

The Market for Fake Reviews

Sherry He
Anderson School of Management
UCLA*

Brett Hollenbeck
Anderson School of Management
UCLA†

Davide Proserpio
Marshall School of Business
USC‡

December, 2020§

[PRELIMINARY DRAFT: COMMENTS WELCOME]

*sherry.he.phd@anderson.ucla.edu

†brett.hollenbeck@anderson.ucla.edu

‡davide.proserpio@marshall.usc.edu

§Authors listed in alphabetical order. Corresponding author: Brett Hollenbeck. We thank the Morrison Center for Marketing Analytics for generous funding. We also thank Elizabeth Paz, Jason Lei, Yoon Sang Moon, and Rachel Ziffer for excellent research assistance and seminar participants at the Kellogg School of Management at Northwestern, Singapore Management University, Frankfurt School of Finance & Management, Olin Business School at Washington University in Saint Louis, USC Marshall marketing workshop, and the UCLA Anderson marketing workshop for helpful suggestions.

Abstract

We study the market for fake product reviews on Amazon.com. These reviews are purchased in large private internet groups on Facebook and other sites. We hand-collect data on these markets to characterize the types of products that buy fake reviews and then collect large amounts of data on the ratings and reviews posted on Amazon for these products, as well as their sales rank, advertising, and pricing behavior. We use this data to assess the costs and benefits of fake reviews to sellers and evaluate the degree to which they harm consumers. The theoretical literature on review fraud shows that there exist conditions when they harm consumers and conditions where they function as simply another type of advertising. Using detailed data on product outcomes before and after they buy fake reviews we can directly determine if these are low-quality products using fake reviews to deceive and harm consumers or if they are possibly high-quality products who solicit reviews to establish reputations. We find that a wide array of products purchase fake reviews including products with many reviews and high average ratings. Soliciting fake reviews on Facebook leads to a significant increase in average rating and sales rank but the effect disappears after roughly one month. After firms stop buying fake reviews their average ratings fall significantly and the share of one-star reviews increases significantly, indicating fake reviews are mostly used by low quality products and are deceiving and harming consumers. We also observe that Amazon deletes large numbers of reviews and we document their deletion policy.

1 Introduction

Online markets have from their first days struggled to deal with malicious actors. These include consumer scams, piracy, counterfeit products, malware, viruses, and spam. And yet online platforms have become some of the world’s largest companies in part by effectively limiting these malicious actors and retaining consumer trust. The economics of these platforms suggest a difficult tradeoff between opening the platform to outside actors such as third party developers and sellers and retaining strict control over access to and use of the platform. Preventing deceptive or fraudulent actions are key to this tradeoff. Third party participants may have strong incentives to manipulate platforms, such as increasing their visibility in search rankings via fake downloads (Li et al., 2016), increasing revenue via bot-driven advertising impressions (Crussell et al., 2014; Cho et al., 2015), manipulating social network influence with fake followers, manipulating auction outcomes, defrauding consumers with false advertising claims (Rao and Wang, 2017; Chiou and Tucker, 2018; Rao, 2018), or manipulating their seller reputation with fake reviews (Mayzlin et al., 2014; Luca and Zervas, 2016).

We study this last form of deception or fraudulent activity: the widespread purchasing of fake product reviews. Fake reviews may be particularly harmful because they not only deceive consumers into purchasing products or visiting firms like restaurants or hotels that may be of low quality and sold at inflated prices, they also erode the long-term trust in the review platforms that is crucial for online markets to flourish (Cabral and Hortacsu, 2010; Einav et al., 2016; Tadelis, 2016). Therefore, if user feedback and product reviews are not trustworthy, in addition to consumers being harmed platform values may suffer as well.

We study the effect of fake reviews on seller outcomes, consumer welfare, and platform value using the following research design. Despite this practice being unlawful, we document

the existence of a large and fast-moving online market for fake reviews.¹ Specifically, this market features product sellers posting in private online groups to promote their products and solicit willing customers to purchase them and leave positive reviews in exchange for compensation.² These groups exist for many online retailers including Walmart and Wayfair but we focus on Amazon because it is the largest and most developed market. We collect data from this market by sending research assistants into these groups to document what products are buying fake reviews and the duration of these promotions. We then carefully track these products' outcomes on Amazon.com including posted reviews, average ratings, prices, and sales rank. This is the first data of this kind, in that it provides direct evidence on both the fake reviews themselves and on detailed firm outcomes from buying fake reviews.³

We track these outcomes both in the short-term, i.e., right after the beginning of, and during the time these promotions are active, and long-term, i.e., after the buying of fake reviews has ended. This is for two main reasons. First, we observe substantial variation in the duration of these promotions which, in turn, can have different immediate and lasting effects on the outcomes we measure. Second, and as we explain later, tracking long-term effects can help us establish whether recruiting fake reviews is a harmful or harmless process for both the platform and consumers.

In general, because consumers value trustworthy information and e-commerce platforms value having good reputations, their incentives should be aligned in that they both want to avoid fake reviews. However, this may not always be the case. In particular, platforms may benefit from allowing fake positive reviews if these reviews increase their revenue via higher

¹The FTC has brought cases against firms alleged to have posted fake reviews, including a case against a weight-loss supplement firm buying fake reviews on Amazon in February 2019. See: <https://www.ftc.gov/news-events/press-releases/2019/02/ftc-brings-first-case-challenging-fake-paid-reviews-independent>

²The requirement that these reviews be positive to receive payment and the lack of disclosure are what differentiate “fake” reviews from the more innocuous practice of “incentivized” reviews. In the latter case, sellers offer discounted or free products to potential reviewers in exchange for posting a review. While incentivized reviews have raised concerns as well since they may be biased upward, in principle, they allow for authentic feedback and typically involve disclosure.

³On May 22, 2020, towards the end of our data collection window, the UK Competition and Markets Authority (CMA) announced it was opening an investigation into these practices. See: <https://www.gov.uk/government/news/cma-investigates-misleading-online-reviews>

sales or prices. It may also be the case that fraudulent reviews are not misleading in the aggregate if higher quality firms are more likely to purchase them than lower quality firms.

Indeed, Dellarocas (2006) shows that this is a possible equilibrium outcome and specifies under what condition: namely that demand increases convexly with respect to user rating. Due to the way ratings influence product rankings in search results in competitive markets, it is plausible that this property may hold. In fact, this result is related to classic signaling models of advertising for experience goods. Nelson (1970) and later Milgrom and Roberts (1986) show that in a signaling model there are separating equilibria where higher quality firms are more likely to advertise because the returns from doing so are higher for them. This is because they expect repeat business or word-of-mouth once consumers have discovered their true quality. To the extent that fake reviews generate higher sales which generate future online ratings, a similar dynamic may play out in our setting. In this case, fake reviews may be seen as harmless substitutes for advertising rather than as being deceptive or malicious. It is therefore ultimately an empirical question whether firms and regulators should view fake reviews as representing a significant threat to consumer welfare.

The above discussion motivates our research focus on the long-term outcomes associated with fake reviews. Are consumers ultimately harmed by fake reviews or are they mainly used by high quality products in a manner akin to advertising? Do fake reviews lead to a self-sustaining increase in sales and organic ratings or are they only useful for boosting short-term sales? These questions can be directly tested using the unique panel nature of our data. On the one hand, if products purchasing fake reviews continue to receive high ratings from consumers after they stop purchasing reviews and continue to enjoy high sales, it suggests the fake reviews are more akin to advertising and the incentive to buy reviews is stronger for high quality products. In this case, consumers may not be harmed by fake reviews and the platform might not want to regulate them too strictly. On the other hand, if when products stop buying fake reviews they start to receive mostly negative ratings from consumers, it suggests that these consumers have essentially been deceived and tricked into

buying low quality products. In this case, the consumer harm is clear and direct, and the platform reputation will be harmed, albeit more slowly and indirectly.

To perform this test we conduct a large-scale data collection from both private Facebook groups where fake reviews are purchased and from Amazon.com. We begin by providing a number of descriptive results on the markets for fake reviews and the products purchasing fake reviews. We continue by presenting a simple discussion of the costs and benefits of buying fake reviews. Then, we move to describe the short-term and long-term outcomes associated with fake review purchases we observe. Finally, we show how Amazon responds to sellers purchasing fake reviews.

Our first finding is that the markets for fake reviews are large and many, many sellers are seen participating in these markets. Over a 9-month period we document more than 20 private Facebook groups where sellers buy fake reviews. These groups average 16,000 members each and feature more than 500 posts per day from sellers soliciting reviews. Interested reviewers respond to these posts and then purchase the product, leave an authentic-seeming five-star review, and then are typically reimbursed via PayPal for the product cost plus tax and fees and in some cases an additional commission.

To understand why these markets are so large and active, we provide a short discussion of the costs and benefits of buying fake reviews based on our observation of how this market works in practice. We show two clear implications. First, for products with high margins the economics of buying fake reviews are quite favorable. Fake reviews are relatively cheap and generating just a few additional sales can justify their cost. Second, our model highlights why we observe the buying of positive fake reviews for the sellers' own products but not the buying of negative reviews for the sellers' competitors. The costs of buying negative fake reviews for competitors is significantly higher because it requires the seller buying the fake review to incur the full price of the competitor's product.

Using research assistants participating in these groups we collect data on a random sample of approximately 1,500 products observed soliciting fake reviews over a 9-month period. We

might expect these products to be new products with very few reviews, or else low quality products with very low ratings from organic reviews that must be countered with fake positive reviews. Instead, we find a wide assortment of product types in many categories, including many products with a very large number of reviews at the time we first observe them buying fake reviews. These products also tend to have high ratings, with a higher average rating than comparable products. We do also observe many new products with few or no reviews at the time, suggesting that some sellers are using fake reviews to get established on the platform. Almost none of the sellers purchasing reviews in these markets are well known brands, consistent with research showing that online reviews are more effective and more important for small independent firms compared to brand name firms (Hollenbeck (2018)). Finally, by matching seller names to trademark data we can determine that over 80% of these sellers are from China, particularly from the Shenzhen area.

We then track the outcomes of these products before and after the buying of fake reviews using the data collected from Amazon. In the weeks after they purchase fake reviews the number of reviews posted increases substantially, indicating that the Facebook posts are successful. Their average rating and share of five-star reviews also increase substantially in this period. Ratings increase by .08 stars on average, and the average number of reviews posted per week increases by 7, roughly doubling the number of reviews they receive compared to before soliciting fake reviews. We also observe a substantial increase in search position and sales at this time, as measured by Amazon sales rank. The increase in average ratings is short-lived, with ratings falling back to the previous level within 2 to 4 weeks, but the increase in the weekly number of reviews, sales rank, and position in search listings remain substantially higher more than four weeks later.

We also track the long-term outcomes associated with the buying of fake reviews. These are of particular interest because they are informative of the consumer harm or lack thereof associated with fake reviews. If organic reviews that follow fake review purchases tend to be low, it would indicate consumers were deceived into purchasing low quality products, and if

they tend to be high, it would suggest fake reviews are simply another form of advertising to boost sales but are not necessarily harmful. We find that the evidence primarily supports the consumer harm view. We track outcomes after the last observed post soliciting fake reviews and find that ratings tend to fall as soon as the seller stops buying fake reviews. The share of reviews that are one star increases substantially at this point as well. This pattern also holds for new products and those with few reviews and is in fact stronger for them, suggesting that even for this type of product fake reviews are associated with consumer harm. We find the largest increase in one-star reviews among new products, more expensive products, and those posting during the October through December period.

Finally, we document some facts regarding how the platform regulates fake reviews. We see that a very large share of reviews are ultimately deleted by Amazon. For the products in our data observed buying fake reviews, roughly one third of their reviews are eventually deleted. We can observe this by comparing our contemporaneous data collection used for the previous results to the set of reviews that remain online several months later. The reviews that are deleted are longer than non-deleted reviews, more likely to contain photographs, and more likely to be five stars. This suggests the reviews are deleted because of suspicion that they are fraudulent rather than because they are perceived to be of low quality. The bulk of deleted reviews are those that are posted within one to two months of the fake review solicitation that we observe, but they are deleted with an average lag of over 100 days, thus allowing the short term boost in average ratings and number of reviews that we previously showed. Altogether our results suggest that while Amazon's review deletion policy should reduce the long-term harm to consumers from fake reviews, it is inadequate in the sense that there is enough of a short-term boost in sales and ratings that firms find it advantageous to participate in this market and there is a clear consumer harm as shown in the subsequent increase in one-star reviews.

We contribute to the empirical study of fake online reviews. Prior work includes Mayzlin et al. (2014), who argue that in the hotel industry, independent hotels with single-unit owners

have a higher net gain from manipulating reviews. They then compare the distribution of reviews for these hotels on Expedia and TripAdvisor and find evidence consistent with review manipulation. Luca and Zervas (2016) uses Yelp’s review filtering algorithm as a proxy for fake reviews, and finds that these reviews are more common on pages for firms with low ratings, independent restaurants, and restaurants with more close competitors. Anderson and Simester (2014) show examples of a different type of fake review: customers rating apparel products on a brand site who never purchased those products. Ananthkrishnan et al. (2020) show using lab experiments that a policy of flagging fake reviews but leaving them posted can increase consumer trust in a platform.

We contribute to this literature in two primary ways. First, we document the actual market where fake reviews are purchased and characterize the sellers participating in this market. This data gives us a direct look at fake reviews, rather than merely inferring their existence. Second, we observe firm outcomes both before and after they purchase fake reviews. This allows us to characterize the costs and benefits to firms of fake reviews. We are also able to use the organic reviews posted after sellers stop buying fake reviews to understand whether and when consumers are harmed by this practice.

This research also contributes to the broader academic study of online reviews and reputation. By now, it is well understood that online reviews affect firm outcomes and improve the functioning of online markets (see Tadelis (2016) for a review.) There is also a growing body of research showing that firms take actions to respond to online reviews, including by leaving responses directly on review sites (Proserpio and Zervas, 2016) and changing their advertising strategy (Hollenbeck et al., 2019). There has always existed a difficult tension in the broader literature on online reviews, coming from the fact that sellers may manipulate or fake their reviews. By documenting the types of sellers purchasing fake reviews and the size and timing of their effects on ratings and reviews, we provide guidance to future researchers on how to determine whether review manipulation is likely in their setting.

Finally, we contribute to the literature on fraudulent activity in marketing. This research

studies practices such as fake news on social media (Chiou and Tucker, 2018), and deceptive online advertising (Rao, 2018; Wu and Geylani, 2020). The theoretical literature on deceptive practices has emphasized that there are generally conditions when these might make markets more efficient and possibly even benefit consumers (Dellarocas, 2006; Rhodes and Wilson, 2018). It is therefore up to empirical researchers to document the use of fraudulent practices to inform the debate on how regulators and firms should respond to these practices.

The rest of the paper proceeds as follows: Section 2 describes our data collection procedure and the settings of our paper; Section 3 present a discussion of the costs and benefits of buying fake reviews; Section 4 documents the short term changes in outcomes like average ratings, number of reviews, and sales rank in the weeks following the buying of fake reviews; Section 5 documents what happens to these outcomes after sellers stop buying fake reviews, Section 6 discusses the Amazon response to the problem of fake reviews; and, finally, Section 7 discusses our findings and provides concluding remarks.

2 Data and Settings

In this section, we document the existence and nature of online markets for fake reviews, and discuss in detail the data collection process and the data we obtained to study fake reviews and their effect on seller outcomes, consumer welfare, and platform value. Specifically, we collected data mainly from two different sources, Facebook and Amazon. From Facebook, we obtained data about sellers and products buying fake reviews, while from Amazon we collect reviews, ratings, and sales data.

2.1 Facebook Groups and Data

Facebook is one of the major platforms that Amazon sellers use to recruit fake reviewers. To do so, Amazon sellers create Facebook private groups where they promote their products by soliciting users to purchase their products and leave a five-star review in exchange for a full

refund and in some cases an additional payment. Discovering these groups is straightforward for interested reviewers; it only requires using the Facebook search engine to retrieve a list of them by searching for “Amazon Review”. We begin by documenting the nature of these markets and then describe how we collect product information from them.

Discovering groups We collected detailed data on Facebook group activity during a four months period, from Mar 28, 2019 to July 12, 2019. Each day, we collect the Facebook group statistics for the top-30 groups by search rank, and only including groups where sellers recruit fake reviewers. During this period, on average, we identify about 23 fake review related groups every day. These groups are large and quite active, with each having about 16,000 members on average, and about 568 fake review requests posted per day.

Within these Facebook groups, sellers can obtain a five-star review that looks organic. Figure 1 shows examples of Facebook posts aimed at recruiting reviewers. Usually, these posts contain words such as “need reviews”, “refund after pp (Paypal)” with product pictures. To avoid being detected by Amazon’s algorithm, sellers do not directly give reviewers the product link; instead, the sellers would ask reviewers to search for specific keywords associated with the product and find it by identifying the product title picture. On Amazon, the product title picture is a unique identifier for each product listing.

The vast majority of sellers buying fake reviews compensate the reviewer by refunding the cost of the product via a PayPal transaction after the five-star review has been posted (most sellers advertise that they also cover the cost of the PayPal fee and sales tax). Moreover, we observe that roughly 15% of products also offer a commission on top of refunding the cost of the product. The average commission value is \$6.24 with the highest observed commission for a review being \$15. Therefore, the vast majority of the cost of buying fake reviews is the cost of the product itself.

Reviewers are compensated for creating realistic seeming five-star reviews unlike reviews posted by bots or cheap foreign workers with limited English skills, which are more likely to

be filtered by Amazon’s fraud detection algorithms. First, the fact that the reviewer buys the product means that the Amazon review is listed as a “Verified Purchase” review; second, reviewers are encouraged to leave lengthy, detailed reviews including photos and videos to mimic authentic and organic reviews.⁴

Finally, sellers recruit only reviewers located in the United States, with an `amazon.com` account, and with a history of past good reviews.

This process differs from “incentivized reviews”, where sellers offer free or discounted products in exchange for reviews. The payment for “incentivized” reviews is not conditional on the review being positive, and these also typically involve disclosure in the form of a disclaimer contained in the review itself. Amazon has at times experimented with allowing incentivized reviews through its Vine program and through its “Early Reviewer Program,” but considers fake reviews a violation of its terms of service by both sellers and reviewers, leaving them subject to being banned from the platform if caught.

Discovering products To discover products that are promoted we rely on research assistants. We assign a few active Facebook groups to each one of them and ask them to select Facebook posts randomly. Given a Facebook post, the goal of the research assistants is to retrieve the Amazon URL of the product. To do so, they use the keywords provided by the seller. For example, in Figure 1, the search words would be “shower self”, “toilet paper holder” and “cordless vacuum”.

After a research assistant successfully identified the product, we ask them to document the following variables: search keywords, product ID, product subcategory (from the Amazon product page), date of the Facebook post, the earliest post date from the same seller for the same product (if older posts promoting the same product exist), and the Facebook group name.

We use the earliest Facebook post date as a proxy for when the seller began to recruit

⁴The fact that these fake reviews are from verified purchases indicates that an identification strategy like that used in Mayzlin et al. (2014) will not work in settings like ours.

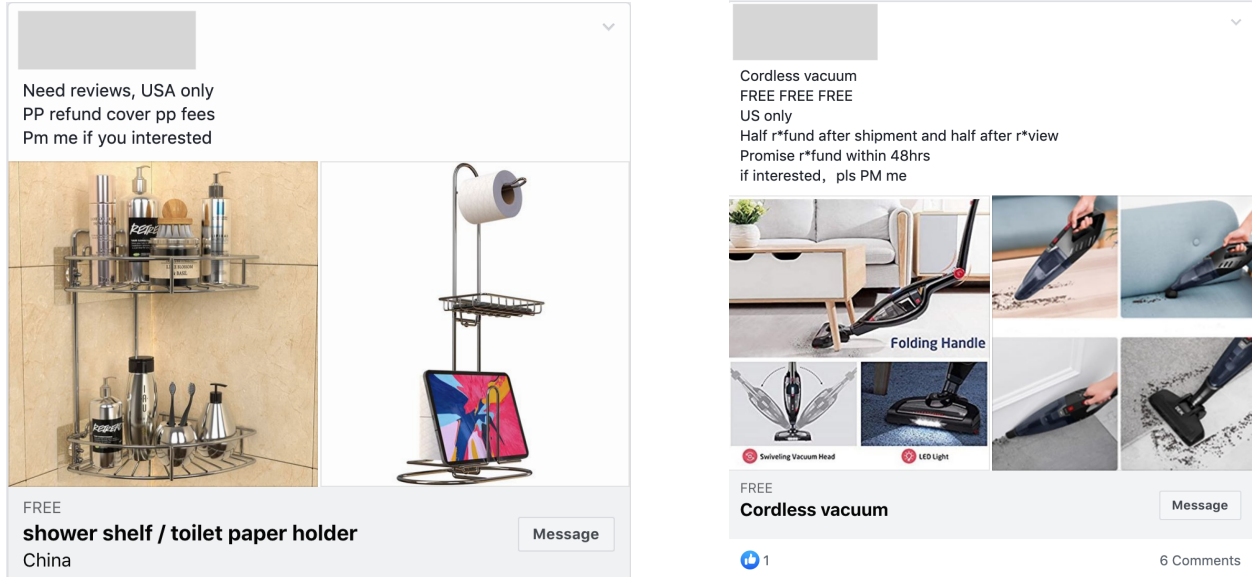


Figure 1: Examples of Fake Review Recruiting Posts

fake reviewers. To identify when a seller stops recruiting fake reviews for a product we continuously monitor each group and record any new posts regarding the same product by searching for the seller’s Facebook name or the product keywords. We then use the date of the last observed post as a proxy for when the seller stopped recruiting fake reviews.

We collect data from these Facebook fake review groups using this procedure on a weekly basis from October 2019 to June 2020 and the result is a sample of roughly 1,500 unique products. This provides us with the rough start and end dates of when fake reviews are solicited in addition to the product information.

2.2 Amazon Data

We use the Amazon information obtained by the research assistants to collect Amazon products information. Specifically, we collect information related to products in the same category of products asking for fake reviews, review and ratings, sales rank data, and sellers information. We describe this data in detail next.

Category Data Using the search keywords for the products buying fake reviews, we collect the search page results, i.e., the products returned as a result of the query, from Amazon. This information is useful to form a competitor set of products for each focal product.

We collect this information daily and store all consumer-facing information available on these pages including price, coupon, badges (best seller/ amazon’s choice), displayed rating, number of reviews, search page number, shipping information, whether the product is sponsored, and product position.

Review Data We collect the reviews and ratings for each of the products observed buying fake reviews on a daily basis. For each review we store the following variables: review data, ratings, product ID, review text, helpful votes.

Additionally, we collect the full set of reviews for each product on a bi-monthly basis. The reason for this is that it allows us to measure to what extent Amazon respond to sellers recruiting reviews by deleting reviews that it deems as potentially fake.

Product Rank Data We use the subcategory of the product asking for fake reviews to collect the corresponding subcategory Top-100 Best Seller product rank information (see Figure 2). The information we save includes rank, price, product ID, number of reviews, and displayed ratings. Additionally, we collect the same information for the parent category and corresponding children subcategories if they are likely to have a relationship (substitute or complements) with the focal product buying fake reviews. For example, products in the parent category “Children Dental Care” are likely to be complements or substitutes with products in the categories “electric toothbrushes” and “manual toothbrushes”. However, products in the parent category “Home” and “Kitchen Appliance” might not have a strong substitution or complementary relationship with the products in the subcategories “humidifier” and “dehumidifier”.

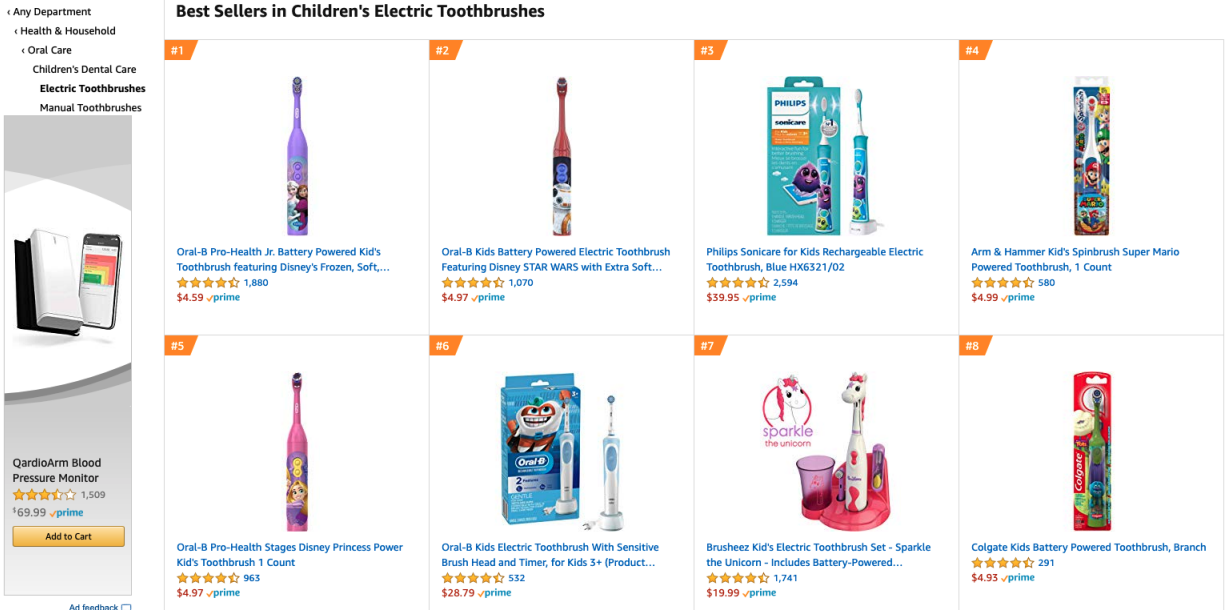


Figure 2: Screenshot of Amazon Subcategory Best Seller Page for Children's Electric Toothbrushes

Sales Rank Data We rely on Keepa.com and its API to collect sales rank data for the products soliciting fake reviews. On a bi-weekly basis, we collect this data for focal products and any products that appear in the category data discussed above. Amazon reports a measure called Best Seller Rank, whose exact formula is a trade secret, but which translates actual sales within a specific period of time into a ranking of products by sales levels.

Sellers data In addition to obtaining information about the focal products we identify as collecting fake reviews, we also collect data on other products sold by the same sellers.

2.3 Descriptive Statistics

In this subsection we provide descriptive statistics on the set of roughly 1,500 products collected between October 2019 to June 2020.

We use this sample of products to characterize the types of products that purchase fake reviews. On the one hand, we might expect these products to be primarily new products with few or no reviews who are trying to jumpstart sales by establishing an online reputation.

On the other hand, these might be products with many reviews and low average ratings, whose sellers resort to fake reviews to improve the product reputation and therefore increase sales.

Table 1: Focal Product Categories and Subcategories

Category	N	Subcategory	N
Beauty & Personal Care	193	Humidifiers	17
Health & Household	159	Teeth Whitening Products	15
Home & Kitchen	148	Power Dental Flossers	14
Tools & Home Improvement	120	Sleep Sound Machines	12
Kitchen & Dining	112	Men’s Rotary Shavers	11
Cell Phones & Accessories	81	Vacuum Sealers	11
Sports & Outdoors	77	Bug Zappers	10
Pet Supplies	62	Electric Back Massagers	10
Toys & Games	61	Cell Phone Replacement Batteries	9
Patio, Lawn & Garden	59	Light Hair Removal Devices	9
Electronics	57	Outdoor String Lights	9
Baby	42	Cell Phone Charging Stations	8
Office Products	30	Electric Foot Massagers	8
Automotive	29	Meat Thermometers & Timers	8
Arts, Crafts, & Sewing	21	Aromatherapy Diffusers	7
Camera & Photo	19	Blemish & Blackhead Removal Tools	7
Clothing, Shoes & Jewelry	14	Cell Phone Basic Cases	7
Computers & Accessories	12	Portable Bluetooth Speakers	7

Table 1 shows a breakdown of the top 20 categories and subcategories for our sample of products. The use of fake reviews is widespread across products and product categories. The top categories are “Beauty & Personal Care”, “Health & Household”, and “Home & Kitchen”, but the full sample of products comes from a wide array of categories as the most represented category still only accounts for just 13% of products, and the most represented product in

our sample, Humidifiers, only accounts for roughly 1% of products.

We observe substantial variation in the length of the Facebook fake reviews recruiting period, with some products being promoted for a single day and others promoted for over a month. The average length of the Facebook promotion period is 23 days and the median is 6 days.

Turning to the product age (measured using the first date the product was listed on Amazon), we find that the mean and median product age when they first begin soliciting fake reviews is 229 days and 156 days, respectively. This suggests that products collecting fake reviews are rarely new and without any reputation. Indeed, out of the 1,500 products we observe, only 17 of them solicit fake reviews in their first week after the product appears on Amazon, and only 94 solicit fake reviews in their first month.

Next, we compare the characteristics of our focal products to a set of competitor products. We define competitor products as those products that appear on the same page of search results for the same product keywords as our focal products. Even with these restrictions, we obtain a set of about 200,000 competitor products.

Table 2 compares the focal products with their competitors over several characteristics. We observe that while they are not extremely new when soliciting fake reviews, the focal products are significantly younger than competitor products, with a median age of roughly 5 months compared to 15 months for products not observed buying fake reviews. Moreover, our focal products charge slightly lower average prices than their competitors, with a mean price of \$33 compared to \$45 for other products. However, this result is mainly driven by the right tail of the price distribution. Fake review products actually charge a higher median price than their competitors but there are far fewer high priced products among the fake review products than among competitors. This may reflect the fact that a primary cost of buying fake reviews is compensating the reviewer for the price of the product. In other words, the more expensive a product is, the costly is to buy fake reviews.⁵

⁵We show this using a simple model in Section 3

Table 2: Comparison of Focal Products to Comparable Products

	Count	Mean	SD	Min	25%	50%	75%	Max
<i>Avg Rating</i>								
Fake Review Products	1,315	4.4	0.5	1.0	4.1	4.5	4.8	5.0
All Products	203,480	4.2	0.6	1.0	4.0	4.3	4.6	5.0
<i>Number of Reviews</i>								
Fake Review Products	1,425	183.1	493.5	0.0	10.0	45.0	167.0	8,777.0
All Products	203,485	451.4	2,619.0	1.0	13.0	59.0	250.0	348,107.0
<i>Price</i>								
Fake Review Products	1,425	33.4	45.0	0.1	16.0	24.0	35.0	1,199.0
All Products	236,542	44.7	154.8	0.0	13.0	21.0	40.0	26,999.0
<i>Sponsored</i>								
Fake Review Products	1,425	0.1	0.3	0.0	0.0	0.0	0.0	1.0
All Products	236,542	0.1	0.3	0.0	0.0	0.0	0.0	1.0
<i>Keyword Position</i>								
Fake Review Products	1,425	21.4	16.1	1.0	8.0	16.0	33.0	64.0
All Products	236,542	28.2	17.3	2.0	13.0	23.0	43.0	67.0
<i>Age (days)</i>								
Fake Review Products	1,305	229.8	251.1	0.0	77.0	156.0	291.0	3,554.0
All Products	153,625	757.8	797.1	0.0	257.0	466.0	994.0	7,831.0
<i>Sales Rank</i>								
Fake Review Products	1,300	73,292.3	151,236.4	2.0	7,893.3	26,200.5	74,801.5	2,111,680.0
All Products	5,647	89,926.1	323,028.9	3.0	5,495.0	21,610.0	72,563.5	13,034,008.0

Turning to ratings, we observe that products purchasing fake reviews have, at the time of their first observed post, relatively high product ratings. The mean rating is 4.4 stars and the median is 4.5 stars, which are both higher than the average ratings of competitor products. Although, we note that ratings may of course be influenced by previous unobserved Facebook campaigns. Only 14% of products have initial ratings below four stars and only 1.2% have ratings below three stars, compared to 19.5% and 3% for competitor products. Thus, it appears that products purchasing fake reviews do not seem to do so because they have a bad reputation.

We also examine the number of reviews. The mean number of reviews for focal products

is 183, which is driven by a long right tail of products with more than 1,000 reviews. The median number of reviews is 45 and roughly 8% of products have zero reviews at the time they are first seen soliciting fake reviews. These numbers are relatively low when compared to the set of competitor products which has a median of 59 reviews and a mean of 451 reviews. Despite these differences, it seems that most of the focal products are not buying fake reviews because they have very few or no reviews.

The last comparison is in terms of sales. We observe that the focal products have slightly lower sales than competitor products as measured by their sales rank, but the difference is relatively minor.

Turning to brand names, we find that almost none of the sellers in these markets are well-known brands. Brand name sellers may use other channels or avoid buying fake reviews altogether to avoid damage to their reputation. This result is also consistent with research showing that online reviews have larger effects for small independent firms relative to firms with well-known brands (Hollenbeck, 2018).

Finally, to better understand which type of sellers are buying fake reviews we collect one additional piece of information. We take the sellers' name from Amazon and check the U.S. Trademark Office for records on each seller. We find a match for roughly 70% of products. Of these products, the vast majority, 84%, are located in China, more precisely in Shenzhen or Guangzhou in the Guangdong province, an area associated with manufacturing and exporting.

To summarize, we observe purchases of fake reviews from a wide array of products across many categories. These products are slightly younger than their competitors but only a small share of them are truly new products. They also have relatively high ratings, a large number of reviews, and similar prices to their competitors.

3 The Simple Economics of Fake Reviews

We build on the results from the previous section on how the market of fake reviews functions, and briefly show the costs and benefits of buying fake reviews. We start by focusing on the costs to the seller of buying a fake review.

First, to buy 1 fake review, a seller must pay to the reviewer:

$$P(1 + \tau + F_{PP}) + Commission \quad (1)$$

Where P is the product's list price, τ is the sales tax rate, F_{PP} is the PayPal fee, and *Commission* refers to the additional cash offered by the seller, which is often zero but is sometimes in the \$5-10 range. After the reviewer buys the product, the seller receives a payment from Amazon of:

$$P(1 - c)$$

Where c is Amazon's commission on each sale. So the difference in payments or net financial cost of 1 review is:

$$P(1 + \tau + F_{PP}) + Commission - P(1 - c) = P(\tau + F_{PP} + c) + Commission$$

This is the share of the list price that is lost to PayPal, Amazon, and taxes, along with the potential cash payment. Along with this financial cost the seller bears the production cost of the product (MC), making the full cost of 1 fake review:

$$Cost = MC + P(\tau + F_{PP} + c) + Commission \quad (2)$$

If we define the gross margins rate as λ such that $\lambda = \frac{P-MC}{P}$, we can show that equation 2 becomes

$$Cost = P(1 - \lambda + \tau + F_{PP} + c) + Commission \quad (3)$$

This defines the marginal cost of a fake review to the seller. The benefit of receiving 1 fake review is a function of how many organic sales it creates Q_o and the profit on those sales, which is:

$$Benefit = Q_o P(\lambda - c) \quad (4)$$

where again c refers to Amazon's commission from the sale. Setting equations 3 and 4 equal allows us to calculate the breakeven number of organic sales Q_o^{BE} . This is the number of extra incremental sales necessary to exactly justify buying 1 fake review. If the seller does not offer an additional cash commission, and the vast majority of sellers do not, this can be written as:

$$Q_o^{BE} = \frac{1 - \lambda + \tau + F_{PP} + c}{\lambda - c} \quad (5)$$

Where the direct effect of price drops out and this is just a function of the product markup and observable features of the market. We take these market features as known:

- $\tau = .0656^6$
- $F_{PP} = 2.9\%$
- Amazon commission c varies by category but is either 8% or 15% in almost all cases.⁷

The result for products in the 8% commission categories is:

$$Q_o^{BE} = \frac{1.175 - \lambda}{\lambda - .08} \quad (6)$$

Thus the breakeven level of incremental sales needed to justify buying 1 fake review is a simple expression of a product's price-cost margin. It is clear that products with larger markups require fewer incremental organic sales in order to justify a fake review purchase. This is for two reasons that this analysis makes clear. First, because the cost of a fake review

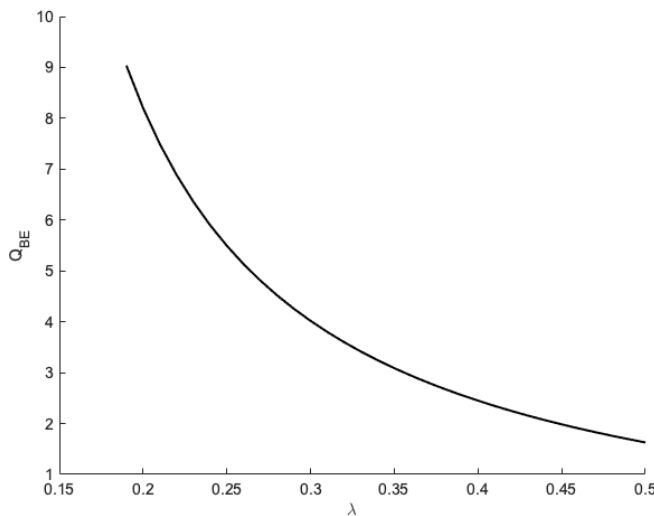
⁶<https://taxfoundation.org/2020-sales-taxes/>, we aggregate by taking an average of state and local sales taxes.

⁷<https://sellercentral.amazon.com/gp/help/external/200336920>

is lower since conditional on price the marginal cost is lower, and second because the benefit of an organic sale is larger for products with larger markups.

Figure 3 plots equation 6 where the X-axis is λ and the Y-axis is Q_o^{BE} . It shows that for products with relatively low markups the breakeven number of organic sales approaches 10 but for products with relatively high markups this number is below 2 and even below 1 for very high markups. Two implications immediately follow from this. First, the economics of fake reviews are quite favorable since a fairly small number of organic sales are needed to justify their cost. Second, this is especially the case for low quality products.

Figure 3: Organic Sales Needed to Justify 1 Fake Review



To take a concrete example, imagine two products that both list a price of \$25. Product A costs \$15 to produce and product B costs \$20 to produce because A is of lower quality than B. For product A $Q_o^{BE} = 2.4$ and for product B $Q_o^{BE} = 8.1$. The lower cost product needs far fewer organic sales to justify the expense of 1 fake review.

This analysis also makes clear why we are unlikely to observe fake negative reviews applied to competitor products, as in Luca and Zervas (2016) and Mayzlin et al. (2014). The cost of a fake review for a competitor product is significantly higher because it requires the firm buying the review to incur the full price of the competitor's product and the benefit is likely

to be lower because the negative effect on competitor sales is indirect and dispersed across potentially many other products.

4 Short-term Outcomes from Buying Fake Reviews

In this section, we describe the short-term effects associated with buying fake reviews for important outcomes such as ratings, reviews, and sales rank, as well as other marketing activities such as advertising and promotions. One of the unique features of our data is this detailed panel on firm outcomes observed both before and after sellers buy fake reviews. We therefore quantify the extent to which this practice affects ratings, reviews and sales, although we stress that these results are descriptive in nature. We do not observe the counterfactual outcomes in which these sellers do not buy fake reviews and so the outcomes we measure are not to be interpreted strictly as causal effects.

To evaluate these outcomes, we partition the time around the earliest Facebook recruiting post date (day 0) in 7-day intervals. For example, the interval 0 includes the days in the range $[0,7)$ and the interval -1 includes the days in the range $[-7,0)$. We then plot the quantity of interest for eight 7-day intervals before fake reviews recruiting start and four 7-day intervals after fake reviews recruiting starts. We focus on roughly four weeks after fake reviews recruiting starts because in this section we are interested in discussing short-term effects (recall that the mean length in days of a Facebook campaign is 23 days in our dataset).

Ratings and reviews We start by looking at how ratings and reviews change after the seller begins buying fake reviews. In the left panel of Figure 4 we plot the weekly average rating. Several interesting facts emerge from this figure. First, the average ratings increase by about 5%, from 4.3 stars to 4.5 stars at its peak, after Amazon sellers start recruiting fake reviewers. Second, this positive effect is short and it starts dissipating just two weeks after the beginning of the recruiting of fake reviews; despite this, even after four weeks after the

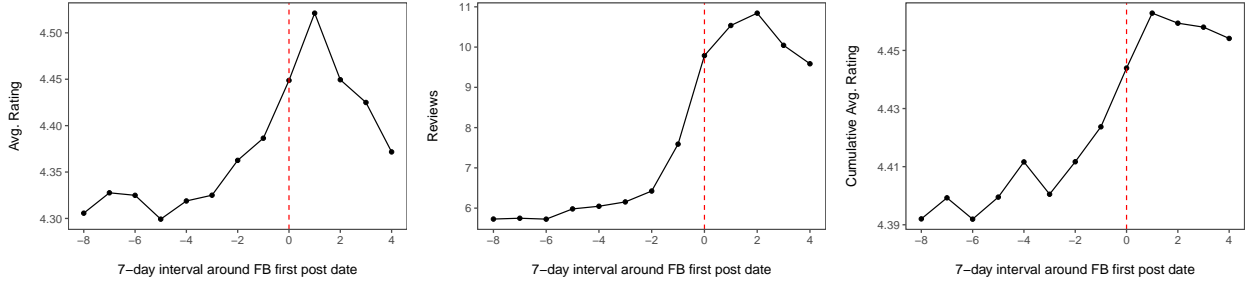


Figure 4: 7-day average ratings, 7-day average number of reviews, and cumulative average ratings before and after fake reviews recruiting begins. The red dashed line indicates the first time we observe Facebook fake review recruiting.

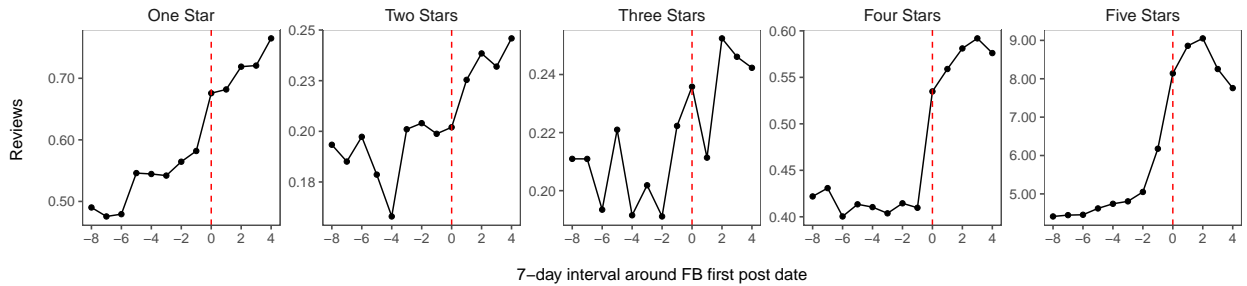


Figure 5: 7-day average number of reviews by star-rating, before and after fake reviews recruiting begins. The red dashed line indicates the first time we observe Facebook fake review recruiting.

beginning of the promotion, average ratings are still slightly higher than ratings in the pre-promotion period. Third, the average star-rating starts increasing roughly two weeks before the first Facebook post we observe, suggesting that we may not be able to capture with high precision the exact date at which sellers started promoting their products on Facebook. Despite this limitation, our data seems to capture the effect of recruiting fake reviewers fairly well.

Next, we turn to the number of reviews. In the middle panel of Figure 4, we plot the weekly average number of posted reviews. We observe that the number of reviews increases substantially around interval 0, nearly doubling, and confirming the expectation that recruiting fake reviewers is effective at generating new product reviews at a fast pace. Moreover, and differently from the average rating plot, the effect of recruiting fake reviewers on the number of reviews persist during the entirety of the post-promotion period. Finally,

Figure 4 confirms that we are not able to capture the exact data at which the Facebook promotion started.

Intuitively, because we observe an increase in average ratings and because average ratings are already pretty high (equal or above 4.3 stars), the increase in the number of reviews should be mostly driven by five-star reviews. To verify this hypothesis, in Figure 5 we plot the 7-day interval average number of reviews by star-rating. As expected, the largest change in the number of reviews is for five-star reviews, which increase by roughly 80% at its peak when compared with the level in the pre-promotion period.

Does this increase in positive reviews lead to higher displayed product ratings? To answer this question, in the right panel of Figure 4, we plot the cumulative average rating before and after the Facebook promotion starts. We observe a positive change centered around the beginning of the promotion and that stabilized for about two weeks after the promotion begins, after which the effect starts to dissipate.

Finally, we investigate how the effect of recruiting fake reviewers changes with the length of the campaign duration. As discussed in Section 2, there is substantial variation in the length of the Facebook promotion across the products in our dataset. We therefore plot the average rating, reviews, and cumulative average rating by Facebook campaign duration quartiles in Figure 6.⁸ Looking at average ratings (left panel), we observe that products with lower ratings in the pre-promotion period seem to be promoted for longer periods than products with higher ratings. Further, and as we would expect, the positive effect of fake reviews is much shorter for products that are promoted for shorter periods. Turning to the number of reviews (middle panel), we observe that products with fewer reviews are promoted for shorter periods; we argue that this might be related to the fact that changing the average ratings of products with a lot of reviews would require collecting more reviews (and therefore more time) than for product with a few reviews. Finally, the cumulative average rating plot (right panel) show patterns similar to those observed for the average ratings; however, the

⁸The four quartiles refer to Facebook campaign duration of 1 day, 2-6 days, 7-31 days and more than 31 days.

Facebook promotion effect last for a longer period for all products.

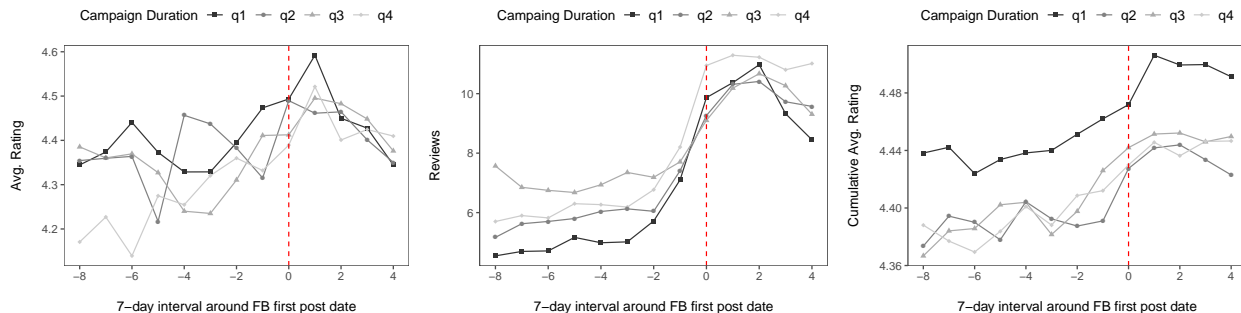


Figure 6: 7-day average ratings, 7-day average number of reviews, and cumulative average ratings before and after fake reviews recruiting begins, by Facebook campaign duration quartiles. The red dashed line indicates the first time we observe Facebook fake review recruiting.

Sales rank In the left panel of Figure 7 we plot the average log of sales rank. This figure reveals several facts. First, the figure shows that the sales rank of products that are eventually promoted is increasing between the intervals -8 and -3. This suggests that Amazon sellers tend to promote products for which sales are falling. Second, the effect of recruiting fake reviewers is negative (i.e. the sales rank decrease in magnitude and therefore the product sales increase) and large (in the post-promotion period sales rank goes back to the lowest levels observed in the pre-promotion period). Finally, the effect lasts almost all of the post-promotion period.

As we did for review and ratings, we investigate how the effect of recruiting fake reviewers changes with the length of the campaign duration. In the right panel of Figure 7, we plot the focal products sales rank by Facebook campaign duration quartiles. The first thing that emerges from this plot is that there is a clear difference in rank between products with different Facebook promotion lengths: products promoted for shorter periods have a higher sales rank. Second, we observe that the negative effect on sales rank is more pronounced for products running shorter campaigns, but also (and as expected) this effect lasts for shorter periods as the length of the campaign decreases.

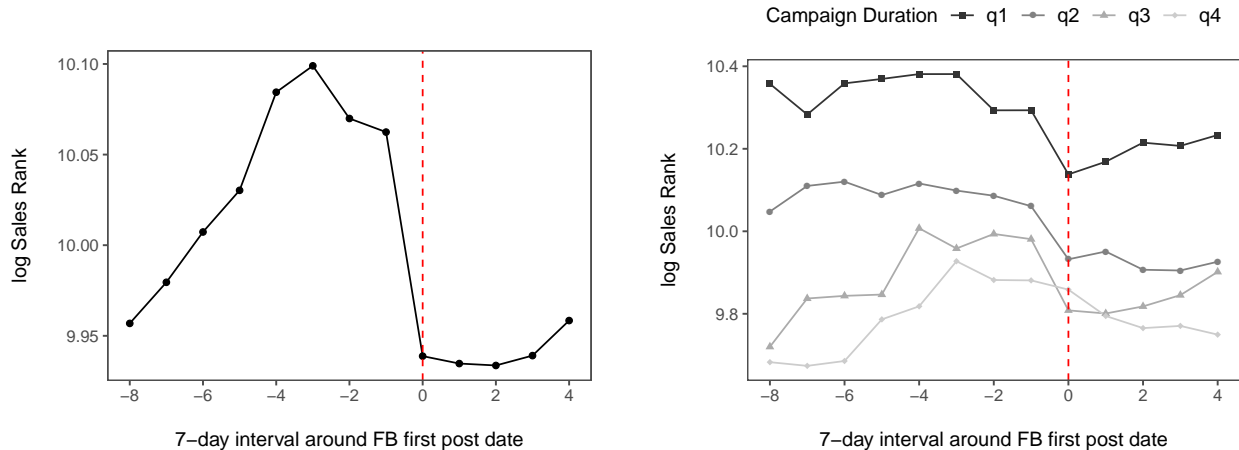


Figure 7: 7-day average sales rank before and after fake reviews recruiting begins overall products (left) and by campaign duration quartiles (right). The red dashed line indicates the first time we observe Facebook fake review recruiting.

Keywords search position So far we have shown that recruiting fake reviews improves ratings, reviews, and sales. One reason for observing higher sales is that higher ratings signal higher quality to consumers, who then are more likely to buy the product. A second reason that could drive sales is that products recruiting fake reviews will be ranked higher in the Amazon search results due to them having higher ratings and more reviews (both factors that are likely to play a role in determining a product search rank). To investigate whether this is the case, in Figure 8 we plot the search position rank of products recruiting fake reviews. In the left panel, we plot the search rank for all products, while in the right panel we divide products based on their campaign duration. We observe a large drop in search position rank corresponding with the beginning of the Facebook promotions, indicating that products recruiting fake reviews improve their search position substantially. Moreover, this seems to be a lasting effect as the position remains virtually constant in the post-promotion period. Turning to the heterogeneous effect by campaign duration (right panel), we observe similar patterns across all groups with minimal differences in the effect of the Facebook campaign.

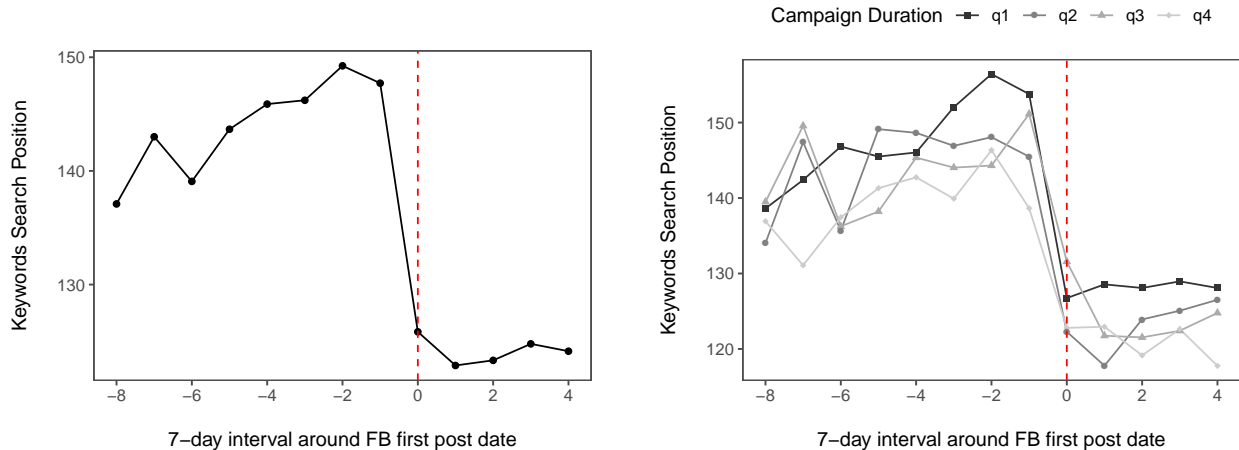


Figure 8: 7-day keywords search position before and after fake reviews recruiting begins overall products (left) and by campaign duration quartiles (right). The red dashed line indicates the first time we observe Facebook fake review recruiting.

Verified purchases and photos Next, we investigate the effect of recruiting fake reviewers on whether the review is written by someone who actually bought the product (Amazon ‘Verified purchase’ reviews) and the number of photos associated with the reviews. An important aspect of the market for fake reviews is that reviewers are compensated for creating realistic reviews, meaning they actually buy the product and can therefore be listed as a verified reviewer, and they are encouraged to post long and detailed reviews. We plot these two quantities in Figure 9. In the left panel, we show changes in 7-day interval average verified purchase reviews. Despite being quite noisy in the pre-promotion period, the figure suggests that verified purchases increase with the beginning of the promotion. Turning to the number of photos (right panel) we observe a sharp decline in the 7-day interval period $[-8,5]$ and a sharp increase that begins around interval -1 suggesting a positive effect associated with the beginning of the Facebook promotion.

Marketing activities Finally, we investigate to what extent recruiting fake reviewers is associated with changes in other marketing activities such as promotions (sponsored listings and coupons). We plot these quantities in Figure 10. We observe a substantial negative change in prices (left panel) that persists for several weeks. We also observe a persistent

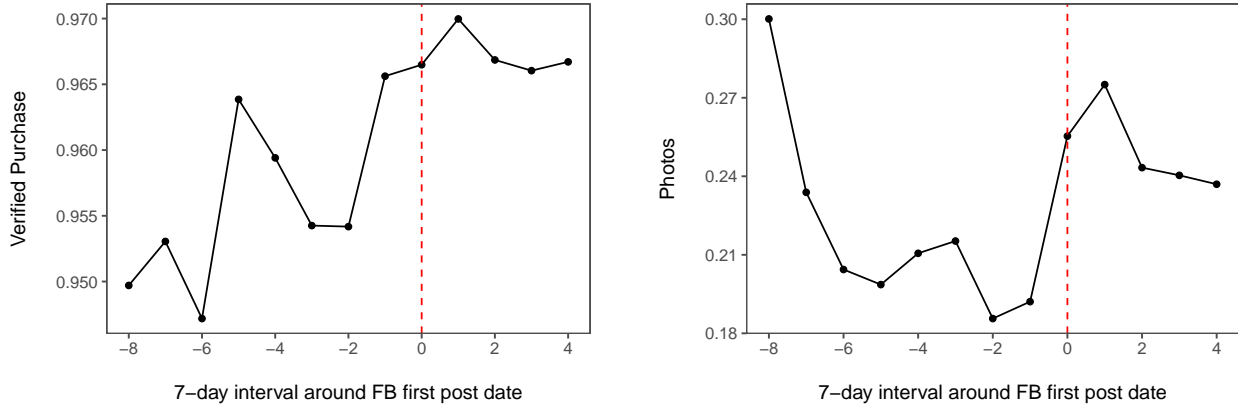


Figure 9: 7-day average verified purchase and number of photos before and after fake reviews recruiting begins. The red dashed line indicates the first time we observe Facebook fake review recruiting.

increased use of sponsored listings suggesting that Amazon sellers complement the Facebook promotion with advertising activities. This result contrasts with Hollenbeck et al. (2019) which finds that online ratings and advertising are substitutes and not complements in the hotel industry, an offline setting with capacity constraints. Finally, we observe a small negative effect (albeit noisy) on the use of coupons.

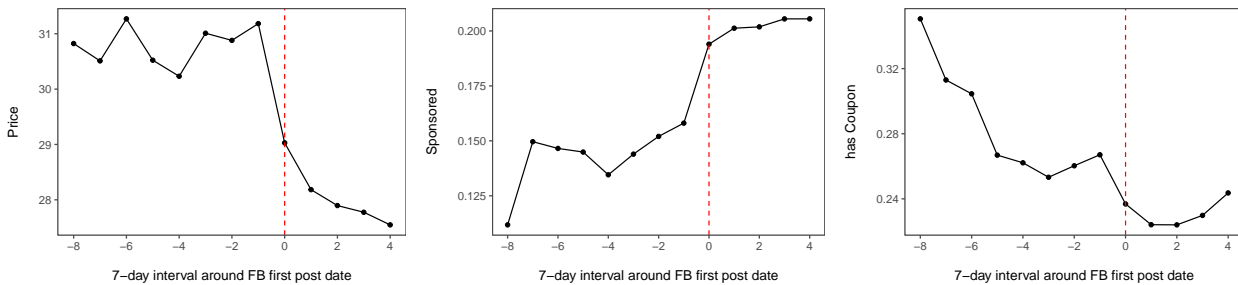


Figure 10: 7-day average sponsored listings and coupon. The red dashed line indicates the first time we observe Facebook fake review recruiting.

Regression Results We have so far shown the effects of fake reviews visually. We now show the same results in a regression context in order to test whether the changes in outcomes we observe are statistically meaningful relative to the normal amount of noise in the data as well as to quantify the size of these changes.

We take the 12 weeks before and after fake review recruitment starts and regress each

outcome variable on a dummy for the time period from 0 to 2 weeks afterward, as well as an additional dummy for the time period after that. This divides up our sample into three periods: a before period, a period in which short-term effects should be present, and a period in which longer-term effects should be present. In each case we include calendar month dummies and product fixed effects. In addition, we control for product age fixed effects to account for potential trends in the number of reviews, sales rank, or other variables over the life cycle of a product.

The results for each variable are shown in Table 3. Consistent with our visual analysis, we see significant short-term improvements in average rating, cumulative rating, number of reviews, sales, and search position (keyword rank). The increase in weekly average rating and cumulative rating are roughly .07-.08 stars, and the increase in the weekly number of reviews is 7. We also see significantly higher use of sponsored listings in this period and a significant increase in the share of reviews that are from verified purchases. There are also positive coefficients for the long-term dummy for the number of reviews and search position, confirming that these are long-lasting effects.

Table 3: Short-term Effects of Recruiting Fake Reviews

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Avg. Rating	Cum. Rating	Reviews	Log(Rank)	Keyword Rank	Sponsored	Coupon	Photos	Verified
0-2 Wks After FB	0.14*** (0.021)	0.16*** (0.023)	5.27*** (0.281)	-0.19*** (0.020)	-31.8*** (1.582)	0.039*** (0.006)	0.0026 (0.010)	0.022*** (0.004)	0.019*** (0.003)
>2 Wks After FB	0.032 (0.027)	0.056 (0.029)	2.77*** (0.358)	-0.082** (0.025)	-31.6*** (2.005)	0.040*** (0.008)	-0.013 (0.012)	0.0068 (0.005)	0.013** (0.004)
Product FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age in Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21567	21567	21567	19184	14721	14555	14555	21567	21567
R^2	0.283	0.254	0.641	0.824	0.579	0.616	0.576	0.189	0.220

Note: Robust standard errors in parentheses.
Significance levels: * p<0.05, ** p<0.01, *** p<0.001.

Overall, we observe that when sellers purchase fake reviews there is an immediate and substantial increase in the number of reviews they receive and average ratings. Additionally, these products increase their use of marketing activities such as sponsored listings at this time, and the net effect of these are a large increase in their sales that persist for several weeks. In the next section, we document the long-term trends in these outcome variables

after sellers stop buying fake reviews.

5 Long-term Outcomes from Buying Fake Reviews

In this section, we describe what happens after sellers stop buying fake reviews. Using the procedure described in section 2.1 we construct an end date for each product after they are no longer observed recruiting fake reviews. In particular, we are interested in using the long-term outcomes for these products to assess the potential consumer harm from fake reviews. For instance, if we observe these products continue to receive high organic ratings and high sales after recruiting fake reviews stops, we might conclude that fake reviews are a potentially helpful way to solve the cold start problem of selling online with limited reputation. If, by contrast, we see a large number of one-star reviews and declining sales and ratings it suggests that the sellers buying fake reviews are using them to mask low quality products and deceive consumers. An increase in one-star reviews indicates that these products are either low quality or overpriced relative to their quality. This analysis is also important from the platform’s perspective; an increase in one-star reviews should indicate fake reviews are a major problem since they reflect negative consumer experiences that should erode the sense of trust the platform’s reputation system is meant to provide.

We therefore track the long-term trends for ratings, reviews, and sales rank. We also track the share of reviews that come with ratings of one star, the lowest possible rating, as an indicator of low product quality or consumers who feel they have been deceived into buying the product. Lastly, we perform a detailed text analysis of the post-recruiting one-star reviews to see if they are distinctive compared to other one-star reviews and, if so, what text features they are associated with.

Similar to Section 4, we partition the time around the last Facebook recruiting post date (day 0) in 7-day intervals, and plot the quantity of interest for four 7-day intervals before fake reviews recruiting stop (thus covering most of the period where products recruited

fake reviews) and eight 7-day intervals after fake reviews recruiting starts. Doing so, we are comparing the Facebook promotion effect (negative intervals) with the post-promotion effect (positive intervals) where fake recruiting stopped.

Ratings and Reviews The long term effects of fake reviews on ratings and reviews are shown in Figure 11. We observe that the effect of buying fake reviews is fairly short. After one-to-two weeks from the end of the Facebook promotion, both the number of reviews and average ratings (left and middle panel, respectively) start to decrease substantially. The right panel of Figure 11 clearly explains why the average rating is decreasing. The share of one-star reviews starts to increase considerably once recruiting fake reviews stops. Interestingly, these products end up having average ratings that are significantly worse than when they started recruiting fake reviews (approximately interval -4).

We next plot these outcomes for different duration of the Facebook campaign in Figure 12. Focusing on reviews (left panel), we observe that while all products experience a reduction in the number of reviews they receive, the decrease for products with shorter campaign duration is less pronounced. In fact, these products seem to continue to receive more reviews than when the Facebook promotion started. Instead, products with longer campaign duration are in a worse situation with the number of reviews going back to pre-Facebook campaign levels. We observe similar patterns for average ratings and share of one-star reviews (although all products seem to go back to levels worse than the pre-Facebook promotion period). These findings suggest that there might be a negative correlation between the quality of the product promoted and how long they are promoted.

Finally, we explore the long-term effect on the share of one-star reviews for different types of products. It may be the case that while one-star reviews increase after fake review purchases stop, certain products are able to retain high ratings. For example, new products (i.e., products with few reviews or that have been listed on Amazon for a brief period of time) might use fake reviews to bootstrap their reputation that then they can sustain if these

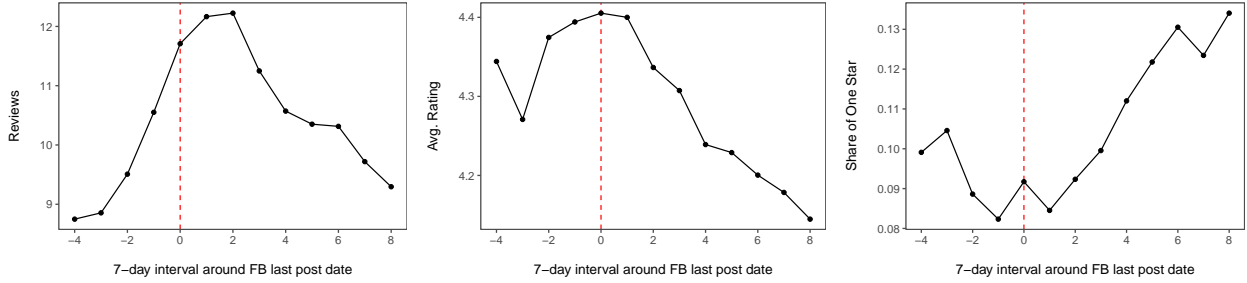


Figure 11: 7-day average number of reviews, average ratings, and average share of one-star reviews before and after fake reviews recruiting stops. The red dashed line indicates the last time we observe Facebook fake review recruiting.

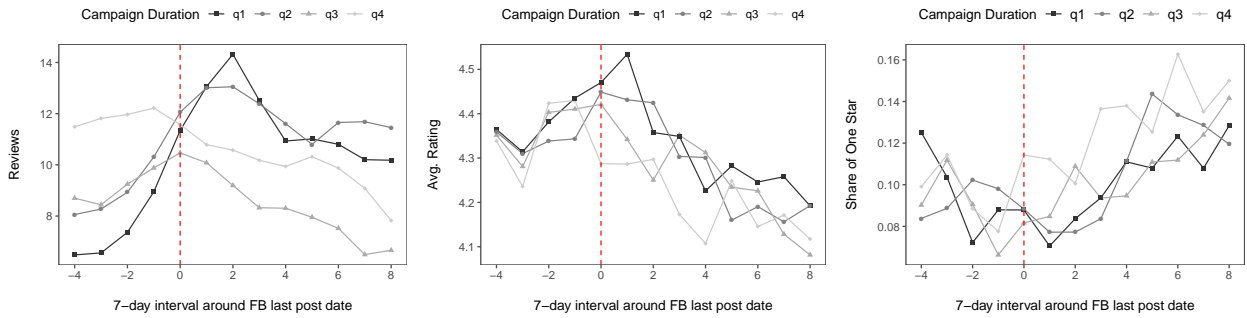


Figure 12: 7-day average number of reviews, average ratings, and average share of one-star reviews before and after fake reviews recruiting stops, by Facebook campaign duration quantities. The red dashed line indicates the last time we observe Facebook fake review recruiting.

products are high quality products.

To test this hypothesis, we replicate the right panel of Figure 11 but segmenting products by number of reviews and age. The left panel of Figure 13 shows how the share of one-star reviews changes for products with fewer than 50 reviews at the time they started recruiting fake reviews compared to all other products. The products with few reviews show a somewhat sharper increase in one-star ratings. The right panel of Figure 13 shows the same outcome, but for products that have been listed on Amazon for fewer than 60 days when they started recruiting fake reviews (very young products), compared to all other products. The young products have a much larger increase in one-star reviews compared to other products, with more than 20% of their ratings being one-star ratings two months after they stop recruiting fake reviews. Overall, these results do not support our hypothesis. Instead, they suggest

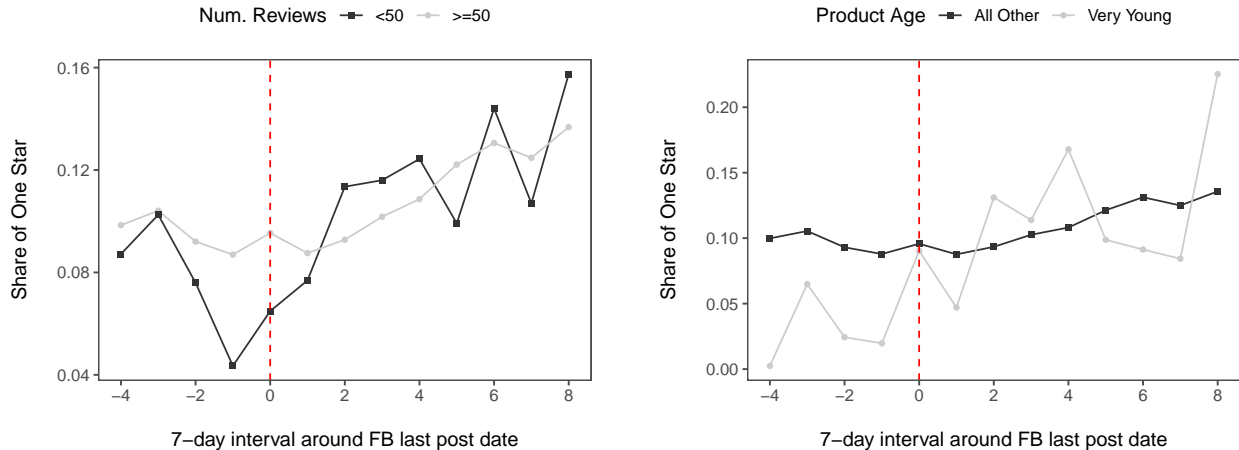


Figure 13: 7-day average sales rank before and after fake reviews recruiting stops for sub-samples of products with few reviews and new products (posted less than 60 days). The red dashed line indicates the last time we observe Facebook fake review recruiting.

that new products recruiting fake reviews are likely to be low quality products that use fake reviews to inflate their ratings and sales.

Sales Rank Figure 14 shows the average log of sales rank. These two figures reveal several facts. First, sales decline substantially after the last observed Facebook post. This suggests that the effect of recruiting fake reviews is not long-lasting. It does not create a self-sustaining set of sales and positive reviews, in other words.

Keywords search position Figure 15 shows the average keywords search position over all products (left panel) and by campaign duration (right panel). We observe that after the Facebook campaign stops, the search position trend reverses. However, the increase in search position is slow, therefore all products enjoy better search position rank for a relatively long period after fake review recruiting stops. Turning to the figure by campaign duration (right panel), we observe a similar pattern that confirms the almost steady state of search position rank after fake reviews recruiting stops.

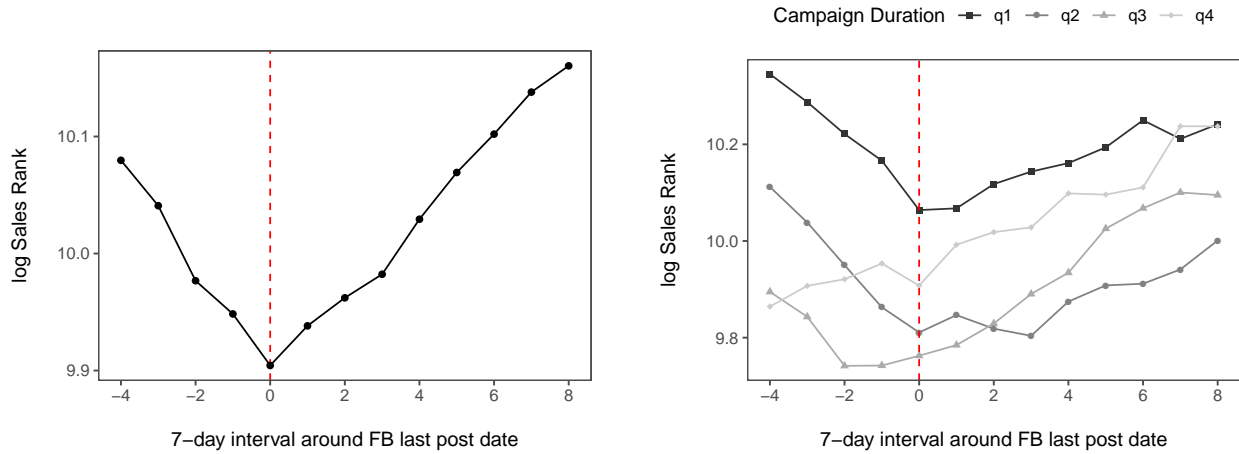


Figure 14: 7-day average sales rank before and after fake reviews recruiting stops overall focal products (left) and by campaign duration quartiles (right). The red dashed line indicates the last time we observe Facebook fake review recruiting.

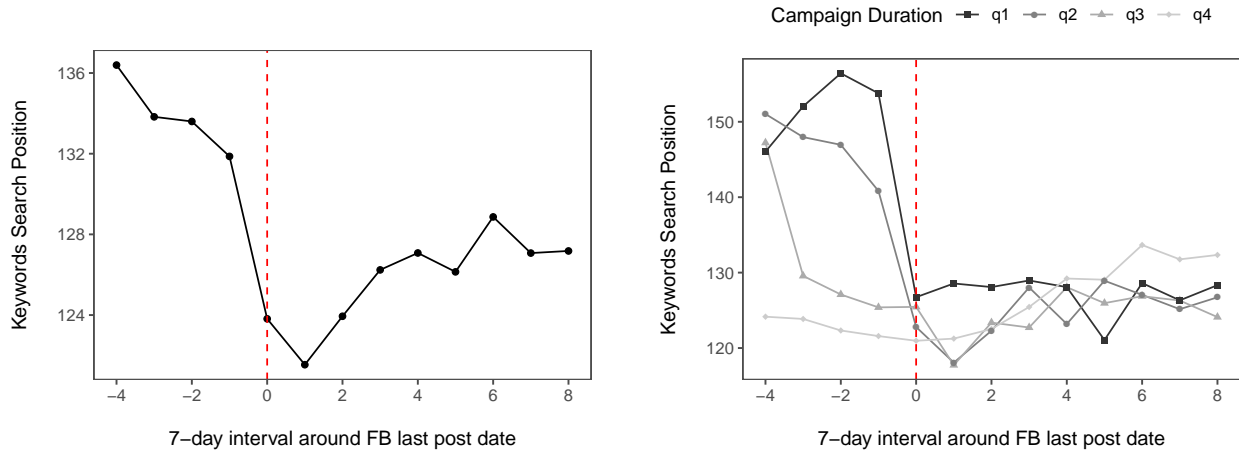


Figure 15: 7-day keywords search position before and after fake reviews recruiting stops overall products (left) and by campaign duration quartiles (right). The red dashed line indicates the last time we observe Facebook fake review recruiting.

5.1 Text Analysis

So far, we have shown increases in the number and share of one-star reviews to provide evidence that consumers are harmed by the fake reviews that Amazon sellers buy through Facebook groups. Here, we provide additional evidence in support of this hypothesis by analyzing the text of these reviews.

We use state-of-the-art machine learning algorithms to analyze and compare the text

of a random sample of the one-star reviews that product recruiting fake reviews received once the Facebook campaign ended with a random sample of one-star reviews of comparable products. The goal of this analysis is two-fold. First, we want to check whether negative reviews of the focal products are distinctive compared to those of competitors; second, if they are indeed different, we want to identify the text features that differentiate them.

We start by sampling 5,000 one-star reviews for each product type: those recruiting fake reviews and their competitors, defined as products showing up on the same results page for a keyword search. Then, we train a text-based classifier to predict whether each one-star review is from a product recruiting fake reviews or not. Following standard practice, we split the dataset into an 80% training sample and a 20% test sample. We present the results using a Naive Bayes classifier based on tf-idf. Depending on the configuration of the classifier (we can change the number of text features used by the classifier by removing very rare and very popular words), we achieve an accuracy rate that varies between 63% and 75% and a ROC-AUC score that varies between 69% and 83%.⁹ These results suggest that the text of the reviews is sufficiently distinctive for the classifier to distinguish between the two types of one-star reviews. In other words, despite the products themselves being highly similar and the reviews having the same star-rating, the one-star reviews for products that had recruited fake reviews used a significantly different set of terms from one-star reviews for products that did not recruit fake reviews.

We next look at what are the most predictive text features for distinguishing the two product types. In Table 4, we report the top-30 features for the model achieving an accuracy rate and ROC-AUC score of 63% and 69%, respectively. What emerges from this table is that one-star reviews written for products recruiting fake reviews are predicted by text features mostly related to product quality (qualiti, stop work, work, etc.), value/price (waste money, money, disappoint, etc.), and past reviews (review) that may have deceived consumers into buying these products; instead, competitors' one-star reviews are predicted by text

⁹Other types of classifiers lead to similar performance.

features mostly related to idiosyncratic product characteristic (second attach, fade, reseal, etc.). Overall, these results add further evidence that consumers who bought products that recruited fake reviews felt deceived in thinking that the products were of higher quality than they really were.

Table 4: Top-30 Most Predictive Features of the Text Classifier

Products...	Text Features
recruiting fake reviews	work, product, money, return, use, time, stop, wast, month, would, like, wast money, charg, even, broke, stop work, week, disappoint, good, back, light, first, tri, bought, qualiti, review, turn, batteri, recommend, great
not recruiting fake reviews	reseal, command, bang, fixtur, apart piec, septemb, product dont, fade, ignit, use never, use standard, terrier, compani make, desktop, love idea, wifi connect, bead, solar panel, inexpens, within year, return sent, compani product, second attach, pure, cycl, thought great, solar charg, blame, bought march, price paid

5.2 Which product characteristics predict consumer harm?

Next, we take the number and share of one-star reviews as indicators of low quality products and plausible consumer harm, and study what product characteristics predict increases in these variables after fake review recruiting stops. To do so, we start by measuring changes in these variables between the post and pre-fake review recruiting periods using a 20-week window (10 weeks for each period). Additionally, we allow for possible lagged effects by excluding the 4-week window around the first Facebook post. The changes are thus defined using weeks $[-12,-2]$ for the pre-fake review recruiting period and $[2,12]$ for the post-fake review recruiting period.

We then regress these changes on product characteristics at the time of the first observed fake review post. We include product age, price, category, total number of reviews, sales rank, whether the seller is known to be located in China, and when the post occurred. This last variable is broken into seasons which we describe as Spring (March through May), Summer (June through September), Fall (October through December), and Winter (January through February).

These regressions are meant to capture correlations only since the decision to purchase fake reviews is endogenous and related to product characteristics. However, these correlations are informative and can show which characteristics predict larger or smaller consumer harm from fake reviews. We report these results in Table 5. We observe that the estimates are not sensitive to the definition of consumer harm (either the number or the share of 1-star reviews). We find that younger and more expensive products see significantly larger increases in 1-star reviews after they stop recruiting fake reviews. In addition, lower ranked products and products in several categories have significantly larger increases in 1-stars. There are also significantly larger increases for products that recruit fake reviews in the Spring and Fall. The largest increase is for products posted in the Fall (October through December) when online shopping for Christmas may motivate low quality products to temporarily boost sales by buying fake reviews.

Overall, these results suggest that consumer harm can increase depending on both product characteristics and time of the purchase.

6 Amazon Response

In this section we provide evidence that Amazon is aware of the fake review problems and it is taking some measures to remove these reviews.

While we cannot observe reviews that are filtered by Amazon’s fraud detection practices and never made public, by collecting review data on a daily basis and bi-monthly we can

Table 5: Long-Term Outcomes After Recruiting Fake Reviews

	(1) Change in # 1 Star Reviews	(2) Change in Share of 1 Star Reviews
Log(Age)	-0.17** (0.054)	-0.035*** (0.009)
Log(# Reviews)	-0.060 (0.035)	0.0022 (0.006)
Log(Price)	0.11* (0.052)	0.027** (0.009)
Log(Sales Rank)	-0.21*** (0.031)	-0.0086 (0.005)
China	-0.021 (0.073)	-0.026* (0.012)
Spring	0.30*** (0.087)	-0.0024 (0.015)
Summer	0.076 (0.217)	-0.059 (0.037)
Fall	0.46*** (0.112)	0.015 (0.019)
Computers & Accessories	0.41 (0.363)	0.12* (0.062)
Office Supplies	-0.22 (0.427)	0.12 (0.073)
Beauty & Personal Care	0.64* (0.309)	0.091 (0.053)
Arts, Crafts, & Sewing	0.57 (0.533)	0.081 (0.091)
Clothing, Shoes & Jewelry	-0.25 (0.827)	0.0045 (0.141)
Electronics	0.27 (0.307)	0.070 (0.052)
Health & Household	0.54 (0.312)	0.089 (0.053)
Home & Kitchen	0.45 (0.306)	0.067 (0.052)
Automotive	0.23 (0.358)	0.044 (0.061)
Toys & Games	0.81* (0.336)	0.12* (0.057)
Patio, Lawn & Garden	0.42 (0.327)	0.098 (0.056)
Sports & Outdoors	0.39 (0.328)	0.090 (0.056)
Tools & Home Improvement	0.41 (0.315)	0.082 (0.054)
Baby	0.095 (0.428)	-0.017 (0.073)
Observations	948	948
R^2	0.129	0.059

Note: Robust standard errors in parentheses.

Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

observe if reviews are posted and then later deleted. To characterize Amazon’s current policy we present data on the share of reviews being deleted, the timing of their deletion, and the characteristics of deleted reviews.

How Many Reviews are deleted? We start by measuring the difference in the number of posted reviews between consecutive days for all products in our dataset (i.e., products that we observe and we don’t observe buying fake reviews). This difference is zero in 66.7% of cases, positive in 30.75% of cases, and negative in 2.55% of cases. The ratio of the total observed deleted reviews over the total additional (new) reviews is 0.46, which means that on average, one out of two reviews is deleted by Amazon.¹⁰

Next, we then compare the change in number of posted reviews between consecutive days for products we observe buying fake reviews and products we don’t observe buying fake reviews. We intend to document Amazon’s response to the focal products and non-focal products and infer whether Amazon can detect fake review well. For products we observe buying fake reviews, we observe a decrease in the number of reviews 9.6% of the time, an increase 49.7% of the time, and no change 40.7% of the time. For the remaining products, we observe a decrease in the number of reviews 2.5% of the time, an increase 30.6% of the time, and no change 66.9% of the time. This means that, compared to products we don’t observe buying fake reviews, products that buy fake reviews are both more likely to see their reviews increase and decrease. This is likely because these products buy fake reviews (which increase reviews) and then these reviews are deleted by Amazon.

We can also confirm this by looking at the ratio of the total observed deleted reviews over the total additional (new) reviews. For products we observe buying fake reviews the ratio is 0.47, while it is 0.47 for products that we don’t observe buying fake reviews. Besides suggesting that it is much more common for reviews to be deleted for products buying fake reviews, the ratio also tells us that these deleted reviews are a large fraction of all the reviews

¹⁰It is worth noting that because review deletion can coincide with new reviews being posted, this is a lower bound for the amount of reviews that Amazon deletes.

posted for these products. Instead, Amazon deletes reviews for products we don't observe buying fake reviews much more rarely but when it does so, these deleted reviews represent a large share of all posted reviews. This suggests that Amazon can identify fake reviews to some extent. For the products we don't observe purchasing fake reviews, on average Amazon's review deletion behavior is different and these deletions are likely motivated by factors unrelated to review fraud, for example review consolidation across different variations of the same product.

Characteristics of Deleted Reviews In Table 6, we report the mean and standard deviation for several review characteristics for existing reviews and deleted reviews, respectively. Following the literature on fake reviews, we focus on characteristics that are often found to be associated with fake reviews. Specifically, we focus on whether the reviewer purchased the product through Amazon (verified purchase), review rating, number of photos associated with the review, whether the reviewer is part of Amazon's "Early Reviewer Program", i.e. is one of the first users to write a review for a product the length of the review title, and the length of the review.¹¹

We find that deleted reviews have higher average ratings than non-deleted reviews. This is driven by the fact that the vast majority of deleted reviews are five-star reviews (see Figure 16).

Deleted reviews are also associated with more photos, shorter review titles, and longer review text. In general, we might expect longer reviews, those that include photos, and those from verified purchases to be less suspicious. The fact that these are more likely to be deleted suggests that Amazon is fairly sophisticated in targeting potentially fake reviews.¹² Finally, we find no difference for whether the review is associated with a verified purchase

¹¹For more details about the "Early Reviewer Program", we refer the reader to <https://smile.amazon.com/gp/help/customer/display.html?nodeId=202094910>

¹²This result contrasts with Luca and Zervas (2016), who find that longer reviews are less likely to be filtered as fake by Yelp.

or tagged as “Amazon Earlier Reviews”.¹³

Table 6: Summary Statistics of Existing Reviews

	All Reviews Mean	Deleted Reviews Mean	t-stat
Verified purchase	0.96 (0.19)	0.97 (0.17)	10.37
Review rating	4.41 (1.24)	4.57 (1.10)	34.05
Number of photos	0.22 (0.75)	0.28 (0.85)	27.65
Early reviewer	0.01 (0.08)	0.00 (0.00)	-18.01
Title length	19.81 (15.15)	16.94 (15.40)	-54.01
Review length	225.07 (242.11)	240.15 (233.50)	17.10

Standard deviations in parentheses.

When Reviews are Deleted? Finally, we analyze when Amazon deleted fake reviews for focal products. We do so by plotting the number of products for which reviews are deleted over time relative to the first Facebook post, i.e., the beginning of the buying of fake reviews. To do so, we partition the time in days around the Facebook first post, and then plot the number of products for which reviews are deleted. Because our focal products have different campaign duration, we do this analysis by campaign duration quartiles. Figure 17 shows the results of this analysis. What clearly emerges from this figure is that Amazon starts deleting reviews for more products after the Facebook campaign begins and often it does so only after the campaign terminated. Indeed, it seems that most of the review deletion happens during the period covering the two months after the first Facebook post date, but most campaigns are shorter than a month. Simple calculations suggest that reviews are deleted only after a quite large lag (when compared with the duration of the Facebook campaign). The mean

¹³We find that Amazon does not delete any reviews tagger as “Amazon Earlier Reviews” potentially because Amazon process to identify and select early reviewers drastically reduces the possibility of these reviews to be fake.

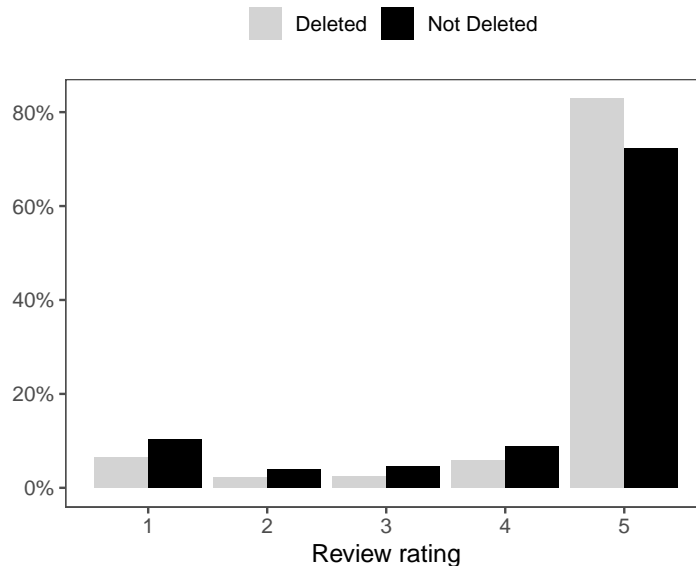


Figure 16: Histogram of Review Ratings of Reviews

time between when a review is posted and when it is deleted is over 100 days, with a median time of 53 days.

This analysis suggests the deleted reviews may be well-targeted at fake reviews, but that there is still a significant lag between when the reviews are posted and when they are deleted; and this lag allows sellers buying fake reviews to enjoy the short-term benefits of this strategy discussed in Section 4.

7 Discussion and Conclusions

Fake reviews are becoming the de-facto standard for online sellers to manipulate their reputation on online platforms. In this paper, we study the market for fake reviews of one of the world’s largest e-commerce platforms, Amazon. We do so by collecting data from sellers recruiting fake reviews from Facebook groups and combine it with several Amazon sellers’ performance metrics.

We start by showing that the market for fake reviews is large and fast-moving, with hundreds of sellers approaching thousands of potential reviewers every day. We then track a

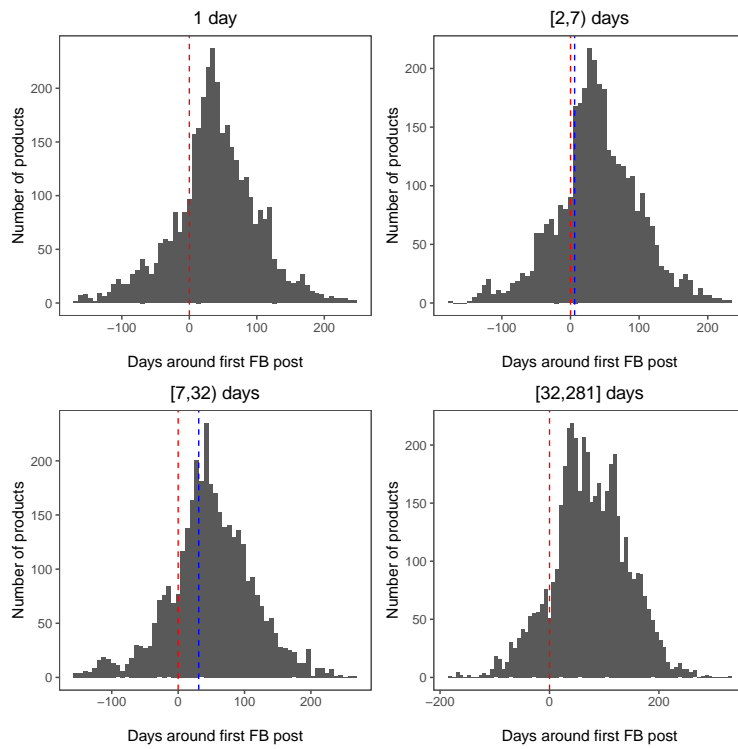


Figure 17: Number of products for which reviews are being deleted over time relative to the first Facebook post date. The red dashed line indicates the first time we observe Facebook fake review recruiting, and the blue dashed line indicates the last time we observe Facebook fake review recruiting.

random sample of these sellers and products to study the effect of recruiting fake reviews on sellers' outcomes, consumer welfare, and platform value. We study these effects both in the short-term, i.e., right after the beginning of, and during the time these promotions are active, and in the long-term, i.e., after the fake review solicitation has ended. These two analyses allow us to study the immediate and lasting effects of fake reviews on sellers' outcomes, and they are important to understand whether these reviews are harming consumers and online platforms or not.

Several interesting findings emerge from our analyses. First, we find that products promoted on Facebook groups span a wide variety of categories and they already have many reviews, with average ratings that are often higher than those of comparable products not recruiting fake reviews. Second, we find that the Facebook promotion is extremely effective at improving several sellers' outcomes such as number of reviews, ratings (average and displayed), search position rank, and sales rank, in the short-term. However, these effects are short-lived as many of these outcomes return to pre-promotion levels a few weeks after the fake reviews recruiting stops. This is explained by an increase in the share of one-star reviews once the Facebook promotion ends.

Overall, these results suggest that these fake reviews deceive consumers into buying products that then turned out to be of lower quality than expected. Therefore, our results are consistent with a story in which fake reviews are harmful to both consumers and the platform itself.

Finally, we document the platform response to sellers recruiting fake reviews. We find that Amazon responds by deleting reviews at a very high rate. Moreover, we find that Amazon's response is quite sophisticated. The timing of review deletion suggests that it is able to identify which sellers and reviews are likely fake despite these reviews being very hard to discover (these reviews always come from verified buyers and are associated with detailed information about the product reviewed). However, while review deletion seems well targeted, there is a large lag (compared with the Facebook promotion duration) between

these reviews being posted and then deleted. In practice, this means that currently, Amazon is not able to completely eliminate the short-term effects that these reviews have on sellers' outcomes.

Firms are continuously improving and perfecting their platforms' manipulation strategies, and fake reviews continue to be one of the main approaches that firms take to improve economic outcomes. Despite the quite large body of research studying fake reviews, this area is evolving at a very fast pace so that findings that were true only a few years ago, or strategies that could have worked in the past to eliminate fake reviews, might be outdated today. This is why studying and understanding how firms create fake reviews continue to be an extremely important topic of research for both academics and practitioners. As a testament to this, in 2019 alone, Amazon spent over \$500 million and employed over 8,000 people to reduce fraud and abuse on its platform.¹⁴

Our paper is one of the first that tracks and documents the market for fake reviews and its impact on sellers' outcomes. Doing so required the implementation of a sophisticated and extensive data collection that employed both research assistants and scrapers spanning over a year. This effort led to a complete characterization of a new market for fake reviews based on Facebook groups, including its effects on sellers, consumers, and platforms.

¹⁴See: <https://themarkup.org/ask-the-markup/2020/07/21/how-to-spot-fake-amazon-product-reviews>

References

- Ananthakrishnan, U., Li, B., and Smith, M. (2020). A tangled web: Should online review portals display fraudulent reviews? *Information Systems Research*, forthcoming.
- Anderson, E. and Simester, D. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, 51.
- Cabral, L. and Hortacsu, A. (2010). The Dynamics Of Seller Reputation: Evidence From Ebay. *Journal of Industrial Economics*, 58(1):54–78.
- Chiou, L. and Tucker, C. (2018). Fake news and advertising on social media: A study of the anti-vaccination movement.
- Cho, G., Cho, J., Song, Y., and Kim, H. (2015). An empirical study of click fraud in mobile advertising networks. In *2015 10th International Conference on Availability, Reliability and Security*, pages 382–388.
- Crussell, J., Stevens, R., and Chen, H. (2014). Madfraud: Investigating ad fraud in android applications. *MobiSys 2014 - Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*.
- Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management science*, 52(10):1577–1593.
- Einav, L., Farronato, C., and Levin, J. (2016). Peer-to-peer markets. *Annual Review of Economics*, 8(1):615–635.
- Hollenbeck, B. (2018). Online reputation mechanisms and the decreasing value of chain affiliation. *Journal of Marketing Research*, 55(5):636–654.
- Hollenbeck, B., Moorthy, S., and Proserpio, D. (2019). Advertising strategy in the presence of reviews: An empirical analysis. *Marketing Science*, pages 793–811.
- Li, X., Bresnahan, T. F., and Yin, P.-L. (2016). Paying incumbents and customers to enter an industry: Buying downloads.
- Luca, M. and Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427.
- Mayzlin, D., Y., D., and Chevalier, J. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *The American Economic Review*, 104:2421–2455.
- Milgrom, P. and Roberts, J. (1986). Prices and Advertising Signals of Product Quality. *Journal of Political Economy*, 94:297–310.
- Nelson, P. (1970). Information and consumer behavior. *Journal of political economy*, 78(2):311–329.

- Proserpio, D. and Zervas, G. (2016). Online Reputation Management: Estimating the Impact of Management Responses on Consumer Reviews. *Marketing Science*.
- Rao, A. (2018). Deceptive claims using fake news marketing: The impact on consumers.
- Rao, A. and Wang, E. (2017). Demand for “healthy” products: False claims and ftc regulation. *Journal of Marketing Research*, 54.
- Rhodes, A. and Wilson, C. M. (2018). False advertising. *The RAND Journal of Economics*, 49(2):348–369.
- Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8(1):321–340.
- Wu, Y. and Geylani, T. (2020). Regulating deceptive advertising: False claims and skeptical consumers. *Marketing Science*.