



# Leveraging Social Media to Buy Fake Reviews

By Sherry He, Brett Hollenbeck, and Davide Proserpio

## Abstract

**We study the market for fake product reviews on Amazon.com. Reviews are purchased in large private groups on Facebook and other sites. We hand collect data on these markets and then collect a panel of data on these products' ratings and reviews on Amazon, as well as their sales rank, advertising, and pricing policies. We find that a wide array of products purchase fake reviews, including products with many reviews and high average ratings. Buying fake reviews on Facebook is associated with a significant but short-term increase in average rating and number of reviews. We exploit a sharp but temporary policy shift by Amazon to show that rating manipulation has a large causal effect on sales. Finally, we examine whether rating manipulation harms consumers or whether it is mainly used by high-quality in a manner like advertising or by new products trying to solve the cold-start problem. We find that after firms stop buying fake reviews, their average ratings fall and the share of one-star reviews increases significantly, particularly for young products, indicating rating manipulation is mostly used by low-quality products.**

## 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 practices and earning consumer trust. The economics of platforms suggest a difficult trade-off between opening the platform to outside actors such as third-party sellers and retaining strict control over actions taken on the platform. Preventing fraudulent or manipulative actions is key to this trade-off.

One such practice is manipulating reputation systems with fake product reviews. Conventional wisdom holds that fake reviews are particularly harmful because they inject noise and deception into systems designed to alleviate asymmetric information, cause consumers to purchase products that may be of low quality, and erode the long-term trust in the review platforms that is crucial for online markets to flourish.<sup>1,3,13</sup>

We study the economics of rating manipulation and its effect on seller outcomes, consumer welfare, and platform value. Despite being illegal, we document the existence of large and active online markets for fake reviews. Sellers post in private online groups to promote their products and pay customers to purchase them and leave positive reviews. 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

which products are buying fake reviews and when.<sup>a</sup> We then track these products' outcomes on Amazon.com, including their reviews, ratings, prices, and sales rank. This is the first data of this kind, providing direct evidence of the fake reviews themselves and the outcomes of buying fake reviews.

The mere existence of such a large and public market for fake reviews on the largest e-commerce platform presents a puzzle. Given the potential reputation costs, why does Amazon allow this? In the short run, platforms may benefit from allowing fake positive reviews if these reviews increase revenue by generating sales or allowing for higher prices. It is also possible that fraudulent reviews are not misleading on average if high-quality firms are more likely to purchase them than low-quality firms. They could be an efficient way for sellers to solve the "cold-start" problem and establish a good reputation. Indeed, Dellarocas<sup>2</sup> shows that this is a potential equilibrium outcome. In an extension of the signal-jamming literature on how firms can manipulate strategic variables to distort beliefs, he shows that fake reviews are mainly purchased by high-quality sellers and, therefore, increase market information under the condition that demand increases convexly with respect to user rating. Given how ratings influence search results, it is plausible that this condition holds. Other attempts to model fake reviews have also concluded that they may benefit consumers and markets. The mechanism is different, but intuitively this outcome is similar to signaling models of advertising for experience goods. Nelson<sup>11</sup> and later Milgrom and Roberts<sup>10</sup> shows that separating equilibria exist 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 positive word-of-mouth once consumers have discovered their true quality. If fake reviews generate sales which, in turn, generate future organic ratings, a similar dynamic could play out. In this case, fake reviews may be seen as harmless substitutes for advertising rather than as malicious. Therefore, we are left with an empirical question as to whether or not to view rating manipulation as representing a significant threat to consumer welfare and platform reputations.

Our research objective is to answer a set of currently unsettled questions about online rating manipulation. How does this market work, in particular, what are the costs

<sup>a</sup> While technically the seller buys the fake reviews, not the product, because our analysis is done at the product level and sellers often have many products, for clarity we refer to products buying fake reviews.

The original version of this paper is entitled "The Market of Fake Reviews" and was published in *Proceedings of the 22<sup>nd</sup> ACM Conf. on Economics and Computation*, June 2021.

and benefits to sellers from buying fake reviews? What types of products buy fake reviews? How effective are they at increasing sales? Does rating manipulation ultimately harm consumers or are they mainly used by high-quality products? That is, should they be seen more like advertising or outright fraud? Do fake reviews lead to a self-sustaining increase in sales and organic ratings? These questions can be directly answered using the unique nature of our data.

## 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 rating manipulation and its effect on seller outcomes, consumer welfare, and platform value. 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 collected product information such as reviews, ratings, and sales rank data.

### 2.1. Facebook groups and data

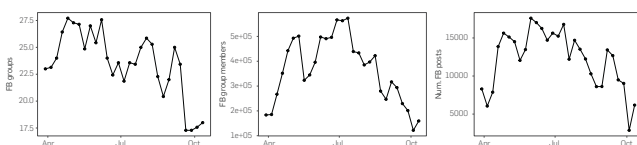
Facebook is one of the major platforms that Amazon sellers use to recruit fake reviewers. To do so, sellers create private Facebook groups where they promote their products by soliciting users to purchase their products and leave a five-star review in exchange for a full refund (and in some cases an additional payment). Discovering these groups is straightforward by searching for “Amazon Review.” We begin by documenting the nature of these groups and then describe how we collect product information from them.

**Discovering groups.** We collected detailed data on the extent of Facebook group activity from March 28, 2020 to Oct 11, 2020. Each day, we collected the Facebook group statistics for the top 30 groups by search rank. 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 568 fake review requests posted per day per group. We observe that Facebook periodically deletes these groups but that they quickly reemerge. Figure 1 shows the weekly average number of active groups, number of members, and number of posts between April and October 2020.<sup>b</sup>

Within these Facebook groups, sellers can obtain a five-star review that looks organic. Usually, these posts contain words such as “need reviews,” “refund after pp [PayPal]”

<sup>b</sup> The total number of members and posts likely overstates the true amount of activity due to double-counting the same sellers and reviewers across groups.

**Figure 1. Weekly average number of FB groups, members, and seller posts.**



with product pictures. The reviewer and seller then communicate via Facebook private messages. To avoid being detected by Amazon’s algorithm, sellers do not directly give reviewers the product link; instead, sellers ask reviewers to search for specific keywords associated with the product and then find it using the title of the product, the product photo, or a combination of the two.

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. The fact that the reviewer buys the product means that the Amazon review is listed as a “Verified Purchase” review and reviewers are encouraged to leave lengthy, detailed reviews that include photos and videos to mimic authentic and organic reviews.<sup>c</sup> Finally, sellers recruit only reviewers located in the United States, with an Amazon.com account, and with a history of past reviews.

**Discovering products.** We use a group of research assistants to discover products that are promoted. Facebook displays the posts in a group in an order determined by some algorithm that factors in when the post was made as well as engagement with the post via likes and comments. Likes and comments for these posts are relatively rare and so the order is primarily chronological. We directed our research assistants to randomize which products were selected by scrolling through the groups and selecting products in a quasi-random way while explicitly ignoring the product type/category, amount of engagement with the post, or the text accompanying the product photo.

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. After a research assistant successfully identifies the product, we ask them to document the 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 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

<sup>c</sup> The fact that these fake reviews are from verified purchases indicates that an identification strategy like the one used in Mayzlin and Chevalier<sup>9</sup> will not work in settings like these.

the seller’s Facebook name and 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 random 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 1500 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

After identifying products whose ratings are manipulated, we collect data for these products on Amazon.com.

**Search results data.** For each product buying fake reviews, we repeatedly collect all information from the keyword search page results, that is, the list of products returned as a result of a keyword search query. This set of products is useful to form a competitor set for each focal product. We collect this information daily, including price, coupon, displayed rating, number of reviews, search page number, whether the product buys sponsored listings, and the product position in each page.

**Review data.** We collect the reviews and ratings for each of the products on a daily basis. For each review, we observe rating, product ID, review text, presence of photos, and helpful votes.

Additionally, twice per month we collect the full set of reviews for each product. The reason for this is that it allows us to measure to what extent Amazon responds by deleting reviews that it deems as potentially fake.

In addition to collecting this data for the focal products, we collect daily and twice-monthly review data for a set of 2714 competitor products to serve as a comparison set. To do so, for each focal product we select the two competitor products that show up most frequently on the same search page as the focal product in the seven days before and seven days after their first FB post. The rationale is that we want to create a comparison set of products that are in the same subcategory as the focal products and have a similar search rank. We collect these products’ review data from 14 August 2020 to 22 January 2021.

**Sales rank data.** We rely on Keepa.com and its API to collect sales rank data twice a week for all products. 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 an ordinal ranking of products.

2.3. Descriptive statistics

Here, we provide descriptive statistics on the set of roughly 1500 products collected between October 2019 to June 2020. We use this sample of products to characterize the types of products that sellers promote with fake reviews. On the one hand, we might expect these products to be primarily products that are new to Amazon.com with few or no reviews whose sellers are trying to jump-start sales by establishing a good 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 shows a breakdown of the top 15 categories and subcategories in our sample. Fake reviews are 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, and the most represented product in our sample, Humidifiers, only accounts for roughly 1% of products. Nearly all products are sold by third-party sellers.

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

The focal products are significantly younger than competitor products, with a median age of roughly five months compared with 15 months for products not observed buying fake reviews. But with a mean age of 229 days, the products collecting fake reviews are not generally new to Amazon and without any reputation. Indeed, out of the 1500 products we observe, only 94 solicit fake reviews in their first month.

Focal products charge slightly lower average prices than their competitors, having a mean price of \$33 (compared with \$45 for the comparison 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.

Turning to ratings, we observe that products purchasing fake reviews have, at the time of their first Facebook 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. Only 14% of focal products have ratings below four stars, compared with 19.5% for competitor products. Thus, it appears that products purchasing fake reviews do not seem to do so because they have a bad reputation. Although, we note that ratings may of course be influenced by previous unobserved Facebook campaigns.

We also examine the number of reviews. The mean number of reviews for focal products is 183, which is driven by

Table 1. Focal product top 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

a long right tail of products with more than 1000 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 with the set of competitor products, which has a median of 59 reviews and a mean of 451 reviews. Despite these differences, it seems that only a small share of the focal products have very few or no reviews. We also 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 still be buying fake reviews via other (more private) channels, or they may 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 on small independent firms relative to firms with well-known brands.<sup>6</sup>

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. DESCRIPTIVE RESULTS ON PRODUCT OUTCOMES AFTER BUYING FAKE REVIEWS

In this section, we quantify the extent to which buying fake reviews is associated with changes in average ratings, number of reviews, and sales rank, as well as other marketing activities such as advertising and promotions. To do so we take advantage of a unique feature of our data in that it contains a detailed panel on firm outcomes observed both before and after sellers buy fake reviews. We stress that, in this section, the 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. We present results on the causal effects of fake reviews on sales outcomes in the full version of this paper.<sup>5</sup>

We first present results in the short term, that is, immediately after sellers begin buying fake reviews for their listings. We then show results for the persistence of these effects after the recruitment period has ended. Finally, we show descriptive results on the extent to which Amazon responds to this practice by deleting reviews.

#### 3.1. Short-term outcomes after buying fake reviews

We begin by quantifying the extent to which buying fake reviews is associated with changes in average ratings, reviews, and sales rank in the short term. To evaluate these outcomes, we partition the time around the earliest Facebook recruiting post date (day 0) in 7-day intervals.<sup>d</sup> We then plot outcomes for eight 7-day intervals before and four 7-day intervals after the first fake review recruitment post.

<sup>d</sup> 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).

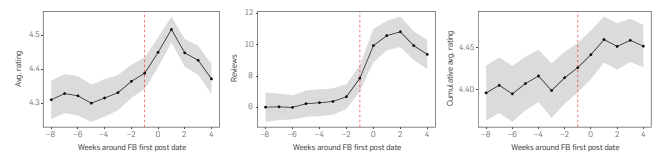
**Ratings and reviews.** We first examine ratings and reviews. In the left panel of Figure 2, we plot the weekly average rating after rating manipulation begins. We see that, first, the average ratings increase by about 5%, from 4.3 stars to 4.5 stars at its peak. Second, this increase in rating is short-lived, and it starts dissipating just two weeks after the beginning of the fake review recruiting; despite this, even after four weeks after the beginning of the promotion, average ratings are still slightly higher than ratings in the pre-promotion period. Third, the average star rating increases slightly 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 beginning date of the fake review recruitment fairly well because the largest change in outcome is visible after or on interval zero.

Next, we turn to the number of reviews. In the middle panel of Figure 2, we plot the weekly average number of posted reviews. We observe that the number of reviews increases substantially around interval zero, nearly doubling, providing suggestive evidence that recruiting fake reviewers is effective at generating new product reviews at a fast pace. Moreover, and differently from the average rating plot, the increase in the weekly number of reviews persists for more than a month. This increase in the number of reviews likely reflects both the fake reviews themselves and additional organic reviews that follow naturally from the increase in sales we document below. Finally, Figure 2 confirms that we are not able to capture the exact date on which the Facebook promotion started.

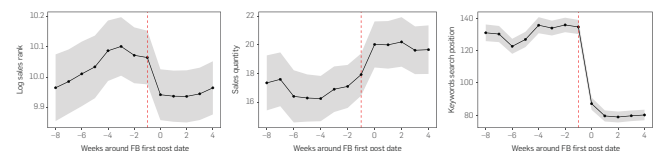
Does the increase in reviews lead to higher displayed ratings? To answer this question, in the right panel of Figure 2, we plot the cumulative average rating before and after the Facebook promotion starts. We observe that ratings increase and then stabilize for about two weeks, after which the increase starts to dissipate.

**Sales rank.** In the left panel of Figure 3, we plot the average

**Figure 2. 7-day average ratings (left), number of reviews (center), and cumulative average ratings (right) before and after fake reviews recruiting begins. The red dashed line indicates the last week of data before we observe Facebook fake review recruiting.**



**Figure 3. 7-day average sales rank (left), sales in units (center), and keyword search position (right) before and after fake reviews recruiting begins. The red dashed line indicates the last week of data before we observe Facebook fake review recruiting.**





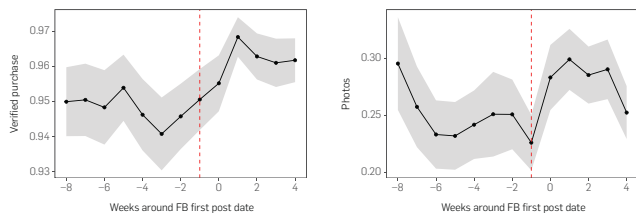
log of sales rank. The figure shows that the sales rank of these products increases between the intervals  $-8$  and  $-3$ , meaning that rating manipulation typically follows a period when sales are falling. When the recruiting period begins, we observe a large increase in weekly sales (i.e., sales rank falls). This increase is likely reflecting both the initial product purchases by the reviewers paid to leave fake reviews as well as the subsequent increase in organic sales that follow. The increase in sales lasts for at least several weeks.

The center panel of Figure 3 plots sales in units sold. Amazon does not display this metric but it is possible to measure sales in units for a subset of products and then estimate the relationship between rank and units.<sup>4</sup> We plot the observed sales and point estimates of estimated sales around the time of the first Facebook post and see a sharp increase in average units sold, from around 16 units per week to roughly 20.

**Keyword search position.** So far we have shown that recruiting fake reviews is associated with improvements in ratings, reviews, and sales. One reason for observing higher sales may be that higher ratings signal higher quality to consumers, who then are more likely to buy the product. A second reason is that products recruiting fake reviews will be ranked higher in the Amazon search results due to them having higher ratings and more reviews. To investigate whether this is the case, in the right panel of Figure 3, we plot the search position rank of products recruiting fake reviews. 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 change seems to be long-lasting as the position remains virtually constant for several weeks.

**Verified purchases and photos.** An important aspect of the market for fake reviews is that reviewers actually buy the product and can therefore be listed as verified reviewers. In addition, they are compensated for creating realistic reviews, that is, they are encouraged to post long and detailed reviews including photos and videos. In the left panel of Figure 4, we show changes in the average share of 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. In

**Figure 4. 7-day average verified purchase (left) and the number of photos (right) before and after fake reviews recruiting begins. The red dashed line indicates the last week of data before we observe Facebook fake review recruiting.**



the right panel, we observe a sharp increase in the share of reviews containing photos.

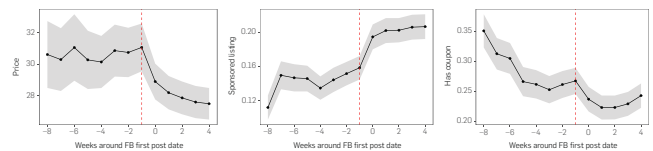
**Marketing activities.** Finally, we investigate to what extent rating manipulation is associated with changes in other marketing activities such as promotions (rebates, sponsored listings, and coupons). We plot these quantities in Figure 5. We observe a substantial drop in prices (left panel) that persists for several weeks and an increase in the use of sponsored listings, suggesting that Amazon sellers complement the Facebook promotion with advertising activities. This result is in contrast with<sup>7</sup> who find 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 (albeit noisy) change in the use of coupons.

### 3.2. Long-term outcomes after buying fake reviews

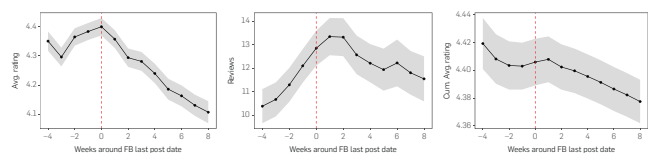
In this subsection, we describe what happens after sellers stop buying fake reviews. We are particularly interested in using the long-term outcomes to assess whether rating manipulation generates a self-sustaining increase in sales or organic reviews. If we observe that these products continue to receive high organic ratings and have high sales after they stop recruiting fake reviews, we might conclude that fake reviews are a potentially helpful way to solve the cold-start problem of selling online with a limited reputation.

We, therefore, track the long-term trends for ratings, reviews, and sales rank. Similar to Section 3.1, we partition the time around the last Facebook recruiting post date in 7-day intervals, and plot the outcomes for four weeks before fake reviews recruiting stops (thus covering most of the period where products recruited fake reviews) and eight weeks after fake reviews recruiting starts. Doing so, we compare the Facebook promotion period (negative intervals) with the post-promotion period (positive intervals).

**Figure 5. 7-day average prices (left), sponsored listings (center), and coupon (right) before and after fake reviews recruiting begins. The red dashed line indicates the last week of data before we observe Facebook fake review recruiting.**



**Figure 6. 7-day average number of average ratings (left), number of reviews (center), and cumulative average ratings (left) before and after fake reviews recruiting stops. The red dashed line indicates the last week of data in which we observe Facebook fake review recruiting.**



**Ratings and reviews.** Long-term trends in ratings and reviews are shown in Figure 6. We observe that the increase that occurs when sellers buy fake reviews is fairly short. After one to two weeks from the end of the Facebook promotion, both the weekly average rating and the number of reviews (left and middle panel, respectively) start to decrease substantially. The cumulative average rating (right panel) drops as well. Interestingly, these products end up having average ratings that are significantly worse than when they began recruiting fake reviews (approximately interval  $-4$ ).

**Sales rank.** The left panel of Figure 7 shows the long-term trend in the average log sales rank. It shows that sales decline substantially after the last observed Facebook post. This suggests that the increase associated with recruiting fake reviews is not long-lasting as it does not lead to a self-sustaining set of sales and positive reviews.

The middle panel of Figure 7 shows sales in units. The result is consistent with sales rank, showing that sales peak during the week of the last Facebook post and subsequently decline.

**Keyword search position.** The right panel of Figure 7 shows the long-term trend in average keyword search position. We observe that after the Facebook campaign stops, the downward trend in search position stops but does not substantially reverse even after two months. Therefore, products enjoy a better ranking in keyword searches for a relatively long period after fake review recruiting stops.

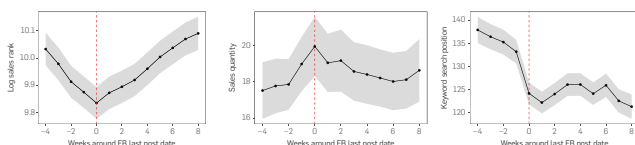
The relatively stable and persistent increase in search position suggests that this measure may have a high degree of inertia. After an increase in sales and ratings causes a product's keyword rank to improve, it does not decline quickly, even when sales are decreasing. This also suggests that the decrease in sales shown in Figure 7 does not come from reduced product visibility but from the lower ratings and increase in one-star reviews. Finally, while we demonstrate in the next section that Amazon deletes a large share of reviews from products that recruit fake reviews, the inertia in keyword rank suggests that Amazon does not punish these sellers using the algorithm that determines organic keyword rank. This could therefore serve as an additional policy lever for the platform to regulate fake reviews.

### 3.3. Amazon's response

In this subsection, we provide evidence of the extent to which Amazon is aware of the fake review problem and what steps it is taking to remove these reviews.

While we cannot observe reviews that are filtered by Amazon's fraud detection practices and never made public,

**Figure 7. 7-day average sales rank (left), sales in units (center), and keyword rank (right) before and after fake review recruiting stops. The red dashed line indicates the last week of data in which we observe Facebook fake review recruiting.**



by collecting review data on a daily and twice-monthly basis, we can observe if reviews are posted and then later deleted. We calculate the share of reviews that are deleted by comparing the full set of observed reviews from our daily scraper with the set of reviews that remain posted at the end of our data collection window. We find that for the set of products observed recruiting fake reviews, the average share of posted reviews that are ultimately deleted is about 43%, compared to 23% for products not observed recruiting fake reviews. This suggests that, to some extent, Amazon can identify fake reviews.

To further characterize Amazon's current policy, we next analyze the characteristics of deleted reviews and the timing of review deletion.

**Characteristics of deleted reviews.** In Table 2, we report the mean and standard deviation for several review characteristics for deleted and non-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," that 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.

We find that deleted reviews have higher average ratings than non-deleted reviews. 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 reviews are more likely to be deleted suggests that Amazon is fairly sophisticated in targeting potentially fake reviews.<sup>e</sup> Finally, we find no difference for whether the review is associated with a verified purchase or tagged as "Amazon Earlier Reviews."<sup>f</sup>

**When are reviews deleted?** Finally, we analyze when

<sup>e</sup> This result contrasts with Luca and Zervas,<sup>8</sup> who find that longer reviews are less likely to be filtered as fake by Yelp.

<sup>f</sup> We find that Amazon does not delete any reviews tagged as "Amazon Earlier Reviews" potentially because Amazon's process to identify and select early reviewers drastically reduces the possibility of these reviews being fake.

**Table 2. Comparing deleted and non-deleted reviews characteristics.**

	Deleted reviews	Non-deleted reviews
Verified purchase	0.98 (0.16)	0.96 (0.20)
Review rating	4.65 (0.98)	4.24 (1.37)
Number of photos	0.35 (0.93)	0.19 (0.72)
Early reviewer	0.00 (0.00)	0.01 (0.11)
Title length	9.81 (13.94)	21.08 (13.80)
Review length	236.73 (222.88)	198.75 (231.68)

Note: Standard deviations in parentheses.

Amazon deletes 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, that is, the beginning of the buying of fake reviews. To do so, we partition the time in days around the first Facebook post and then plot the number of products for which reviews are deleted. Because products recruit fake reviews for different time periods, we perform this analysis by segmenting products based on the quartiles of campaign duration. Figure 8 shows the results of this analysis.

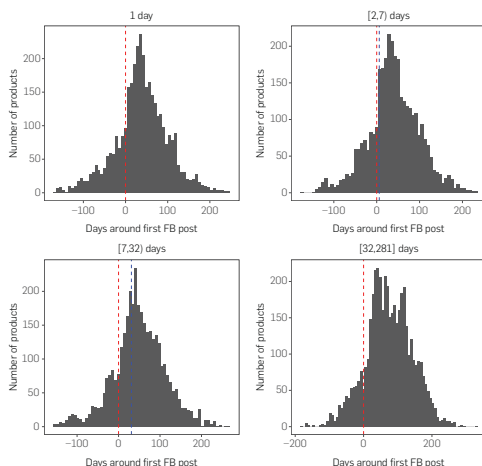
What emerges from this figure is that Amazon starts deleting reviews for more products after the Facebook campaign begins (red-dashed line) and often it does so only after the campaign terminated (blue-dashed line). Indeed, it seems that most of the review deletion happens during the period covering the two months after the first Facebook post date, but most campaigns are shorter than a month. A simple calculation suggests that reviews are deleted only after a quite large lag. The mean 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 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 3.1. In the next section, we show that there is one time period in our data during which Amazon's deletion policy changes significantly; we use this period to identify the causal effects of fake reviews on sales.

#### 4. EVIDENCE OF CONSUMER HARM FROM FAKE REVIEWS

We conclude the paper by evaluating whether consumers are harmed by fake reviews. To do so, we analyze the products' ratings after they stop buying fake reviews. If they continue receiving high ratings after rating manipulation ends it would

**Figure 8. 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.**



be evidence that fake reviews are used by high-quality products in a manner akin to advertising. This would be consistent with the predictions of theoretical results in Dellarocas<sup>2</sup> and others. If, by contrast, we see declining ratings and observe a large number of one-star reviews, it would suggest fake reviews are bought to mask low product quality and deceive consumers.

There is an inherent limitation in using ratings to infer welfare because consumers leave ratings for many reasons and generally ratings are not a literal expression of utility. But we argue that when products receive low ratings and a large number of one-star reviews, it indicates that the actual quality of these products is lower than what most customers expected at the time of their purchase. The low ratings are either a direct expression of product quality or an attempt to realign the average rating back toward the true level and away from the manipulated level. In this latter case, we still infer consumer harm, either because it indicates consumers paid a higher price than what they would have if the product was not overrated due to rating manipulation, or because the fake reviews caused them to buy a lower quality product than the closest alternative. This analysis is also important from the platform's perspective. An increase in one-star reviews would indicate that fake reviews are a significant problem since they reflect negative consumer experiences that erode trust in the platform's reputation system.<sup>8</sup>

##### 4.1. One-star ratings and reviews

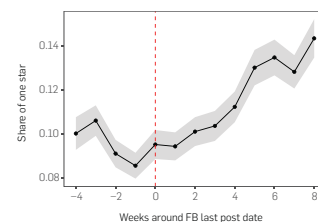
We previously showed in Figure 6 in Section 3.2 that average ratings fall after fake review recruiting ends. Figure 9 shows why. The share of one-star reviews increases by approximately 70% after fake review recruiting stops. The increase in the share of one-star ratings and the increase in the total number of ratings means that the absolute number of one-star reviews increases by even more.

Next, we explore how this pattern varies for different types of products. It may be the case that ratings stay high for certain products. 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, which they can sustain if these products are high quality.

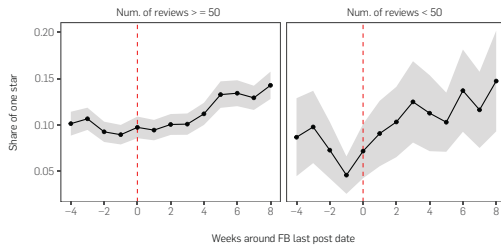
To test this, we segment products by the number of reviews and age. Figure 10 shows how the share of one-star reviews changes for products with fewer than 50 reviews.

<sup>8</sup> Nosko and Tadelis<sup>12</sup> shows that when a buyer has a bad product experience with a third-party seller on a platform, they are significantly less likely to shop at that platform again.

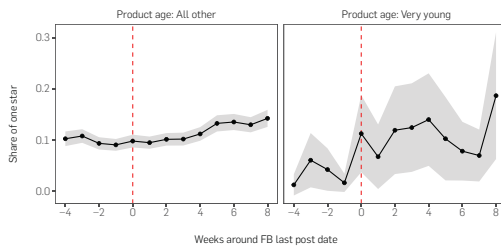
**Figure 9. 7-day average share of one-star reviews before and after fake reviews recruiting stops. The red dashed line indicates the last time we observe Facebook fake review recruiting.**



**Figure 10. 7-day average share of one-star reviews before and after fake reviews recruiting stops by the number of reviews accumulated prior to the fake review recruiting. The red dashed line indicates the last time we observe Facebook fake review recruiting.**



**Figure 11. 7-day average share of one-star reviews before and after fake reviews recruiting stops by product age (very young products are those listed for fewer than 60 days). The red dashed line indicates the last time we observe Facebook fake review recruiting.**



The increase in one-star ratings is sharper for products with few reviews. Figure 11 makes the same comparison for products that have been listed on Amazon for fewer than 60 days. The young products experience a much larger increase in one-star reviews than the other products, with more than 20% of their ratings being one-star two months after they stop recruiting fake reviews. Overall, these results refute the idea that “cold-start” products use fake reviews efficiently. Instead, these products seem to be of especially low quality.

## 5. DISCUSSION AND CONCLUSION

It has become commonplace for online sellers to manipulate their reputations on online platforms. In this paper, we study the market for fake Amazon product reviews, which take place in private Facebook groups featuring millions of products. We find that soliciting reviews on Facebook is highly effective at improving several sellers’ outcomes, such as the number of reviews, ratings, search position rank, and sales rank. However, these effects are often short-lived as many of these outcomes return to pre-promotion levels a few weeks after the fake reviews recruiting stops. In the long run, this boost in sales does not lead to a positive self-sustaining relationship between organic ratings and sales, and both sales and average ratings fall significantly once fake review recruiting ends. Rating manipulation is not used efficiently by sellers to solve a cold-start problem, in other words.

We also find evidence that this practice is likely harmful to consumers, as fake review recruiters ultimately see a large decrease in ratings and an increase in their share of one-star reviews. An important implication is that rating manipulation is also likely to harm honest sellers and the platform’s

reputation itself. If large numbers of low-quality sellers are using fake reviews, the signal value of high ratings could decrease, making consumers more skeptical of new, highly rated products. This, in turn, would make it more difficult for high-quality sellers to enter the market and would likely reduce innovation.

Firms are continuously improving and perfecting their manipulation strategies 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 manipulate their ratings continue to be an extremely important topic of research for both academics and practitioners.

We also document that Amazon does delete large numbers of reviews and that these deletions are well targeted, but there is a large lag before these reviews are deleted. The result is that this deletion policy does not eliminate the short-term profits from these reviews or the consumer harm they cause.

Of course, Amazon has other potential policy levers at its disposal to regulate fake reviews. But we do not observe Amazon deleting products or banning sellers as a result of them manipulating their ratings. Nor do we observe punishment in the products’ organic ranking in keyword searches. This keyword ranking stays elevated several months after fake review recruiting has ended, even when Amazon finds and deletes many of the fake reviews posted on the platform. Reducing product visibility in keyword rankings at the time fake reviews are deleted could potentially turn fake reviews from a profitable endeavor into a highly unprofitable one.

It is not obvious whether Amazon is simply under-regulating rating manipulation in a way that allows this market to continue to exist at such a large scale, or if it is assessing the short-term profits that come from the boost in ratings and sales and weighing these against the long-term harm to the platform’s reputation. Quantifying these two forces is, therefore, an important area of future research.

## References

- Cabral, L., Hortacsu, A. The dynamics of seller reputation: Evidence from eBay. *J. Indus. Econ.* 58, 1 (2010), 54–78.
- Dellarocas, C. Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Manage. Sci.* 52, 10 (2006), 1577–1593.
- Einav, L., Farronato, C., Levin, J. Peer-to-peer markets. *Annu. Rev. Econ.* 8, 1 (2016), 615–635.
- He, S., Hollenbeck, B. Sales and rank on amazon.com. Available from: SSRN 3728281. 2020.
- He, S., Hollenbeck, B., Proserpio, D. The market for fake reviews. *Market. Sci.* 41, 5 (2022), 896–921.
- Hollenbeck, B. Online reputation mechanisms and the decreasing value of chain affiliation. *J. Market. Res.* 55, 5 (2018), 636–654.
- Hollenbeck, B., Moorthy, S., Proserpio, D. Advertising strategy in the presence of reviews: An empirical analysis. *Market. Sci.* 38, 5 (2019), 793–811.
- Luca, M., Zervas, G. Fake it till you make it: Reputation, competition, and yelp review fraud. *Manage. Sci.* 62, 12 (2016), 3412–3427.
- Mayzlin, D., Dover, Y., Chevalier, J. Promotional reviews: An empirical investigation of online review manipulation. *Am. Econ. Rev.* 104 (2014), 2421–2455.
- Milgrom, P., Roberts, J. Prices and advertising signals of product quality. *J. Political Econ.* 94 (1986), 297–310.
- Nelson, P. Information and consumer behavior. *J. Political Econ.* 78, 2 (1970), 311–329.
- Nosko, C., Tadelis, S. The limits of reputation in platform markets: An empirical analysis and field experiment. *NBER Working Paper 20830* (2015).
- Tadelis, S. Reputation and feedback systems in online platform markets. *Annu. Rev. Econ.* 8, 1 (2016), 321–340.

**Sherry He and Brett Hollenbeck** ([sherry.he.phd, brett.hollenbeck]@anderson.ucla.edu), Anderson School of Management, University of California Los Angeles, Los Angeles, CA, USA.

**Daide Proserpio** (daide.proserpio@marshall.usc.edu) Marshall School of Business, University of Southern California, Los Angeles, CA, USA.