uncertainty about product quality and, second, when they provide information that contrasts with other information displayed on the page.

Our research contributes to a broader literature that studies the role of online reviews. In contrast to prior research, our findings emphasize the relative importance of individual reviews and suggest this is comparable to other online marketing activities that firms engage in. For firms, our results matter because they suggest that individual reviews can have a disproportionate – and potentially unintended – impact on consumer purchase decisions.

We further see our results as relevant in the context of the broader debate on how algorithms shape consumer decisions (Lambrecht and Tucker 2019). Our findings imply that even when a firm uses a simple algorithm that is blind to the favorability of a review and firms do not strategically manage the order of reviews, the information displayed might still lead consumers to make decisions not in their best interest: For example, if for a product rated negatively on average, positive reviews that are not representative of the overall quality are chronologically listed first, consumers may still buy. On the other hand, if their decision is swayed by an unfavorable first review, they may not purchase an objectively good product.

Our study has two limitations. First, while we document that the position of a review can influence consumers' purchase decisions, our results leave open the question which rule a retailer should apply when determining the order of reviews. The key challenge here is that consumers may infer the underlying decision rule and adjust their behavior. Even assuming that a retailer wishes to display the five most recent reviews on the first page, the question in which order the retailer should display those reviews remains to be answered. Second, while our analyses indicate that the effect of star ratings of individual reviews on purchase is most pronounced when reviews either resolve uncertainty about product quality or contrast with information displayed on the product page, this analysis is not based on random assignment of reviews to products and so not fully conclusive.

References

Brightlocal (2018), Local Consumer Review Survey. Retrieved July 23 2019 at https://www.brightlocal.com/learn/local-consumer-review-survey/.

Chevalier, J. A., & Mayzlin., D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research.* 43(3): 345-354.

Chintagunta, P. K., Gopinath, S. & Venkataraman, S. (2010). The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation across Local Markets. *Marketing Science.* 29(5):944-957.

Danaher, P. J., Smith, M. S., Ranasinghe, K., and Danaher, T. S. (2015). Where, when, and how long: Factors that influence the redemption of mobile phone coupons. *Journal of Marketing Research*, 52(5): 710-725.

De los Santos, B. and Koulayev, S. (2017). Optimizing Click-Through in Online Rankings with Endogenous Search Refinement. *Marketing Science.* 36(4):542-564.

Duan, W., Gu, B., and Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. *Journal of Retailing*. 84(2): 233–242.

Erdem, T., Keane, M. P., and Sun, B. (2008). A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*. 27(6): 1111-1125.

Fong, N. M. (2016). How targeting affects customer search: A field experiment. *Management Science.* 63(7): 2353-2364.

Ghose, A., Goldfarb, A., and Han, S. P. (2012). How is the mobile Internet different? Search costs and local activities. *Information Systems Research.* 24(3): 613-631.

Godes, D., and Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science.* 23(4): 545–560.

Goldfarb, A., & Tucker, C. E. (2014). Standardization and the effectiveness of online advertising. *Management Science.* 61(11): 2707-2719.

Goldfarb, A., and Tucker, C. (2011a). Privacy regulation and online advertising. *Management Science.* 57(1): 57-71.

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

Guo, X., and Jiang, B. (2016). Signaling through price and quality to consumers with fairness concerns. *Journal of Marketing Research.* 53(6): 988-1000.

Ho-Dac, N. N., Carson, S. J., and Moore, W. L. (2013). The effects of positive and negative online customer reviews: do brand strength and category maturity matter? *Journal of*

Marketing. 77(6): 37-53.

Jeziorski, P., and Moorthy, S. (2017). Advertiser prominence effects in search advertising. *Management Science*. 64(3): 1365-1383.

Lambrecht, A. & Tucker, C. (2019). Algorithmic bias? An empirical study into apparent gender-based discrimination in the display of STEM career ads. *Management Science.* 65(7): 2966-2981.

Lewis, M. (2004). The influence of loyalty programs and short-term promotions on customer retention. *Journal of Marketing Research.* 41(3): 281-292.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing.* 70(3): 74–89.

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., and Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing.* 77(1): 87-103.

Narayanan, S., & Kalyanam, K. (2015). Position Effects in Search Advertising and their Moderators: A Regression Discontinuity Approach. *Marketing Science.* 34(3):388-407.

Prentice, R. L., and Gloeckler, L. A. (1978). Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data. *Biometrics.* 34(1): 57-67.

Urban, G.L., Liberali, G., MacDonald, E., Bordley, R. and Hauser, J.R. (2013). Morphing banner advertising. *Marketing Science*. 33(1): 27-46.

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

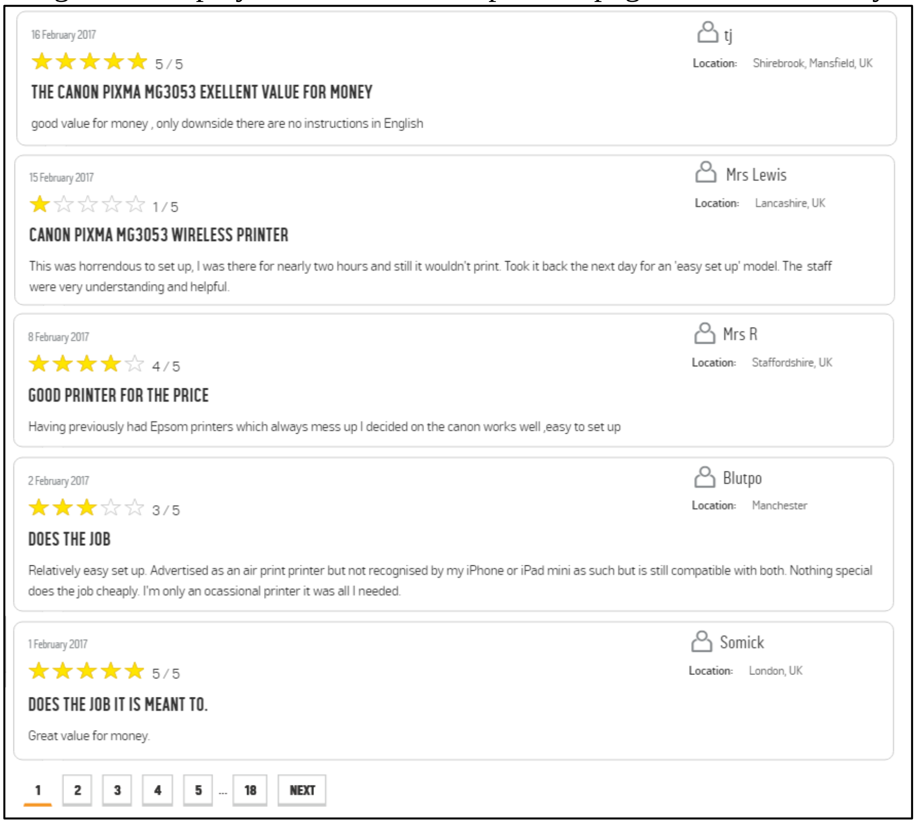
Wang, S. S., and Van Der Lans, R. (2018). Modeling gift choice: The effect of uncertainty on price sensitivity. *Journal of Marketing Research*, 55(4): 524-540.

Wu, C., Che, H., Chan, T.Y., and Lu, X. (2015). The Economic Value of Online Reviews. *Marketing Science.* 34(5):739-754.

Zhao, Y., Yang, S., Narayan, V., and Zhao, Y. (2013). Modeling consumer learning from online product reviews. *Marketing Science.* 32(1): 153-169.

Zhu, F., and Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing.* 74(2): 133-148.

Figure 1: Display of reviews on the product page as of 16 February


16 February 2017
5 / 5
tj
Location: Shirebrook, Mansfield, UK
THE CANON PIXMA MG3053 EXELLENT VALUE FOR MONEY
good value for money , only downside there are no instructions in English

15 February 2017
1 / 5
Mrs Lewis
Location: Lancashire, UK
CANON PIXMA MG3053 WIRELESS PRINTER
This was horrendous to set up, I was there for nearly two hours and still it wouldn't print. Took it back the next day for an 'easy set up' model. The staff were very understanding and helpful.

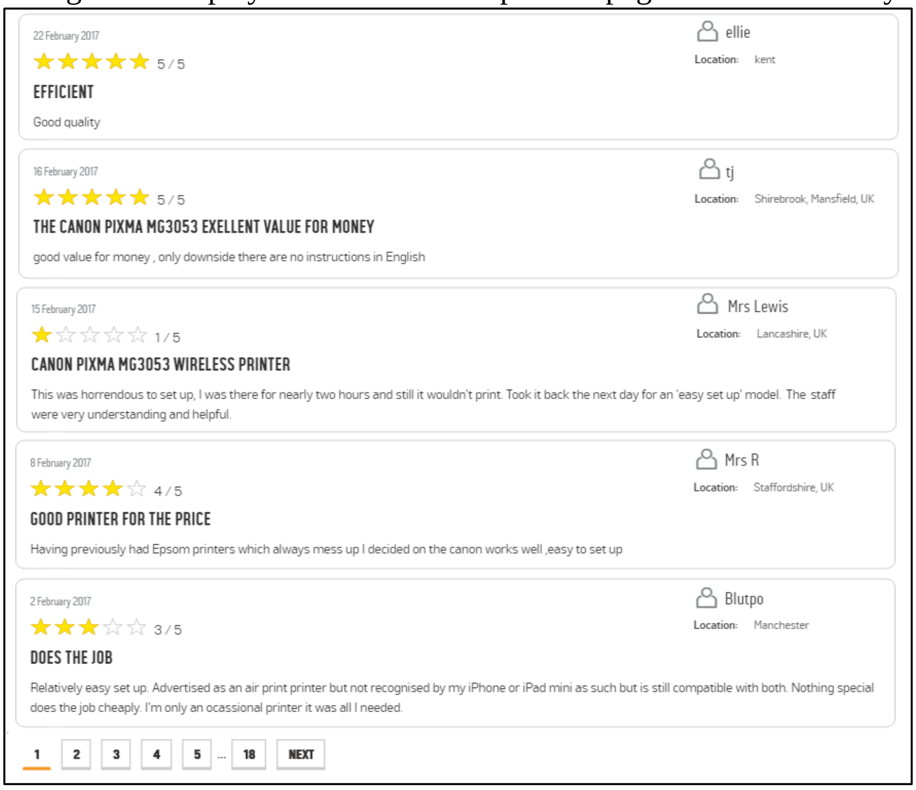
8 February 2017
4 / 5
Mrs R
Location: Staffordshire, UK
GOOD PRINTER FOR THE PRICE
Having previously had Epsom printers which always mess up I decided on the canon works well ,easy to set up

2 February 2017
3 / 5
Blutpo
Location: Manchester
DOES THE JOB
Relatively easy set up. Advertised as an air print printer but not recognised by my iPhone or iPad mini as such but is still compatible with both. Nothing special does the job cheaply. I'm only an ocassional printer it was all I needed.

1 February 2017
5 / 5
Somick
Location: London, UK
DOES THE JOB IT IS MEANT TO.
Great value for money.

1 2 3 4 5 ... 18 NEXT

Figure 2: Display of reviews on the product page as of 22 February


22 February 2017
5 / 5
ellie
Location: kent
EFFICIENT
Good quality

16 February 2017
5 / 5
tj
Location: Shirebrook, Mansfield, UK
THE CANON PIXMA MG3053 EXELLENT VALUE FOR MONEY
good value for money , only downside there are no instructions in English

15 February 2017
1 / 5
Mrs Lewis
Location: Lancashire, UK
CANON PIXMA MG3053 WIRELESS PRINTER
This was horrendous to set up, I was there for nearly two hours and still it wouldn't print. Took it back the next day for an 'easy set up' model. The staff were very understanding and helpful.

8 February 2017
4 / 5
Mrs R
Location: Staffordshire, UK
GOOD PRINTER FOR THE PRICE
Having previously had Epsom printers which always mess up I decided on the canon works well ,easy to set up

2 February 2017
3 / 5
Blutpo
Location: Manchester
DOES THE JOB
Relatively easy set up. Advertised as an air print printer but not recognised by my iPhone or iPad mini as such but is still compatible with both. Nothing special does the job cheaply. I'm only an ocassional printer it was all I needed.

1 2 3 4 5 ... 18 NEXT

Figure 3: Model-free evidence for regression discontinuity design: Two Time Windows on Either Side of Discontinuity

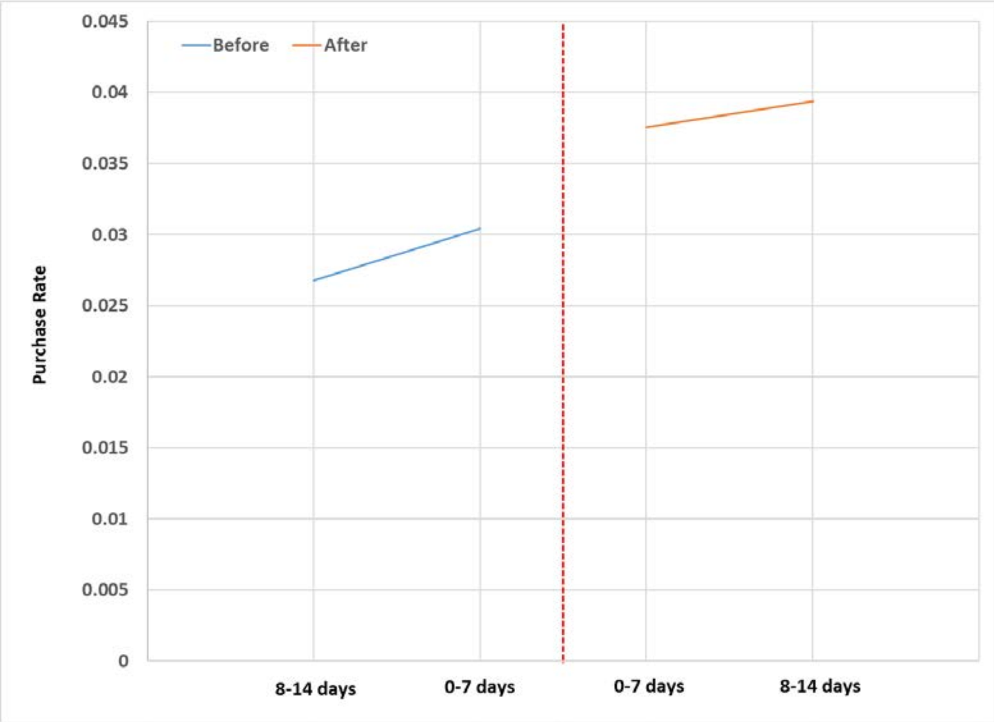


Figure 4: Model-free evidence for regression discontinuity design: One Time Window on Either Side of Discontinuity

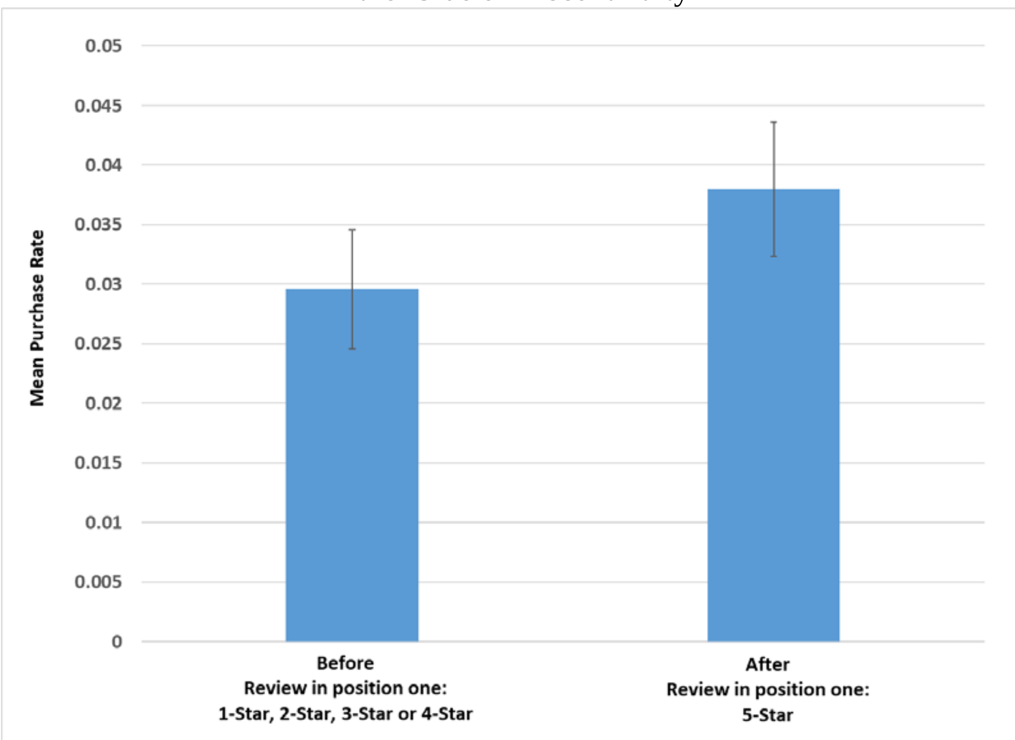


Table 1: Distribution of Old and New Reviews in the First Position (As Used in the RD Analysis)

		Old review in the first position					
		1-Star	2-Star	3-Star	4-Star	5-Star	Total
New review in the first position	1-Star	131	40	61	192	270	694
	2-Star	31	10	15	56	106	218
	3-Star	51	26	39	110	207	433
	4-Star	166	58	132	557	1,016	1,929
	5-Star	239	117	200	1,052	2,302	3,910
	Total	618	251	447	1,967	3,901	7,184

Table 2: Summary of data

Variable	Sessions	Mean	Median	Standard deviation	Minimum	Maximum
Product						
Average rating of a product	76,090	4.24	4.33	0.46	1.15	5
Number of reviews of a product	76,090	159.8	68	273.94	5	4,761
Mean price of a product (£)	39,349	73.12	39.99	87.64	0	1,049
Review						
Rating of a review	43,388	4.18	5	1.2	1	5
Number of words in a review	43,388	30.89	23	24.27	1	218
Other						
Platform used (1=mobile, 0=desktop)	76,090	0.35	0	0.48	0	1
Weekend (1= Weekend, 0 = Weekday)	76,090	0.32	0	0.47	0	1

Table 3: Checks for Exogeneity of Addition of New Reviews

	(I)		(II)			(III)	
	OLS Specification		Cox proportional hazard model			OLS pecification	
Variables	Estimate	SE	Estimate	SE		Estimate	SE
Dependent variable	Average review rating of newly submitted reviews		Time till the addition of a new review (sec)			Time till the addition of a new review (sec)	
Independent variables							
Review previously in position…							
One			-0.0224	0.0187		19,119.83	441,600.80
Two			-0.0271	0.0198		9,169.12	437,873.30
Three			-0.0225	0.0189		642,088.90	497,686.20
Four			0.0049	0.0197		484,839.80	456,554.40
Five			-0.0022	0.0191		146,925.40	454,914.00
Average rating of the product			0.0108	0.1188		-790,621.10	2,903,384.00
Number of reviews of the product			0.0005	0.0001	***	5,304.24	8,353.39
Number of customers browsing the product page	-.00184	.00382					
Product fixed-effects	Yes		Stratified baseline hazard			Yes	
Day fixed-effects	Yes						
N	11,996		6,302			6,302	
R^2	0.40					0.53	
Log likelihood			-4371.06				

**=p<0.1, **=p<0.05, ***=p<0.001*

Table 4: Star rating of the review in position one and percent buying

Star rating of the review in position one	Review impressions with		Total	Percent buy
	Buy	No buy		
1-Star	177	7,165	7342	2.41%
2-Star	70	2,400	2,470	2.83%
3-Star	143	4,747	4,890	2.92%
4-Star	741	19,914	20,665	3.59%
5-Star	1,702	39,031	40,733	4.18%

Table 5: Results based on Multiple Shifts (MS) of Review Positions

		(I)			(II)			(II)		
		LPM: Average rating and average of other four reviews on the page			LPM: Average rating only			LPM: Average rating of other four reviews on the page only		
		Estimate	SE		Estimate	SE		Estimate	SE	
Position of the review	*Star rating*									
One	1-Star									
	2-Star	0.0007	0.0037		-0.0005	0.0037		0.0007	0.0037	
	3-Star	0.0059	0.0032	*	0.0032	0.0031		0.0059	0.0032	*
	4-Star	0.0058	0.0024	**	0.0022	0.0022		0.0057	0.0023	**
	5-Star	0.0088	0.0023	***	0.0041	0.0019	**	0.0087	0.0022	***
Two	1-Star	-0.0041	0.0029		-0.0038	0.0029		-0.0041	0.0029	
	2-Star	-0.0029	0.0039		-0.0038	0.0039		-0.0029	0.0039	
	3-Star	-0.0001	0.0036		-0.0025	0.0036		-0.0001	0.0036	
	4-Star	0.0048	0.0026	*	0.0014	0.0025		0.0048	0.0026	*
	5-Star	0.0050	0.0026	*	0.0003	0.0022		0.0049	0.0025	**
Three	1-Star	-0.0061	0.0031	**	-0.0060	0.0031	*	-0.0061	0.0031	**
	2-Star	0.0027	0.0042		0.0020	0.0042		0.0027	0.0043	
	3-Star	0.0027	0.0032		-0.0001	0.0031		0.0026	0.0032	
	4-Star	0.0024	0.0025		-0.0012	0.0024		0.0024	0.0025	
	5-Star	0.0065	0.0026	**	0.0019	0.0022		0.0064	0.0025	**
Four	1-Star	-0.0060	0.0031	*	-0.0059	0.0031	*	-0.0060	0.0031	*
	2-Star	-0.0095	0.0038	**	-0.0110	0.0038	***	-0.0096	0.0038	**
	3-Star	-0.0032	0.0037		-0.0058	0.0036		-0.0032	0.0037	
	4-Star	0.0037	0.0027		0.0000	0.0025		0.0036	0.0026	
	5-Star	0.0035	0.0026		-0.0012	0.0023		0.0034	0.0025	
Five	1-Star	-0.0025	0.0031		-0.0025	0.0031		-0.0025	0.0031	
	2-Star	-0.0022	0.0038		-0.0037	0.0038		-0.0022	0.0038	
	3-Star	0.0031	0.0030		0.0005	0.0030		0.0030	0.0030	
	4-Star	0.0064	0.0027	**	0.0027	0.0025		0.0064	0.0026	**
	5-Star	0.0057	0.0026	**	0.0011	0.0022		0.0056	0.0026	**
Average rating of the product		-0.0004	0.0057		0.0136	0.0052				
Average rating of other four reviews on the page		0.0097	0.0017	***				0.0096	0.0016	***
Number of reviews of the product		0.0000	0.0000		0.0000	0.0000		0.0000	0.0000	
Mobile platform fixed effect		0.0346	0.0020	***	0.0346	0.0020		0.0346	0.0020	***
Weekend fixed effect		0.0006	0.0018		0.0005	0.0018		0.0006	0.0018	
Number of words in the five positions		Yes			Yes			Yes		
Product fixed effect		Yes			Yes			Yes		
Week specific fixed effect		Yes			Yes			Yes		
Products		3,396			3,396			3,396		
N		380,450			380,450			380,450		
R^2		0.08			0.08			0.08		

**=p<0.1, **=p<0.05, ***=p<0.001*

Standard errors clustered at the product level for all estimations.

Table 6: Difference in the Coefficients of Positions

		Star-rating of review in position one				
		1-star	2-star	3-star	4-star	5-star
Difference in the effect of a review on purchase probability when the review is in position one instead of in position…	Two	-0.0041	0.0036	0.006	0.001	0.0038***
	Three	-0.0061**	-0.002	0.0032	0.0034*	0.0023*
	Four	-0.006*	0.0103**	0.0091**	0.0021	0.0053***
	Five	-0.0025	0.0029	0.0028	-0.0006	0.0031**

**=p<0.1, **=p<0.05, ***=p<0.001*

Table 7: Results Based on Regression Discontinuity (RD)

	(I)			(II)			(III)			(IV)		
	Without bandwidth controls						With bandwidth controls					
	Average effect of all star-ratings of before reviews			Breakdown of effect by star-ratings of before reviews			Average effect of all star-ratings of before reviews			Breakdown of effect by star-ratings of before reviews		
	Estimate	SE		Estimate	SE		Estimate	SE		Estimate	SE	
After	0.0158	0.0066	**				0.0155	0.0067	**			
Before star-rating												
1-star				0.0275	0.0143	*				0.0272	0.0143	*
2-star				0.0375	0.0154	**				0.0380	0.0153	**
3-star				0.0151	0.0170					0.0148	0.0169	
4-star				0.0141	0.0068	**				0.0138	0.0069	**
Bandwidth controls…												
Session within 5-8 days of review addition							0.0024	0.0059		0.0021	0.0059	
Session within 9-12 days of review addition							-0.0029	0.0079		-0.0033	0.0079	
Session within 13-16 days of review addition							0.0116	0.0136		0.0112	0.0135	
Bandwidth censored before the addition of review							0.0024	0.0162		0.0013	0.0162	
Bandwidth censored after the addition of review							0.0072	0.0304		0.0119	0.0293	
Average rating of the product	-0.0941	0.0677		-0.1178	0.0749		-0.0939	0.0686		-0.1183	0.0757	
Average rating of other four reviews on the page	0.0173	0.0099	*	0.0241	0.0111	**	0.0174	0.0101	*	0.0245	0.0112	**
Number of reviews of the product	0.0000	0.0006		0.0000	0.0006		0.0000	0.0006		0.0000	0.0006	
Mobile platform fixed effect	0.0357	0.0058	***	0.0357	0.0058	***	0.0356	0.0058	***	0.0355	0.0058	***
Weekend fixed effect	0.0003	0.0054		0.0002	0.0054		0.0011	0.0053		0.0011	0.0053	
Number of words in each of the five reviews	Yes			Yes			Yes			Yes		
Product fixed effect	Yes			Yes			Yes			Yes		
Products	1,182			1,182			1,182			1,182		
N	8,820			8,820			8,820			8,820		
R^2	0.1816			0.1819			0.1818			0.1821		

**=p<0.1, **=p<0.05, ***=p<0.001*
Standard errors clustered at the product level in all estimations.

Table 8: Multiple Shifts (MS) Analysis Using the Same Sessions as in the Regression Discontinuity (RD) Estimation

		(I)		
		Linear Probability Model		
		Estimate	SE	
Position of the review	*Star rating*			
One	1-Star	-		
	2-Star	-0.0045	0.0089	
	3-Star	0.0142	0.0092	
	4-Star	0.0164	0.0089	*
	5-Star	0.0247	0.0117	**
Two	1-Star	0.0023	0.0094	
	2-Star	-0.0007	0.0089	
	3-Star	0.0111	0.0098	
	4-Star	0.0201	0.0101	**
	5-Star	0.0230	0.0118	*
Three	1-Star	-0.0047	0.0072	
	2-Star	0.0073	0.0102	
	3-Star	0.0011	0.0104	
	4-Star	0.0171	0.0100	*
	5-Star	0.0212	0.0122	*
Four	1-Star	-0.0044	0.0071	
	2-Star	-0.0169	0.0091	*
	3-Star	0.0154	0.0098	
	4-Star	0.0137	0.0102	
	5-Star	0.0205	0.0112	*
Five	1-Star	-0.0074	0.0079	
	2-Star	0.0180	0.0105	*
	3-Star	0.0019	0.0094	
	4-Star	0.0177	0.0098	*
	5-Star	0.0195	0.0115	*
Average rating of the product		-0.1022	0.0676	
Average rating of other four reviews on the page		0.0252	0.0104	**
Number of reviews of the product		-0.0004	0.0007	
Mobile platform fixed effect		0.0347	0.0054	***
Weekend fixed effect		-0.0036	0.0052	
Number of words in the five positions		Yes		
Product fixed effects		Yes		
Week specific fixed effect		Yes		
Products		998		
N		43,215		
R^2		0.18		

**=p<0.1, **=p<0.05, ***=p<0.001*

Standard errors clustered at the product level in all estimations.

Table 9: Regression Discontinuity (RD): Robustness to Different Bandwidths

	(I)			(II)			(III)			(IV)		
	11 day before-after bandwidths						17 day before-after bandwidths					
	Average effect of all star-ratings of before reviews			Breakdown of effect by star-ratings of before reviews			Average effect of all star-ratings of before reviews			Breakdown of effect by star-ratings of before reviews		
	Estimate	SE		Estimate	SE		Estimate	SE		Estimate	SE	
After	0.0128	0.0069	**				0.0157	0.0065	**			
Before star-rating												
1-star				0.0260	0.0149	*				0.0265	0.0138	*
2-star				0.0357	0.0160	**				0.0310	0.0155	**
3-star				0.0076	0.0169					0.0160	0.0162	
4-star				0.0117	0.0072					0.0142	0.0067	**
Bandwidth controls…												
Session within 5-8 days of review addition	0.0018	0.0059		0.0015	0.0059		0.0029	0.0058		0.0027	0.0058	
Session within 9-12 days of review addition	0.0019	0.0092		0.0014	0.0092		-0.0031	0.0077		-0.0034	0.0077	
Session within 13-16 days of review addition							0.0123	0.0111		0.0120	0.0111	
Bandwidth censored before the addition of review	0.0038	0.0135		0.0026	0.0136		0.0015	0.0157		0.0007	0.0157	
Bandwidth censored after the addition of review	0.0125	0.0337		0.0173	0.0326		-0.0012	0.0271		0.0016	0.0265	
Average rating of the product	-0.0820	0.0708		-0.1043	0.0777		-0.1229	0.0700	*	-0.1444	0.0771	*
Average rating of other four reviews on the page	0.0135	0.0103	*	0.0208	0.0115	*	0.0181	0.0097	*	0.0239	0.0109	**
Number of reviews of the product	0.0000	0.0006		0.0000	0.0006		-0.0001	0.0006		-0.0001	0.0006	
Mobile platform fixed effect	0.0356	0.0062	***	0.0356	0.0062	***	0.0338	0.0057	***	0.0338	0.0057	***
Weekend fixed effect	0.0030	0.0058		0.0029	0.0058		0.0030	0.0052		0.0029	0.0052	
Number of words in all five positions	Yes			Yes			Yes			Yes		
Product fixed effect	Yes			Yes			Yes			Yes		
Products	1,182			1,182			1,213			1,131		
N	8,010			8,010			9,336			8,010		
R^2	0.1866			0.1870			0.1800			0.1870		

**=p<0.1, **=p<0.05, ***=p<0.001*
Standard errors clustered at the product level in all estimations.

Table 10: Robustness checks: Robustness to Data Cleaning

		(I) Excluding sessions where reviews were displayed for less than 5 seconds in a session		(II) Excluding product-sessions where consumer visited second page of reviews		(III) Including products where there was no new review added during data collection period		(IV) Quadratic effect of average rating and number of reviews		(V) Including controls for "thumbs up" and "thumbs down" votes	
		Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Position of the review	*Star rating*										
	1-Star										
	2-Star	0.0006	0.0035	0.0040	0.0041	-0.0004	0.0033	0.0010	0.0037	0.0012	0.0038
One	3-Star	0.0065	0.0034 *	0.0074	0.0034 **	0.0050	0.0029 *	0.0062	0.0032 *	0.0062	0.0032 *
	4-Star	0.0052	0.0025 **	0.0099	0.0026 ***	0.0056	0.0023 **	0.0061	0.0024 **	0.0063	0.0024 **
	5-Star	0.0094	0.0025 ***	0.0138	0.0026 ***	0.0084	0.0022 ***	0.0090	0.0023 ***	0.0086	0.0024 ***
	1-Star	-0.0052	0.0029 *	-0.0021	0.0031	-0.0042	0.0027	-0.0041	0.0029	-0.0039	0.0029
	2-Star	-0.0044	0.0039	-0.0005	0.0039	-0.0031	0.0036	-0.0028	0.0039	-0.0033	0.0039
Two	3-Star	0.0002	0.0036	0.0024	0.0039	0.0005	0.0034	0.0002	0.0037	-0.0008	0.0036
	4-Star	0.0055	0.0027 **	0.0081	0.0029 ***	0.0049	0.0025 **	0.0051	0.0026 *	0.0049	0.0026 *
	5-Star	0.0056	0.0027 **	0.0092	0.0028 ***	0.0050	0.0024 **	0.0051	0.0025 **	0.0052	0.0026 **
	1-Star	-0.0054	0.0033 *	-0.0045	0.0033	-0.0058	0.0029 **	-0.0061	0.0031 **	-0.0059	0.0031 *
	2-Star	0.0011	0.0043	0.0043	0.0045	0.0028	0.0039	0.0029	0.0043	0.0034	0.0042
Three	3-Star	0.0015	0.0034	0.0081	0.0034 **	0.0028	0.0030	0.0030	0.0032	0.0027	0.0032
	4-Star	0.0027	0.0027	0.0060	0.0028 **	0.0029	0.0024	0.0027	0.0025	0.0026	0.0026
	5-Star	0.0071	0.0027 ***	0.0123	0.0029 ***	0.0062	0.0024 **	0.0067	0.0026 ***	0.0066	0.0026 **
	1-Star	-0.0067	0.0032 **	-0.0025	0.0033	-0.0061	0.0029 **	-0.0061	0.0031 *	-0.0056	0.0031 *
	2-Star	-0.0092	0.0039 **	-0.0058	0.0040	-0.0110	0.0035 ***	-0.0092	0.0038 **	-0.0090	0.0039 **
Four	3-Star	-0.0026	0.0039	0.0007	0.0038	-0.0030	0.0034	-0.0029	0.0037	-0.0026	0.0037
	4-Star	0.0036	0.0029	0.0086	0.0029 ***	0.0041	0.0026	0.0039	0.0027	0.0039	0.0027
	5-Star	0.0039	0.0029	0.0082	0.0029 ***	0.0033	0.0025	0.0036	0.0026	0.0035	0.0027
	1-Star	-0.0043	0.0033	-0.0015	0.0032	-0.0021	0.0029	-0.0025	0.0031	-0.0022	0.0032
	2-Star	-0.0053	0.0040	0.0019	0.0041	-0.0017	0.0036	-0.0020	0.0039	-0.0025	0.0039
Five	3-Star	0.0036	0.0033	0.0044	0.0033	0.0028	0.0029	0.0034	0.0031	0.0028	0.0030
	4-Star	0.0061	0.0029 **	0.0108	0.0028 ***	0.0062	0.0025 **	0.0068	0.0027 **	0.0063	0.0027 **
	5-Star	0.0058	0.0028 **	0.0095	0.0029 ***	0.0056	0.0025 **	0.0058	0.0026 **	0.0056	0.0027 **
Average rating of the product		-0.0029	0.0060	-0.0072	0.0066	0.0016	0.0053	-0.0407	0.0323	0.0007	0.0058
Average rating of the product^2								0.0051	0.0044		
Avg. rating of four other reviews on the page		0.0108	0.0018 ***	0.0125	0.0020 ***	0.0095	0.0017 ***	0.0154	0.0069 **	0.0091	0.0018 ***
Avg. rating of four other reviews on the page^2								-0.0008	0.0010		
Number of reviews of the product		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Number of reviews of the product^2								0.0000	0.0000		
Mobile platform fixed effect		0.0326	0.0022 ***	0.0327	0.0022 ***	0.0324	0.0019 ***	0.0346	0.0020 ***	0.0346	0.0020 ***
Weekend fixed effect		0.0022	0.0019	0.0004	0.0019	0.0009	0.0017	0.0005	0.0018	0.0004	0.0018
Number of "thumbs up" votes										-0.0001	0.0002
Number of "thumbs down" votes										0.0005	0.0002 ***
Number of words in all five positions		Yes		Yes		Yes		Yes		Yes	
Product fixed effect		Yes		Yes		Yes		Yes		Yes	
Week specific fixed effect		Yes		Yes		Yes		Yes		Yes	
Products		3,395		3,395		4,107		3,396		3,383	
N		301,320		306,140		419,875		380,450		375,210	
R^2		0.08		0.09		0.08		0.08		0.08	

**=p<0.1, **=p<0.05, ***=p<0.001* , standard errors clustered at the product level in all estimations.

Table 11: Robustness Check: Panel of Consumers

	(I)		
	Panel of product-sessions		
	Estimate	SE	
Position of the review			
One	0.0024	0.0008	***
Two	0.0022	0.0008	***
Three	0.0027	0.0009	***
Four	0.0029	0.0009	***
Five	0.0017	0.0008	**
Average rating of the product	0.0005	0.0059	
Number of reviews of the product	0.0000	0.0000	
Mobile platform fixed effect	0.0350	0.0021	***
Weekend fixed effect	0.0008	0.0019	
Number of words in all five positions	Yes		
Product fixed effect	Yes		
Week specific fixed effect	Yes		
Products	3,279		
N	74,302		
R^2	0.08		

**=p<0.1, **=p<0.05, ***=p<0.001*

Standard errors clustered at the product level.

Table 12: Multiple Shifts (MS) Results by Product Description and Review Variance

		(I)			(II)			(III)			(IV)		
		Product Description <= 68 words			Product Description > 68 words			High review variance			Low review variance		
		Estimate	SE		Estimate	SE		Estimate	SE		Estimate	SE	
Position of the review	*Star rating*												
One	1-Star												
	2-Star	0.0014	0.0051		0.0011	0.0054		0.0000	0.0044		0.0045	0.0073	
	3-Star	0.0082	0.0048	*	0.0033	0.0041		0.0057	0.0037		0.0081	0.0064	
	4-Star	0.0090	0.0035	**	0.0033	0.0034		0.0069	0.0029	**	0.0059	0.0049	
	5-Star	0.0141	0.0035	***	0.0035	0.0032		0.0093	0.0028	***	0.0092	0.0049	*
Two	1-Star	-0.0047	0.0043		-0.0038	0.0038		-0.0033	0.0033		-0.0064	0.0074	
	2-Star	-0.0054	0.0059		-0.0008	0.0050		-0.0014	0.0044		-0.0061	0.0086	
	3-Star	0.0025	0.0049		-0.0024	0.0053		0.0008	0.0042		-0.0004	0.0071	
	4-Star	0.0084	0.0039	**	0.0013	0.0035		0.0054	0.0032	*	0.0054	0.0053	
	5-Star	0.0087	0.0039	**	0.0010	0.0035		0.0049	0.0030		0.0055	0.0053	
Three	1-Star	-0.0027	0.0047		-0.0098	0.0037	***	-0.0054	0.0034		-0.0076	0.0080	
	2-Star	0.0109	0.0063	*	-0.0061	0.0053		0.0039	0.0045		0.0006	0.0104	
	3-Star	0.0052	0.0048		-0.0013	0.0042		0.0027	0.0036		0.0042	0.0066	
	4-Star	0.0042	0.0038		0.0005	0.0034		0.0028	0.0029		0.0033	0.0052	
	5-Star	0.0113	0.0038	***	0.0018	0.0036		0.0073	0.0031	**	0.0067	0.0052	
Four	1-Star	-0.0050	0.0045		-0.0069	0.0043		-0.0055	0.0037		-0.0028	0.0077	
	2-Star	-0.0070	0.0058		-0.0106	0.0051	**	-0.0063	0.0043		-0.0142	0.0089	
	3-Star	-0.0035	0.0054		-0.0033	0.0051		-0.0043	0.0044		0.0006	0.0070	
	4-Star	0.0074	0.0040	*	0.0005	0.0038		0.0057	0.0034	*	0.0028	0.0052	
	5-Star	0.0071	0.0039	*	0.0000	0.0036		0.0055	0.0031	*	0.0024	0.0053	
Five	1-Star	0.0000	0.0045		-0.0053	0.0043		-0.0031	0.0035		0.0039	0.0075	
	2-Star	0.0041	0.0052		-0.0090	0.0057		0.0022	0.0041		-0.0113	0.0088	
	3-Star	0.0054	0.0044		0.0006	0.0042		0.0017	0.0034		0.0075	0.0064	
	4-Star	0.0108	0.0039	***	0.0020	0.0037		0.0061	0.0031	*	0.0078	0.0052	
	5-Star	0.0105	0.0039	***	0.0010	0.0036		0.0082	0.0030	***	0.0045	0.0053	
Average rating of the product		-0.0078	0.0092		0.0026	0.0087		-0.0076	0.0057		0.0142	0.0135	
Avg. rating of four other reviews on the page		0.0127	0.0026	***	0.0069	0.0025	***	0.0103	0.0020	***	0.0081	0.0033	**
Number of reviews of the product		0.0000	0.0000		0.0000	0.0000		0.0000	0.0000		0.0000	0.0000	
Mobile platform fixed effect		0.0375	0.0030	***	0.0315	0.0026	***	0.0278	0.0025	***	0.0412	0.0032	***
Weekend fixed effect		-0.0006	0.0026		0.0017	0.0025		-0.0025	0.0022		0.0033	0.0028	
Number of words in all five positions		Yes			Yes			Yes			Yes		
Product fixed effect		Yes			Yes			Yes			Yes		
Week specific fixed effect		Yes			Yes			Yes			Yes		
Products		2,096			1,441			1,699			1,697		
N		190,975			189,475			186,540			193,910		
R^2		0.09			0.07			0.08			0.08		

**=p<0.1, **=p<0.05, ***=p<0.001*
Standard errors clustered at the product level

Table 13: Multiple Shifts (MS) Results by Average Rating and Price Within Category

		(I)			(II)			(III)			(IV)		
		Average rating <= 4.3245			Average rating > 4.3245			Average price below median category price			Average price above median category price		
		Estimate	SE		Estimate	SE		Estimate	SE		Estimate	SE	
Position of the review	Star rating												
One	1-Star												
	2-Star	0.0010	0.0042		-0.0025	0.0077		-0.0058	0.0090		0.0014	0.0133	
	3-Star	0.0068	0.0034	**	0.0048	0.0071		0.0111	0.0087		0.0127	0.0128	
	4-Star	0.0079	0.0026	***	0.0002	0.0053		0.0131	0.0062	**	0.0054	0.0096	
	5-Star	0.0126	0.0028	***	0.0030	0.0052		0.0149	0.0058	***	0.0134	0.0089	
Two	1-Star	0.0001	0.0032		-0.0144	0.0071	**	-0.0016	0.0083		-0.0197	0.0110	*
	2-Star	0.0015	0.0041		-0.0077	0.0098		-0.0080	0.0100		-0.0109	0.0180	
	3-Star	0.0068	0.0039	*	-0.0094	0.0083		0.0014	0.0083		-0.0033	0.0154	
	4-Star	0.0102	0.0031	***	-0.0033	0.0055		0.0075	0.0064		0.0051	0.0096	
	5-Star	0.0108	0.0030	***	-0.0036	0.0055		0.0121	0.0062	*	-0.0032	0.0097	
Three	1-Star	-0.0004	0.0036		-0.0257	0.0066	***	-0.0123	0.0085		-0.0142	0.0123	
	2-Star	0.0045	0.0043		0.0027	0.0122		0.0014	0.0103		0.0143	0.0182	
	3-Star	0.0059	0.0037		-0.0049	0.0065		0.0030	0.0082		-0.0036	0.0131	
	4-Star	0.0055	0.0030	*	-0.0024	0.0053		0.0046	0.0064		-0.0027	0.0096	
	5-Star	0.0100	0.0031	***	0.0005	0.0054		0.0107	0.0062	*	0.0095	0.0097	
Four	1-Star	-0.0011	0.0036		-0.0185	0.0071	***	-0.0098	0.0081		-0.0234	0.0120	*
	2-Star	-0.0010	0.0041		-0.0281	0.0094	***	-0.0210	0.0089	**	-0.0159	0.0152	
	3-Star	-0.0014	0.0040		-0.0058	0.0087		-0.0048	0.0085		-0.0185	0.0158	
	4-Star	0.0092	0.0031	***	-0.0045	0.0056		0.0073	0.0065		-0.0014	0.0101	
	5-Star	0.0093	0.0030	***	-0.0048	0.0055		0.0067	0.0063		-0.0028	0.0095	
Five	1-Star	-0.0002	0.0034		-0.0030	0.0084		-0.0027	0.0083		-0.0160	0.0129	
	2-Star	0.0061	0.0042		-0.0286	0.0084	***	0.0016	0.0091		-0.0176	0.0176	
	3-Star	0.0069	0.0034	**	-0.0013	0.0065		0.0060	0.0076		-0.0068	0.0115	
	4-Star	0.0107	0.0031	***	-0.0003	0.0056		0.0143	0.0067	**	0.0016	0.0097	
	5-Star	0.0094	0.0030	***	-0.0018	0.0054		0.0115	0.0064	*	0.0047	0.0095	
Average rating of the product		-0.0141	0.0102		-0.0071	0.0185		-0.0203	0.0137		0.0120	0.0301	
Avg. rating of four other reviews on the page		0.0108	0.0020	***	0.0097	0.0035	***	0.0168	0.0032	***	0.0180	0.0056	***
Number of reviews of the product		0.0000	0.0000		0.0000	0.0000		0.0000	0.0000	**	0.0000	0.0000	
Mobile platform fixed effect		0.0263	0.0026	***	0.0427	0.0031	***	0.0619	0.0050	***	0.0673	0.0057	***
Weekend fixed effect		-0.0018	0.0022		0.0024	0.0028		0.0016	0.0042		-0.0001	0.0059	
Number of words in all five positions		Yes			Yes			Yes			Yes		
Product fixed effect		Yes			Yes			Yes			Yes		
Week specific fixed effect		Yes			Yes			Yes			Yes		
Products		1,932			1,835			851			564		
N		190,220			190,230			125,250			71,495		
R^2		0.09			0.08			0.06			0.07		

**=p<0.1, **=p<0.05, ***=p<0.001*
Standard errors clustered at the product level.

Table 14: Comparison of Effect Sizes Across Marketing Activities

Study	Marketing activity	Empirical context	Result	Location of result in study
Goldfarb and Tucker (2011a)	Online ads	Purchase intention of an ad during a privacy regulation change	An exposure to an ad increases purchase probability by 0.75%.	Table 5, Column 1
Goldfarb and Tucker (2011b)	Online ads	Purchase intention of a contextual ad	An exposure to an ad increases purchase probability by 0.75%.	Table 2, Column 4
Urban et al. (2014)	Online ads	Purchase likelihood when using vs not using ad morphing technology	Ad morphing increases purchase likelihood by 0.23% to 0.49%	Table 3 and 4
Goldfarb and tucker (2014)	Online ads	Large database of online ad field tests by a US media metrics agency.	Being exposed to ad increases purchase probability by 0.33%	Table 4, Column 5
Lewis (2014)	Email coupons	Internet grocer that runs a loyalty program and email promotions	Loyalty program increases purchase incidence by 0.5% from 19.2% to 19.7%. Email coupon increases purchase incidence by 0.2% from 19.7% to 19.9%	Table 7 and Table 8
Fong (2016)	Targeted email offers	Large online wine retailer	Purchase rate increased by 1.41% to 1.54% due to targeted email offers.	Table 5
Danaher et al. (2015)	Mobile coupons	Coupons for the stores of a large mall in a major city of a Western country.	0.22% increase in purchase probability due to mobile coupons	Table 7