

ORIGINAL ARTICLE

Optional purchase verification in e-commerce platforms: More representative product ratings and higher quality reviews

Marios Kokkodis¹  | Theodoros Lappas²  | Gerald C. Kane¹ 

¹Carroll School of Management, Boston College, Chestnut Hill, Massachusetts, USA

²School of Business Administration, Athens University of Economics and Business, Athens, Greece

Correspondence

Marios Kokkodis, Carroll School of Management, Boston College, Chestnut Hill, MA 02467, USA.
Email: kokkodis@bc.edu

Handling Editor: Subodha Kumar

Abstract

E-commerce platforms struggle to create and maintain high-quality reputation systems. One promising option is “purchase verification,” which confirms that the user reviewing a product purchased the product from the platform. Previous works comparing platforms that require purchase verification with platforms that do not offer purchase verification found that review manipulation is easier in the latter. But what happens in platforms where purchase verification is optional? In such platforms, there is no monetary cost for posting fake reviews. Yet, optional purchase verification (OPV) might introduce indirect costs for fake reviewers through expectation disconfirmation, hence positively affecting the reputation ecosystem of an e-commerce platform. To investigate, we use a quasi-experimental setup to analyze 336,043 book reviews. We find empirical evidence that introducing OPV reduces fake reviews, most of which are positive. This reduction of fake reviews results in lower, more representative product ratings and longer and more helpful reviews posted by more experienced reviewers. These new findings extend our understanding of how OPV can improve a platform’s reputation ecosystem and suggest managerial interventions for platforms that have yet to develop a verification mechanism.

KEYWORDS

fake reviews, online reputation ecosystems, optional purchase verification, quasi-experimental setup

1 | INTRODUCTION

The rapid growth of e-commerce marketplaces (e.g., Amazon) has triggered the offering of a large and diversified set of products (Scrapehero, 2018; Statista, 2018). To increase trust (Awad & Ragowsky, 2008; Pavlou, 2003; Pavlou & Gefen, 2004) and reduce search cost (Bakos, 2001) and information asymmetry (Akerlof, 1978), e-commerce marketplaces have developed various reputation mechanisms based on user-generated content. These mechanisms, which include product ratings and reviews, significantly drive product sales (Archak et al., 2011; Bolton et al., 2004; Chen & Lurie, 2013; Chevalier & Mayzlin, 2006; Chintagunta et al., 2010; Clemons et al., 2006; Cui et al., 2018; Dellarocas, 2003; Dellarocas et al., 2007; Duan et al., 2008a, 2008b; Gu et al., 2012; Han et al., 2018; Hu et al., 2008; Kokkodis & Ipeirotis, 2016; Lau et al., 2018; Li et al., 2011; Reinstein & Snyder, 2005; Zhao et al., 2013). In fact, specific review

characteristics have a stronger effect on sales than actual product information (Liu & Karahanna, 2017), while even negative reviews can increase sales through product awareness (Berger et al., 2010).

Given their established economic value (Wu et al., 2015), it is not a surprise that sellers often submit fake reviews to benefit their products or hurt their competitors. Previous research has found both fraudulent positive and negative reviews in the hotel (Mayzlin et al., 2014) and restaurant (Luca & Zervas, 2016) industries, estimating that as many as 16% of all reviews submitted are fake.

It is in the best economic interest of the platform to ensure that reviews are not only truthful but also high quality. To this end, platforms have developed several mechanisms to reduce fake and low-quality reviews that might distort an item’s reputation and mislead potential buyers (Lappas et al., 2016; Luca & Zervas, 2016; Zhang et al., 2016). One such mechanism is purchase verification, which verifies that the user submitting the review has actually purchased the product on the platform, and, as a result, allows shoppers to identify reviews

Accepted by Subodha Kumar, after two revisions.

that are likely to be truthful (Kokkodis, 2012). Many leading e-commerce platforms, including Amazon, Wayfair, and Target, now use purchase verification.

Despite this widespread use, the little research that has examined the effects of purchase verification on a platform's reputation ecosystem has focused primarily on situations where purchase verification is a prerequisite for reviewing (Mayzlin et al., 2014). However, for many platforms, verifying every purchase may be costly, noisy, or foreign to their business operations. Furthermore, requiring verification could hinder a platform's ability to quickly accumulate a critical mass of opinions and hence result in significantly lower review volume. After all, researchers have identified many legitimate reasons for users to review a product on a different platform than the one they used to purchase (Bateman et al., 2011; Ren et al., 2012; Wasko et al., 2005), and such nonverifiable truthful reviews can subsequently benefit both the reviewed products and the platform. In addition, since most users do not review (Hu et al., 2017), restricting truthful reviewers from participating by requiring purchase verification will result in the loss of valuable but already scarce sources of information.

These observations explain why many platforms (e.g., Amazon, Target) offer *optional purchase verification* (OPV), allowing both users who bought a product on the platform and those who did not to post a review. In these platforms, whether OPV decreases review manipulation remains an open question: Absent a requirement to purchase, fraudulent reviewers can still post a review without an obvious cost. To investigate, we ask: *How does OPV affect the reputation ecosystem of an e-commerce platform?*

We argue that OPV will likely reduce fake reviews because of indirect costs related to the resulting disconfirmation between expected and realized product quality. Because e-commerce platforms offer products with no immediate substitutes (e.g., books), fake reviews are more likely to be overly positive (Luca & Zervas, 2016). Hence, purchase verification will likely result in a reduction of net-positive fake reviews. This reduction will subsequently yield relatively lower, more representative product ratings, and longer, more descriptive reviews posted by more experienced reviewers. Even further, this reduction of fake reviews will likely increase attention to longer truthful reviews, which will increase the average review helpfulness.

To test this theoretical framework we examine 336,043 book reviews from Amazon U.S. and Amazon U.K. The introduction of OPV at different times in the U.S. and U.K. markets forms a quasi-experimental setup that helps isolate the hypothesized effects. In line with our theory, we find that introducing purchase verification results in lower product ratings and reviews that are longer and are posted by more experienced reviewers. Empirical evidence further shows that OPV increases the overall review quality and review helpfulness. Extensive supportive analysis that includes (1) detection of reviews that Amazon has removed as fake, (2) deep learning models that identify fake reviews based on text, (3) models that estimate text authenticity, (4) comparison of review

characteristics between fake and nonfake reviews, (5) evidence that positive fake reviews are followed by relatively negative verified reviews, and (6) analysis of the effect of OPV on the overall review volume provide substantial empirical support of the theorized mechanisms of action. Finally, an extended set of additional tests empirically eliminate alternative explanations that could be driving our observed results.

Our study extends the operations management literature on response bias and fake reviews (Chen et al., 2016; Kumar et al., 2018b, 2019; Lappas et al., 2016; Lee et al., 2018) by explaining how OPV affects both ratings and review characteristics in e-commerce platforms. By showing for the first time that introducing OPV has positive spillover effects on the overall reputation ecosystem of a platform, our findings can guide the business operations of multiple e-commerce and reputation platforms. Platforms that are already offering OPV can now better understand and measure their utility. Platforms that are currently operating without purchase verification (e.g., TripAdvisor, Yelp) have now encouraging empirical evidence that introducing OPV can have multidimensional positive effects on their reputation ecosystem.

2 | RESEARCH CONTEXT AND HYPOTHESES DEVELOPMENT

In the following paragraphs, we provide a brief background of the relevant work in user-generated content in digital and e-commerce platforms, and we hypothesize how the introduction of OPV can change review and reviewer characteristics.

2.1 | Background

Recent studies in the interface of Operations Management–Information Systems (OM-IS) have answered the call to investigate user-generated content in digital platforms (Kumar et al., 2018c). Some focus on understanding and operationalizing user engagement in platforms that depend on content creation (e.g., Twitter, online communities, online reviews; Kokkodis et al., 2020; Kumar et al., 2018a; Mallipeddi et al., 2021). Others focus on the operational value of such content (Cui et al., 2018), and show that (1) proposing sellers' strategies that are consumer-specific yields better quality reviews (Guan et al., 2020), (2) consumer and product characteristics moderate the effect of reviews on product choice (Ba et al., 2020), (3) scalable methods that use review text can identify and extract service dimensions (Mejia et al., 2021), and (4) building predictive models and performing big data analytics and sentiment analysis on user-generated content result in better sales predictions (Cui et al., 2018; Lau et al., 2018).

A stream of OM-IS research closely related to this work focuses on online review biases and manipulation. Specifically, to correct for reporting bias, Chen et al. (2016) model the generating process of online reviews through an inverse probability weighting scheme. Additional research uses

machine learning to build meta-classifiers that identify deceptive reviews (Kumar et al., 2018b), and unsupervised hierarchical approaches to detect anomalies in reviewer behavior and potentially identify fake reviewers (Kumar et al., 2019). Finally, prior studies have also provided evidence of the existence of fake online reviews through text mining, sentiment, and econometric analyses (Lappas et al., 2016; Lee et al., 2018). Our work directly extends this line of OM-IS research by providing, for the first time, evidence that OPV is an effective mechanism against review manipulation, and by showing that its introduction has multidimensional positive spillover effects in the reputation ecosystem of a platform that yield more representative product ratings and longer and more helpful reviews posted by more experienced reviewers.

2.2 | Hypothesis development

2.2.1 | Combating fake reviews with purchase verification

Firms often have difficulty encouraging people to review products on online platforms (Chen et al., 2016). Some estimates suggest that only 0.1% of people actually review the products they purchase on Amazon (Hu et al., 2017). Because the vast majority of people who purchase a product do not provide reviews, sellers have the opportunity to manipulate the reputation of a product by injecting even modest numbers of fraudulent reviews (Lappas et al., 2016; Lee et al., 2018). Such fake reviews undermine the value of online reputation systems (Feng et al., 2012; Sussin & Thompson, 2012). A considerable amount of review fraud originates from professional review-authoring companies that submit fake reviews on major review platforms in exchange for a fee (Lappas et al., 2016). The prevalence and success of these companies have even prompted government action that resulted in \$350,000 in fines (Schneiderman, 2014). Despite such ongoing efforts, many review-manipulating companies are still active, prolific, and successful (Lappas et al., 2016).

To combat fraud and improve the quality of reviews, platforms have implemented various defense mechanisms. For instance, website moderators manually approve reviews on the TripAdvisor and Yelp platforms (Lappas et al., 2016). Suspicious reviews are placed on hold pending examination and can even be eliminated if the website's proprietary filtering process provides enough evidence. Businesses associated with fake reviews are penalized in the platform's rankings, excluded from press releases and top-10 lists, and may even have a relevant banner placed on their page.

Optional and required purchase verification offer another promising strategy for combating fake reviews and increasing trust in online reputation systems (Yuan et al., 2012). When required, purchase verification significantly reduces the number of fake reviews (Mayzlin et al., 2014). When optional, purchase verification reveals which of the available reviews are written by reviewers who have purchased the product on the focal platform, allowing users to distinguish between ver-

ified and nonverified reviews when determining their veracity (Anderson & Simester, 2014; He et al., 2020; Kim et al., 2018b, 2018a; Kokkodis, 2012).

2.2.2 | Types of fake reviews in e-commerce platforms

To understand how introducing OPV affects fake reviews in e-commerce platforms, we first need to better understand what types of fake reviews these platforms attract. In general, fraudulent reviews can either be (1) negative, intended to decrease the ratings of competitors' products, or (2) positive, intended to inflate a product's reputation (Lappas et al., 2016; Lee et al., 2018; Mayzlin et al., 2014). The type of fraudulent reviews depends considerably on the competitive nature of the product (Lee et al., 2018; Luca & Zervas, 2016) or the market (Mayzlin et al., 2014). Fraudulent negative reviews are likely to be more common when competitor products or services are substitutes, as damaging the reputation of one's competitor product is likely to turn consumers to one's own product. For example, fraudulent negative reviews on hotels are more likely when competitors are geographically proximate and able to serve as a ready substitute. In contrast, fraudulent positive reviews are more likely when smaller hotels are trying to stand out and compete with larger competitors (Mayzlin et al., 2014).

In e-commerce settings where such immediate substitutes are improbable, fraudulent negative reviews are less likely to have the intended effect (Anderson & Simester, 2014; Luca & Zervas, 2016). However, fraudulent positive reviews are still likely to have an effect as products seek to differentiate themselves from and gain attention amidst the volume of other products available. The books domain, which is the focus of this study, is a characteristic example of a setting with a large number of competitive products, in which harming a competitor's reputation is unlikely to affect the sales of the fraudulent reviewer's book significantly. However, boosting one's own reputation can still be beneficial. In fact, a considerable number of authors on Amazon positively review their books to increase sales (Smith, 2004).

2.2.3 | Disconfirmation-induced reduction of fake reviews

Even though requiring purchase verification is a successful defense mechanism against fake reviews (Mayzlin et al., 2014), it is not immediately clear why OPV would discourage fake reviews in an e-commerce platform. When purchase verification is required, it introduces economic disincentives for creating fake reviews (Mayzlin et al., 2014). Since optional verification does not create such monetary disincentives for submitting fake reviews, *why would fake reviewers change their behavior after the introduction of OPV?* In fact, one could argue that, because OPV would naturally decrease the value of nonverified reviews (Anderson & Simester,

2014), introducing OPV could create an incentive for fake reviewers to submit *more* fake reviews to compensate for the likely reduced influence of nonverified reviews on purchase decisions.

Response bias (i.e., who chooses to review a product) may provide some rationale of why introducing OPV could in fact reduce the number of fake reviews. In particular, underreporting bias (Hu et al., 2017) recognizes that the vast majority of people do not review products online. People who choose to review often do so because their realized experience with the bought product did not match the rosy picture characterized by the positively skewed reviews (expectation disconfirmation; Ho et al., 2017; Kokkodis & Lappas, 2020). Introducing OPV can increase the value of such expectation disconfirmation incidents: If sellers choose to artificially inflate a product's reputation by submitting nonverified fake positive reviews, then—due to underreporting bias and the fact that not every buyer posts a review—they will also increase the likelihood of reviewers who experience disconfirmation to post negative reviews. As a result, attempts to inflate a product's reputation positively through fake nonverified reviews might consequently result in verified negative reviews from customers who bought the product, experienced it, and found discrepancies between their experience and the posted reviews. Given that verified reviews are relatively more valuable than nonverified ones (Anderson & Simester, 2014), such negative verified reviews will hurt the product significantly more. Hence, and even though introducing OPV does not induce an extra monetary cost to sellers who post fake reviews, we argue that *OPV will end up reducing the amount of fake reviews* due to this indirect disconfirmation-induced cost.

2.2.4 | OPV effects on the reputation ecosystem

The reduction of fake reviews effectively alters the reputation ecosystem (i.e., product ratings, review text, review helpfulness) of an e-commerce platform. Since most fake reviews in such a platform are positive (Section 2.2.2), the disconfirmation-induced reduction of fake reviews suggests that products that have been artificially boosted through positive fake reviews will experience a drop in their average rating after the introduction of OPV. For instance, a product that receives a constant rate of positive fake reviews per year will experience a rate decrease after the introduction of OPV. Hence, all else being equal, this reduction of positive fake reviews will result in a relative *decrease* in product ratings:

Hypothesis 1. *Introducing OPV in an e-commerce platform will result in lower, more representative product ratings.*

Besides product ratings, the reduction of fake reviews also alters the average textual characteristics of the posted reviews. In particular, fake—and paid—reviews tend to be shorter than truthful reviews (Burtch et al., 2017; Li et al.,

2014). This is likely because truthful reviewers describe their genuine experiences, while fake reviewers need to generate and describe fake experiences (Li et al., 2014). Hence, due to the reduction of fake reviews that tend to be shorter, introducing OPV should all else being equal result in relatively longer product reviews.

Hypothesis 2. *Introducing OPV in an e-commerce platform will result in longer product reviews.*

Nonverbal reviewer characteristics are highly predictive of fake reviews (Zhang et al., 2016). In fact, fake reviews are often posted by reviewers who review less and for shorter periods (lower experience) than truthful reviewers (Mukherjee et al., 2013). As a result, the reduction of such fake reviews due to the introduction of OPV will likely lead to a corpus of reviews posted by relatively more experienced reviewers:

Hypothesis 3. *Introducing OPV in an e-commerce platform will result in reviews written by more experienced reviewers.*

Finally, the reduced rate of fraudulent reviews may actually make it easier for consumers to pay attention to the truthful ones that remain (Eppler & Mengis, 2004; Hansen & Haas, 2001). Given that truthful reviews are often longer than fake ones (Burtch et al., 2017; Li et al., 2014), and since longer reviews tend to be more helpful (Mudambi & Schuff, 2010), introducing OPV should also indirectly increase review helpfulness, as consumers can more easily pay attention to these longer, truthful reviews. Thus, we argue that introducing OPV will raise the overall quality of the reviews by reducing shorter fake reviews, which in turn will increase attention to the remaining longer, genuine, and more helpful reviews:

Hypothesis 4. *Introducing OPV in an e-commerce platform will result in more helpful reviews.*

Figure 1 summarizes the hypothesized mechanisms of the expected effects of introducing OPV in an e-commerce platform, while it also illustrates how alternative mechanisms would result in different effects. Next, we discuss the empirical research setting that facilitates the investigation of these mechanisms and the testing of Hypotheses 1 to 4. (Sections 5.2 and 6.3 discuss alternative mechanisms that might overlap with our primary theory presented in Figure 1.)

3 | RESEARCH SETTING

Before describing our research setting and data set, we formally define the focal problem as follows:

Problem definition: Consider an e-commerce platform that does not offer any type of purchase verification. Our goal is to estimate the average effect of introducing OPV on product ratings and review characteristics.

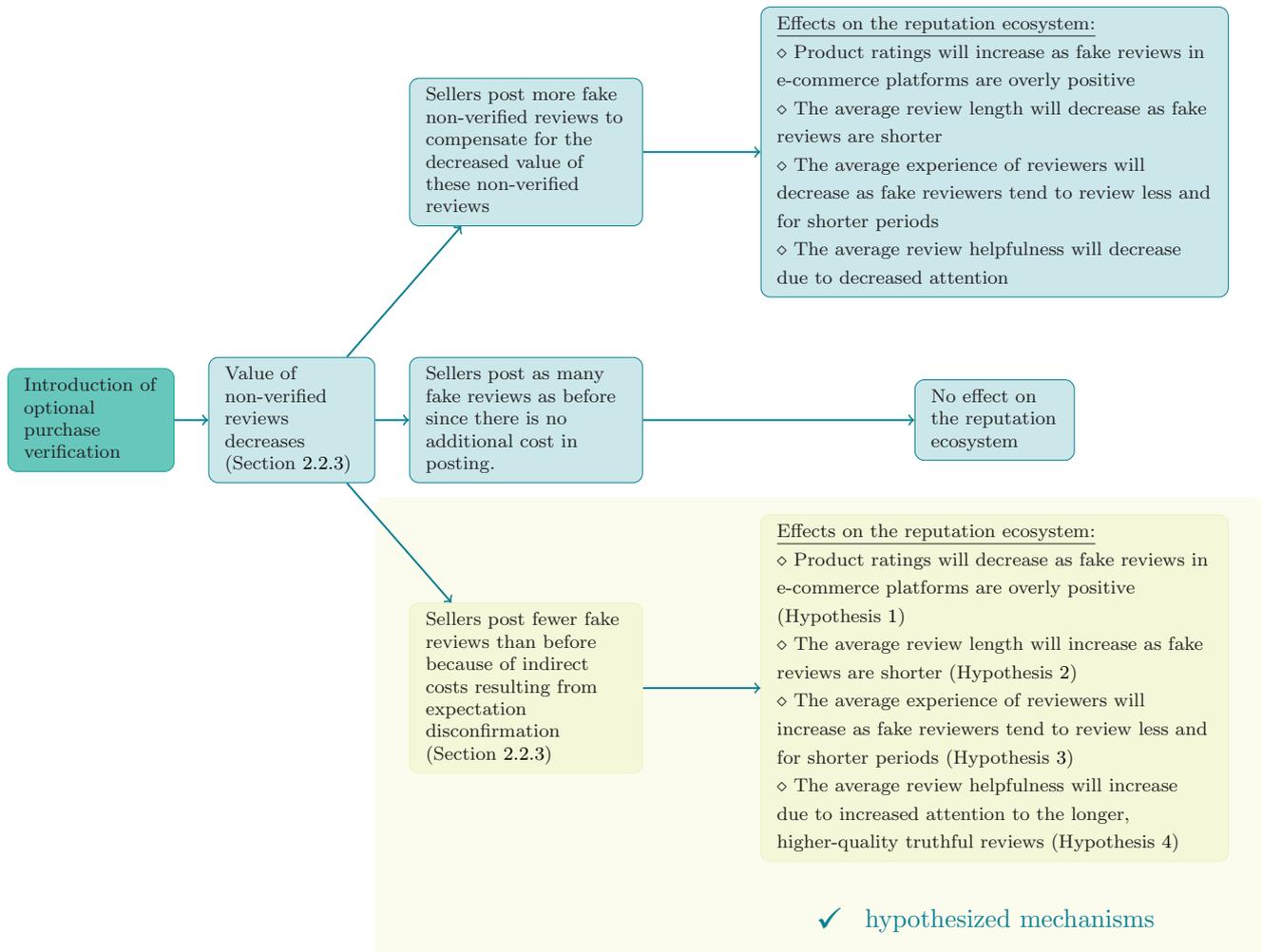


FIGURE 1 Overview of the hypothesized mechanisms

Note: The figure summarizes the hypothesized mechanisms of the expected effects of introducing optional purchase verification. It also shows how alternative underlying mechanisms would have resulted in different effects [Color figure can be viewed at wileyonlinelibrary.com]

It is important to highlight that we do not focus on comparing verified with nonverified reviews—an interesting topic studied by Anderson and Simester (2014) and by Kim et al. (2018a, 2018b). Instead, we focus on how the availability of OPV affects the reputation ecosystem of a platform, which includes both verified and nonverified reviews.

3.1 | Quasi-experimental setup

We use data from Amazon, “the largest internet-based retailer in the world by total sales and market capitalization” (Barney, 2014). Amazon allows users to submit reviews and rate products on a rating scale from one to five stars. To curate a high-quality reputation system, Amazon offers OPV: “If you bought the product you are reviewing from Amazon.com, you [can] label the review as an Amazon Verified Purchase. The Amazon Verified Purchase label offers Amazon.com customers additional context and helps them better gauge the quality and relevance of a product review”



FIGURE 2 An example of an Amazon-verified review header [Color figure can be viewed at wileyonlinelibrary.com]

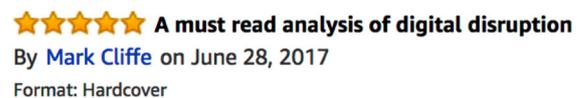


FIGURE 3 An example of an Amazon-nonverified review header [Color figure can be viewed at wileyonlinelibrary.com]

(Amazon, 2009, 2018). Figure 2 shows the header of a verified review. Figure 3 shows the header of a nonverified review.

The theoretical framework in Figure 1 argues that introducing OPV will affect product ratings and review characteristics through the reduction of fake reviews. To empirically test this framework, we design a quasi-experimental setup that

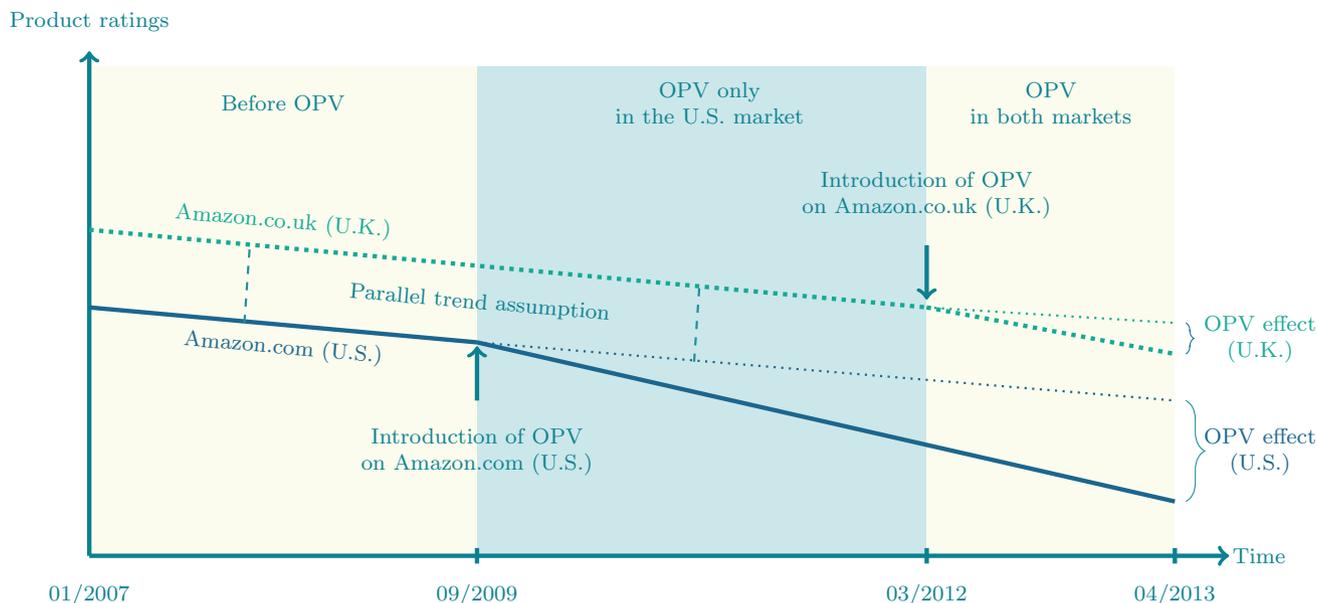


FIGURE 4 The focal quasi-experimental setting

Note: The diagram shows a conceptualization of the hypothesized effect on product ratings. (Figure S4 in Supporting Information J shows the actual parallel trend from our data.) We consider 6.3 years of reviews. Before September 2009, none of the markets were offering any type of purchase verification. Between September 2009 and March 2012, only Amazon U.S. was offering optional purchase verification (OPV). After March 2012, both markets were offering OPV [Color figure can be viewed at wileyonlinelibrary.com]

uses the global dimension of Amazon and the seemingly exogenous (with respect to product ratings and review characteristics) timing differences in the introduction of OPV in different markets. In particular, Amazon U.S. introduced OPV in September 2009 (Amazon, 2018). During the same period, Amazon U.K. was operating without purchase verification. Amazon U.K. finally introduced OPV in March 2012 (Amazon, 2012). Amazon uses the same unique product identifiers (ASIN) across its platforms globally. This consistency allows for studying how product ratings and review characteristics of the same products evolve in these two different markets (U.S.A. vs. U.K.), before and after the introduction of OPV. This setup creates a quasi-experimental design (Adamopoulos et al., 2020; Chen et al., 2018; Chevalier & Mayzlin, 2006; Huang et al., 2017; Mayzlin et al., 2014), where we assume that the introduction of the VP badge is exogenous to product ratings, review length, helpfulness, and reviewer experience.¹ Figure 4 visualizes this quasi-experimental setup and the hypothesized (Hypothesis 1) effects of introducing OPV on assigned ratings. Figure S4 in Supporting Information J shows the actual visualization of the parallel trend of the assigned ratings as derived from our data.

3.2 | Focal, dependent, and control variables

Our analysis focuses on products that were available in both the U.K. and the U.S. markets before and after the introduction of OPV in the two markets. Specifically, we analyze 336,043 book reviews posted between January 2007 and April 2013 in the two markets. *All products in our data have reviews both before and after the introduction of OPV.*

3.2.1 | Focal variable

We capture the focal treatment effect through the binary variable “OPV Introduction” that describes whether or not a review is posted after the introduction of OPV (Figure 4). Note that the treatment applies to a market as a whole. Hence, once OPV is introduced, its treatment applies to both verified and nonverified reviews.

3.2.2 | Dependent variables

Hypotheses 1 to 4 require four dependent variables: assigned rating, review length, reviewer experience, and review helpful votes. The assigned rating and the review helpful votes are extracted directly from the posted reviews. The number of words in a review defines review length. The total number of a reviewer’s posted reviews defines reviewer experience.

3.2.3 | Control variables

To better isolate the effect of the introduction of OPV, we control for various observed and unobserved confounding factors. Product fixed effects eliminate the unobserved time-invariant effects of product peculiarities; market fixed effects control for unobserved systematic differences between the United States and the United Kingdom; biweekly time fixed effects account for time trends.

Various *time-varying* covariates control for the observed population heterogeneity that affects our dependent variables. First, the time since the first posted review for each product

(“Product tenure”) controls for any tenure-specific effects. The current average of prior product ratings (“Accumulated product rating (simple average)”) before the submission of a new review controls for trends in the perceived product quality as well as for effects of previously posted ratings (Moe & Trusov, 2011). Similarly, the standard deviation between already assigned ratings (“Ratings deviation”) captures observed disagreements within a product’s reviewer population at the time of a new review. The total number of posted reviews (“Accumulated product reviews”) right before submission controls for the observed popularity of each product.

Amazon does not estimate product ratings by simply averaging all assigned ratings.² Instead, it uses machine learning models that evaluate multiple criteria to estimate product ratings. Hence, to control for possible discrepancies between the “Accumulated product rating (simple average)” and the machine learning–predicted product rating of Amazon’s that could correlate with our dependent variables we reverse-engineer Amazon’s algorithm through predictive modeling—we discuss the details of this process in Supporting Information E. The result is captured in the variable “Accumulated product rating (reverse engineered)” that measures the expected Amazon’s algorithm rating at each point in time. (Note that our dependent variable “Assigned rating” is not affected by Amazon’s algorithm, as it directly measures the user-assigned ratings.)

Another characteristic that might affect the process of posting reviews is the within-product review rankings. At each point in time, consumers might be intrigued to post a review depending on what reviews are being displayed on the product front page. To control for this variability, we build predictive models that estimate the likelihood of each review to be displayed on the front page (“Likelihood of top-ranked review”). Supporting Information F presents the details of this approach.

Finally, when estimating the effect of introducing OPV on review helpful votes, we control for both the assigned rating and the review length. Both of these measures are visible when a review receives helpful votes, and, as we know from prior studies (Kim et al., 2006; McCallum et al., 1998; Yin et al., 2014), they significantly correlate with the perceived helpfulness of a review.

Table 1 shows the descriptive statistics of the focal, dependent, and control variables. Figure S5 in Supporting Information J shows their correlations. Because we perform our analysis at the product level, we aggregate variables in biweekly observations.³ Sensitivity analyses show results for monthly (Table S12), weekly (Table S13), and nonaggregated observations (Table S14).

4 | EMPIRICAL EVALUATION

A difference-in-differences (DID) specification can empirically isolate the effect of introducing OPV on product ratings and review characteristics (Wooldridge, 2010). For a product p at market m and time t we can estimate the effect of the

introduction of OPV as follows:

$$DV_{pmt} = P + M + T + \gamma OPV_{mt} + \beta X_{pmt} + \varepsilon_{pmt}, \quad (1)$$

where $DV_{pmt} \in \{\text{Assigned rating, Review length, Review helpful votes, Reviewer experience}\}$; P , M , and T are the fixed effects of product, market (U.S., U.K.), and time (biweekly periods—see Tables S12, S13, and S14 for monthly, weekly, and nonaggregated observations); and X_{pmt} is the vector of the observed time-varying variables and the constant term. The coefficient of interest γ , is the DID estimate of the effect of introducing OPV on the dependent variable. If $\gamma > 0$, then the introduction of OPV had an overall positive effect on assigned ratings, length, helpful votes, and reviewer experience. Otherwise, if $\gamma < 0$, the introduction of OPV had a negative effect on these dependent variables.

Our setting quasi-simulates a randomized control trial under the assumption that the introduction of OPV is exogenous to assigned ratings and review characteristics—an assumption that aligns with Amazon’s announcement (Amazon, 2009) and we also empirically test in Supporting Information B.2. Furthermore, if our theoretical mechanism holds and OPV reduces fake reviews (most of which are positive), then the decision to introduce OPV is likely uncorrelated with any omitted variables that drive assigned ratings, review length, review helpful votes, and reviewer experience. Finally, as we empirically show in Section 5.1 and Supporting Information B, it is unlikely that a trend in assigned rating, review length, review helpful votes, and reviewer experience triggered Amazon’s decision to introduce OPV.

4.1 | Results

Table 2 shows the results. For each of the four dependent variables we estimate two specifications: Columns A1, B1, C1, and D1 show the effects of the introduction of OPV by only controlling for time, market, and product fixed effects. Columns A2, B2, C2, and D2 show the effects of the introduction of OPV under the complete specification of Equation (1). All columns support Hypotheses 1 to 4: Introducing OPV results in lower average product ratings (columns A1 and A2) and in longer reviews posted by more experienced reviewers that end up being more helpful (columns B1–D2).

To estimate the magnitude of these effects, we use the following formula:

$$\text{OPV-driven percentage change (\%)} \\ \in \left[\frac{\gamma}{\min(DV)} * 100, \frac{\gamma}{\max(DV)} * 100 \right]. \quad (2)$$

By considering all specifications, we find that the OPV-driven percentage change on assigned rating ranges $\in [-24\%, -1.4\%]$. Similarly, the effect on review length ranges $\in [0.2\%, 1,352\%]$, on reviewer experience $\in [0.1\%, 3,776\%]$,

TABLE 1 Descriptive statistics of the dependent, focal, and control variables

		Mean	Median	SD	Min	Max
Focal variable	OPV introduction	0.46	0	0.5	0	1
Dependent variables	Assigned rating	4.2	5	1.1	1	5
	Review length	120	76	138	1	4173
	Reviewer experience	95	23	331	1	12,490
	Review helpful votes	3.1	1	13	0	906
Time-varying control variables	Product tenure (biweekly)	73	70	48	0	162
	Accumulated product rating (simple average)	4.2	4.3	0.56	1	5
	Likelihood of top-ranked review	0.33	0.28	0.18	0	1
	Ratings deviation	0.96	1	0.39	0	2
	Verified reviews	0.27	0	0.43	0	1
	Accumulated product rating (reverse engineered)	4.3	4.4	0.3	2.8	4.8
	Review length	120	76	138	1	4173
	Accumulated product reviews	69	41	84	0	838
Fixed effects	Time	163 biweekly periods				
	Market	2 markets (U.K., U.S.)				
	Product	2,035 books				

Note: In all 336,043 book reviews over a period of 6.3 years (between January 2007 and April 2013) aggregated in 180,359 biweekly observations. The data come from two markets (U.S. and U.K.). All products appear in both markets before and after the introduction of OPV. Descriptive statistics show biweekly aggregates.

and on review helpful votes $\in [0.1, \infty)$. The OPV-driven percentage changes on the mean of each dependent variable (based on the estimates of the complete specification of Equation (1)) are -5.7% (Assigned rating), 6.7% (Review length), 39.7% (Reviewer experience), and 40.3% (Review helpful votes).

4.2 | Empirical investigation of the underlying mechanisms

The empirical analysis (Table 2) supports Hypotheses 1 to 4. However, it does not provide evidence in support of an underlying mechanism that drives the results. Our theory argues that introducing OPV should reduce the amount of fake reviews, the majority of which tend to be positive in an e-commerce platform (Figure 1). To empirically investigate this mechanism, we first need to identify which reviews are likely to be fake. As a result, we perform the following analyses:

Detection of fake reviews: Over time, Amazon has been identifying and removing fake reviews (Wehner, 2016). Since our data set was collected between 2010 and 2013, we can identify the reviews that Amazon has since removed, and we can label them as “fake.”

Prediction of fake reviews: We can build predictive models that estimate the probability of a review to be fake by using a unique, labeled Amazon reviews data set for deception detection (Saxena, 2018).

Text authenticity: We can estimate the text authenticity of each review through the Linguistic Inquiry and Word Count (LIWC) framework (LIWC, 2018) before and after the introduction of OPV.

We discuss the details and the results of these three approaches next.

4.2.1 | Detection of fake reviews

We collected our main data set between 2010 and 2013. Since then, Amazon has been allocating significant resources toward identifying and removing fake reviews (Wehner, 2016). Hence, a straightforward way to identify fake reviews is to find out which reviews from our corpus have Amazon since deleted. Amazon removed 3863 reviews, which we assume to be fake. Figure 5A1 shows that the biweekly frequency of appearance of fake reviews was higher before the introduction of OPV ($p < 0.001$). This indicates that proportionally more fake reviews were posted before the introduction of OPV. In addition, Figure 5A2 shows that the average product rating of the deleted (fake) reviews was higher ($p < 0.001$) than the product rating of the nondeleted ones, hence showing that the removed fake reviews were indeed more positive than the nonfake ones. As a result, if we assume that the removed reviews by Amazon were fake, then these fake reviews were, on average, more positive than non-fake reviews, and they were found more frequently before the introduction of OPV. This evidence provides empirical support for our main hypothesized mechanism described in Figure 1.

TABLE 2 Effects of the introduction of OPV on product ratings and review characteristics

	Assigned rating (H1)		Review length (H2)		Reviewer experience (H3)		Review helpful votes (H4)	
	(A1)	(A2)	(B1)	(B2)	(C1)	(C2)	(D1)	(D2)
Time (biweekly) FE	✓	✓	✓	✓	✓	✓	✓	✓
Product FE	✓	✓	✓	✓	✓	✓	✓	✓
Market FE	✓	✓	✓	✓	✓	✓	✓	✓
Product tenure		-0.06*		2.41		3.68		-0.93**
		(0.03)		(2.24)		(8.32)		(0.36)
Accumulated product rating (simple average)		0.17***		-4.32***		-3.06		0.13
		(0.01)		(0.83)		(3.19)		(0.11)
Ratings deviation		0.01		-2.73***		-5.60*		-0.46***
		(0.01)		(0.61)		(2.34)		(0.09)
Accumulated product reviews		0.00		-0.79		-4.36**		-0.31**
		(0.01)		(0.72)		(1.56)		(0.10)
Verified reviews		0.14***		-5.80***		-23.30***		0.18***
		(0.01)		(0.42)		(1.27)		(0.04)
Likelihood of top-ranked review		0.16***		83.93***		36.26***		-0.12
		(0.01)		(0.46)		(1.25)		(0.09)
Accumulated product rating (reverse engineered)		-0.06***		3.02**		0.32		-0.11
		(0.01)		(1.09)		(2.88)		(0.16)
Review length								2.06***
								(0.15)
Review length								-1.31***
								(0.06)
OPV introduction	-0.07***	-0.24***	13.52***	8.00***	16.07***	37.76***	1.71***	1.25***
	(0.01)	(0.01)	(1.30)	(1.20)	(3.62)	(4.08)	(0.14)	(0.14)
Observations	180,359	180,359	180,359	180,359	180,359	180,359	180,359	180,359
R ²	0.01	0.04	0.11	0.43	0.01	0.02	0.02	0.05

Note: Clustered standard errors (on products) in parentheses. The constant term is estimated but omitted from the table. FE: fixed effects. (***p < 0.001, **p < 0.01, *p < 0.05).

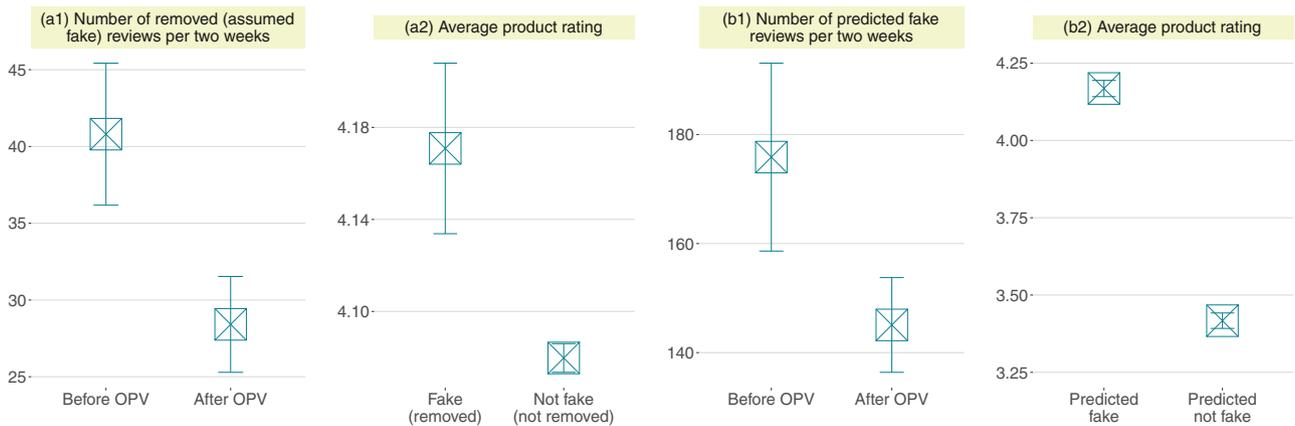


FIGURE 5 After the introduction of OPV, the platform observes fewer fake reviews, which tend to be positive
 Note: Figures A1 and A2 assume as fake reviews those that Amazon subsequently removed (Section 4.2.1). Figures B1 and B2 assume as fake reviews the ones annotated by predictive models (Section 4.2.2). Error bars show 95% confidence intervals [Color figure can be viewed at wileyonlinelibrary.com]

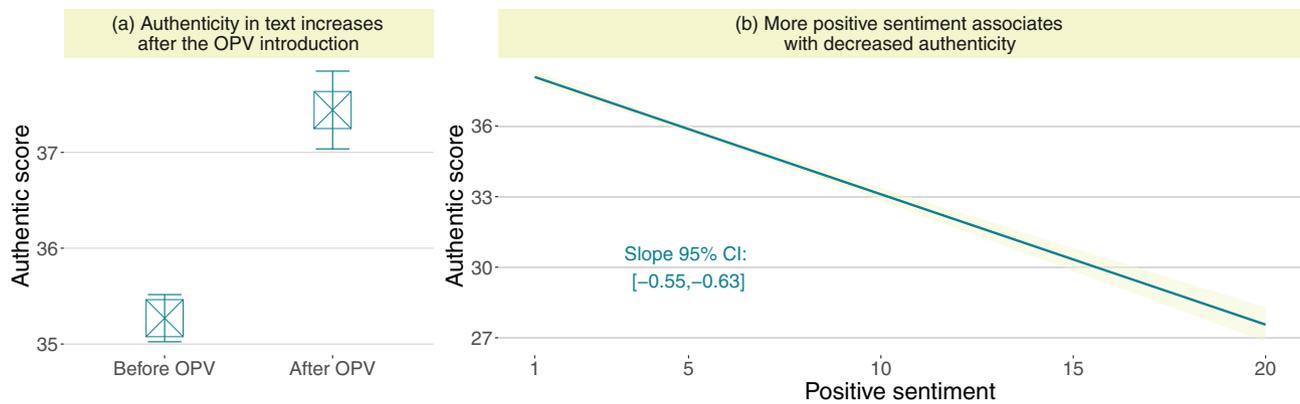


FIGURE 6 The introduction of OPV correlates with more authentic and less positive reviews

Note: The review text was less authentic before the introduction of OPV (A). Furthermore, there is a negative association between text authenticity and positive review sentiment (B). Error bars show 95% confidence intervals. CI: confidence interval [Color figure can be viewed at wileyonlinelibrary.com]

4.2.2 | Prediction of fake reviews

Predicting whether or not a review is fake is an inherently difficult task (Kumar et al., 2018b, 2019; Luca & Zervas, 2016). One of the main difficulties originates from the lack of ground truth data. Recently, researchers have published a data set of labeled Amazon reviews for deception detection (Saxena, 2018). The data set includes 21,000 Amazon reviews, of which 10,500 are labeled as “fake.” We use these data to learn models that predict the likelihood of a review to be fake according to the review text. We discuss this process in Supporting Information G. Figure 5B1, 5B2 shows the results in terms of the number of fake reviews before and after the introduction of OPV and in terms of product ratings for fake and nonfake reviews. The results provide additional support to the hypothesized mechanism: Before the introduction of OPV, the prediction task identified more fake reviews. In addition, the predicted fake reviews were more positive than the nonfake ones.

4.2.3 | Text authenticity

Next, we focus on the raw review text. Recent developments in text analysis allow to detect concepts related to text authenticity. Specifically, we use the LIWC package to analyze the unstructured review text (LIWC, 2018). LIWC adopts a dictionary-based approach and has been successfully used to estimate emotionality in various contexts (Goes et al., 2014; Hong et al., 2016; Sridhar & Srinivasan, 2012; Yin et al., 2014). To capture authenticity, the LIWC package estimates the “Authentic” metric: “higher numbers are associated with a more honest, personal, and disclosing text; lower numbers suggest a more guarded, distanced form of discourse” (LIWC, 2018). Figure 6A shows that after the introduction of OPV, the “Authentic” scores increased significantly ($p < 0.001$). In addition, Figure 6B shows that authenticity has a negative relationship ($p < 0.001$) with the review sentiment in the focal data set. Hence, the two figures provide

further support to the main theoretical mechanism presented in Figure 1.

4.2.4 | Fake reviews are shorter, posted by less-experienced reviewers, and are less helpful

Beyond the main mechanism that focuses on the reduction of positive fake reviews, our theoretical framework further claims that fake reviews are likely shorter, written by less experienced reviewers, and, as a result, are less helpful. Similar to before, we analyze these characteristics of a subset of reviews that our predictive model labels as fake and as truthful (Supporting Information G). Figure 7 shows the results of this analysis. Predicted fake reviews are significantly ($p < 0.001$) shorter than predicted nonfake reviews (Figure 7A). Furthermore, the average reviewer experience of reviews predicted to be truthful is almost four times larger than that of reviewers who posted fake reviews. Finally, on average, predicted truthful reviews are 2.5 times more helpful than predicted fake ones. As a result, these observations provide empirical support of the mechanisms that structure Hypotheses 2, 3, and 4.

4.2.5 | Disconfirmation-driven negative verified reviews

Perhaps the hardest component of our mechanism to test is that the introduction of OPV added an indirect cost on fake reviews due to expectation disconfirmation (Section 2.2.3 and Figure 1). Conceptually, if expectation disconfirmation indeed adds an indirect cost on posting fake positive reviews, then we should be able to observe in the data relatively more negative verified reviews following positive fake reviews. Simply put, positive fake reviews at time t should be followed by relatively more verified negative reviews at times $t + 1$, $t + 2$, and so forth. To test whether we observe this pattern in our data, we estimate for each product the average rating

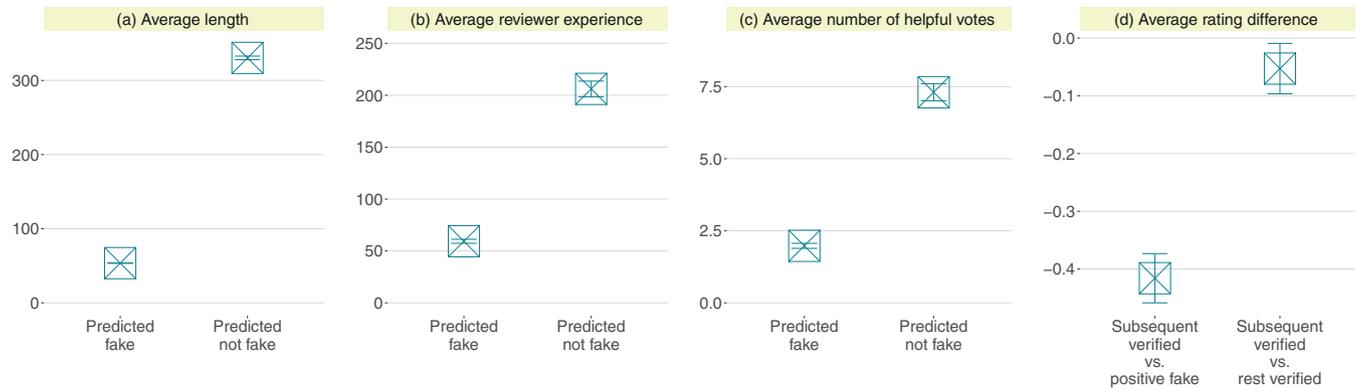


FIGURE 7 Empirical evidence shows that fake reviews are shorter, less helpful, and posted by less-experienced reviewers. Positive fake reviews are followed by relatively more negative verified reviews
Note: Predicted fake reviews are shorter (A), posted by less-experienced reviewers (B), and they accumulate fewer helpful votes (C). Subsequent verified reviews are significantly more negative than (1) positive fake reviews and (2) the rest verified reviews of the focal product. Subsequent verified reviews include verified reviews posted up to 3 months after a positive fake review. Rest verified reviews include all other verified reviews of the product. Positive fake reviews are reviews predicted to be fake that assigned ratings greater than 3. Error bars show 95% confidence intervals [Color figure can be viewed at wileyonlinelibrary.com]

of the verified reviews posted within 3 months after a positive predicted fake review. We name these reviews as “*subsequent verified*.” Then, we compare these ratings with the product-specific ratings of (1) the predicted positive fake reviews (i.e., fake reviews with a product rating greater than 3) and (2) all verified reviews excluding the *subsequent verified* ones (“rest verified”). For our theorized disconfirmation mechanism to hold, the subsequent verified ratings should be significantly lower than both the predicted fake ratings and the rest verified ones. Figure 7D shows exactly this relationship: Subsequent verified ratings are on average 0.053 ($p < 0.05$) more negative compared with the rest verified reviews, and 0.41 ($p < 0.001$) more negative compared with the positive fake reviews.

4.2.6 | The effect of the introduction of OPV on the overall volume of reviews

Our hypothesized mechanism is also connected to the overall review volume. In particular, our theory implies that since the introduction of OPV reduces the number of fake reviews, it will all else being equal subsequently reduce the overall review volume. Table 3 tests this expected effect. The dependent variable in this regression is the biweekly aggregated received reviews per product. Similar to Table 2, we test specifications with and without the vector of time-varying control variables. Columns A1 and A2 show the results: As expected, the introduction of OPV reduced the product-specific volume of reviews.

On top of this analysis, we use the predicted probabilities of a review to be fake to estimate whether the observed effect is more pronounced on fake reviews. Based on the analysis in Supporting Information G, we create a binary variable “Fake review” that describes whether a given review is predicted to be fake. Then, we interact this variable with the introduc-

tion of OPV and rerun the full specification of Equation (1) to see the combined effect on review volume. Conceptually, the coefficient of the interaction variable shows how much stronger the OPV effect is when the reviews are predicted to be fake. Column A3 shows the results: The negative coefficient of the interaction term suggests that the observed review volume reduction is more pronounced on likely fake reviews, therefore further supporting the hypothesized mechanisms in Figure 1.

Overall, six different independent tests provide empirical support for the proposed theory (Figure 1): Introducing OPV reduces the amount of fake reviews, the majority of which tend to be positive in e-commerce platforms. As we discuss in Section 6.3, even though none of these tests is perfect, the fact that all point in the same direction is encouraging that our hypothesized theoretical mechanisms are indeed driving the observed results. Next, we empirically test and eliminate a set of alternative explanations.

5 | ALTERNATIVE EXPLANATIONS

Our study makes a series of assumptions regarding the focal quasi-experimental setting. At the same time, the analysis in Section 4.2 does not rule out alternative mechanisms that could also be driving the observed results. To provide additional empirical support for our theory, in this section we:

- test the validity of our quasi-experimental setup through various statistical tests and subsample analysis (Section 5.1, Supporting Information B),
- empirically test and eliminate alternative theoretical mechanisms that could partially explain the observed results (Section 5.2, Supporting Information C), and
- test the sensitivity of the observed results under various groupings over time as well as through log-transformations

TABLE 3 The volume of (mostly fake) reviews decreases after the introduction of OPV

	Volume		
	(A1)	(A2)	(A3)
Time (biweekly) FE	✓	✓	✓
Product FE	✓	✓	✓
Market FE	✓	✓	✓
Product tenure		−0.22** (0.09)	−0.23** (0.09)
Accumulated product rating (simple average)		0.10*** (0.02)	0.10*** (0.02)
Review length		−0.08*** (0.01)	−0.08*** (0.01)
Assigned rating		0.08*** (0.01)	0.08*** (0.01)
Accumulated product reviews		0.00 (0.06)	0.00 (0.06)
Ratings deviation		0.06*** (0.02)	0.06*** (0.02)
Verified reviews		0.06*** (0.01)	0.06*** (0.01)
Likelihood of top-ranked review		−0.01* (0.01)	−0.02** (0.01)
Accumulated product rating (reverse engineered)		0.06 (0.05)	0.06 (0.05)
OPV introduction	−0.47*** (0.03)	−0.51*** (0.03)	−0.49*** (0.03)
OPV introduction × Fake review			−0.66*** (0.03)
Observations	180,359	180,359	180,359
R ²	0.05	0.05	0.06

Note: Clustered standard errors (on products) in parentheses. The constant term is estimated but omitted from the table. FE: fixed effects. (***) $p < 0.001$, (**) $p < 0.01$, (*) $p < 0.05$.

of the dependent variables and alternative models (Supporting Information H).

Table 4 summarizes these analyses, which we discuss in detail in the following paragraphs and the Supporting Information.

5.1 | Correctness of the quasi-experimental setup

5.1.1 | Market-specific trends

A test that can provide some confidence in support of our DID identification strategy adds market-specific time trends to the list of controls (p. 238 of Angrist & Pischke, 2008; Adamopoulos et al., 2020; Besley & Burgess, 2004).

Specifically, in our context, we estimate the following specification:

$$DV_{pmt} = P + M + \mu * t + \gamma OPV_{mt} + \beta X_{pmt} + \varepsilon_{pmt}, \quad (3)$$

where μ is a market-specific dummy and t is the time trend. This specification allows treated (U.S.) and control (U.K.) markets to follow different trends, hence capturing any systematic time-changing differences between the two markets that could affect the results. “It is heartening to find that the estimated effects of interest are unchanged by the inclusion of these trends, and discouraging otherwise” (Angrist & Pischke, 2008, p. 238). Table S1 shows that the main results are robust to the inclusion of market-specific time trends. Despite controlling for market-specific trends, introducing OPV still yields lower product ratings, and longer and more helpful reviews by more experienced reviewers, hence suggesting that any market-specific trends do not drive our observed results (Angrist & Pischke, 2008).

5.1.2 | Placebo regressions

A different test runs placebo regressions that reveal whether the DID specifications pick up any coincidental spurious entry effects (Chan & Ghose, 2014). Specifically, in our scenario, we randomly sample a period before introducing OPV in the United States, and we assign a placebo the introduction of OPV in the middle of the sampled period. Table 6 in Supporting Information A shows the results. For all dependent variables, the coefficients of the placebo OPV are not statistically significantly different than zero. Hence, the placebo regressions provide empirical evidence that the observed effects are not coincidental.

5.1.3 | Potential contamination from browsing

A valid concern about our quasi-experimental setup is that there might be contamination across treatment and control markets. For instance, U.K. users can browse reviews on U.S. websites and, as a result, be affected by the introduction of OPV (treatment). To empirically test for this possibility, we do the following:

- Estimate fixed-effect panel regressions for the U.S. market alone, before and after the introduction of OPV.
- Randomly select products so that they form independent sets across markets. Through this randomization, each product appears only in a single market.
- Match products based on their observed characteristics through a nearest neighbor approach. Then, we randomly choose matched pairs so that each product appears only in a single market. Through this matching,

TABLE 4 Investigation of alternative explanations

Analysis	H1	H2	H3	H4	Discussion
<u>Alternative explanation:</u> Imperfections of the quasi-experimental DID setting drive the results					
<u>Solution:</u> Test the assumptions of DID:					
Market-specific trends do not drive the results	✓	✓	✓	✓	Section 5.1.1, Table S1
Coincidental events do not drive the results (placebo regressions)	✓	✓	✓	✓	Section 5.1.2, Table S2
Granger causality holds (relative time regressions)	✓	✓	✓	✕	E-C B.1 Figure S3
the introduction of OPV was likely exogenous to the four dependent variables	✓	✓	✓	✓	E-C B.2, Table S9
<u>Alternative explanation:</u> Contamination through global accessibility of the U.S. reviews drives the results					
<u>Solution:</u> Subsample analysis:					
Analysis of U.S.-only reviews	✓	✓	✓	NS	Section 5.1.3, Table S3
Random selection of products (appear in one market)	✓	✓	✓	✓	Section 5.1.3, Table S4
Nearest neighbor matching of products	✓	✓	✓	✓	Section 5.1.3, Table S5
Products with early post-treatment U.K. reviews	✓	✓	✓	✓	Section 5.1.3, Table S6
<u>Alternative explanation:</u> Concurrent events drive the results					
<u>Solution:</u> Fixed effects on concurrent events	✓	✓	✓	✓	Section 5.2.1, Table S7
<u>Alternative explanation:</u> Systematic differences between the two markets drive the results					
<u>Solution:</u>					
Market and time FE control for unobserved static and within period systematic differences	✓	✓	✓	✓	Sections 4, 5.1.1, Tables 2, S1
No empirical evidence of imbalance between the number of fake reviews of the two markets	✓	✓	✓	✓	Section 5.2.2, Figure S1A
Differences in publication dates across the two markets do not drive the results	✓	✓	✓	✓	Section 5.2.3, Table S8
<u>Alternative explanation:</u> Fake reviewers changed their behavior					
<u>Solution:</u> Empirical evidence shows that fake reviewers did not change their behavior	NA	✓	NA	✓	Section 5.2.4, Figure S1B and Figure S1C
<u>Alternative explanation:</u> New restrictions on verified reviews drive the results					
<u>Solution:</u> Empirical evidence shows that these restrictions do not drive the results	✓	✓	✓	✓	Section 5.2.5, Figure S1D
<u>Alternative explanation:</u> Regular reviewers changed their behavior					
<u>Solution:</u> Empirical evidence shows that regular reviewers did not change their behavior	✓	✓	✓	✓	Section 5.2.6, Figure S2
<u>Alternative explanation:</u> Results are driven by products of different quality					
<u>Solution:</u> Empirical evidence shows that results hold across products of varying qualities, with some products being more affected than others	✓	✓	✓	✓	E-C C, Figure 14
<u>Alternative explanation:</u> Results are driven by the chosen aggregation thresholds and models					
<u>Solution:</u> Alternative aggregations and models:					
Regressions on weekly, monthly, and non-aggregated data	✓	✓	✓	✓	E-C H, 16, Tables S13, S14
Logarithmic transformations	✓	✓	✓	✓	E-C H, Table S15
Negative binomial regressions	✓	✓	✓	✓	E-C H, Table S16

Note: ✓: Hypothesis supported, or relevant assumption holds. ✕: There is some pre-trend significance, but post-trend effects are much stronger (Figure S3). NS: The effect of the introduction of OPV on review helpful votes is not statistically significant ($p > 0.05$). NA: Alternative explanation does not apply to the column dependent variable. E-C: Supporting Information.

we can control for matched-pairs fixed effects across both markets (by replacing the product fixed effects of our main specification with matched-pairs fixed effects).

- Identify a group of products that had reviews in the United Kingdom before a review was posted in the United States in the posttreatment period.

Because the first approach focuses only on the U.S. (treated) market, contamination from browsing cannot exist. In the next two approaches, we expect contamination

effects to be absorbed in the product fixed effects, as each product now only appears in one of the two markets. Finally, in our last test, because all reviews in the U.S. were posted in the pre-treatment period, contamination might only come from other products. Tables S3, S4, S5, and S6 in Supporting Information A show the results. In all four scenarios, despite the significant subsampling, the observed effects remain qualitatively unchanged. The only exception is the effect of the introduction of OPV on review helpful votes in Table S3, which is positive but insignificant.

5.1.4 | Additional tests of our quasi-experimental setup

Supporting Information B provides additional tests (relative time regressions, visualization of the parallel trend, exogeneity of the introduction of OPV) that show additional support in favor of the correctness of the DID specification.

5.2 | Empirical elimination of alternative mechanisms

Despite the presented empirical evidence in support of the theoretical framework, alternative mechanisms could work in parallel to partially drive the main results.

5.2.1 | Concurrent events

A valid concern is that our specification of Equation (1) does not control for other concurrent events that cannot be potentially fully captured through the market and time fixed effects and that could have affected the studied dependent variables. To test for this possibility, we capture all relevant concurrent Amazon events—as listed in the History of Amazon Wikipedia page⁴—through dummy variables. In particular, we create dummies for the (1) launch of Amazon Kindle (November 2007), (2) introduction of Nook (October 2009), (3) introduction of iBook (January 2010), (4) launch of Azure (February 2010), (5) launch of Amazon video (February 2011), (6) collection of sales taxes in California (July 2011), (7) announcement of Kindle Fire (September 2011), (8) launch of Amazon Appstore (March 2011), (9) acquisition of Kiva systems (March 2012), (10) collection of sales taxes in Nevada and Texas (April 2012), and (11) announcement of Kindle Fire HD (September 2012). Table S7 shows the results, which are aligned with the main results presented in Table 2. Overall, these external events do not affect the hypothesized and observed effect of the introduction of OPV on the assigned review ratings, length, helpfulness, and observed reviewer experience.

5.2.2 | Imbalance of fake reviews between the two markets

An alternative explanation argues that the two markets are systematically different in the way that they attract fake reviews (i.e., one market attracts more fake reviews than the other). Figure S1A shows that this is likely not the case. The two markets do not significantly differ in terms of predicted fake reviews (Supporting Information G), as shown by the box plots that are not statistically significantly different.

5.2.3 | Differences in publication dates

An alternative argument suggests that there might be a systematic lag of the book publication dates between the two markets (e.g., the U.S. market publishes books earlier than the U.K. market). In general, it might be more likely to post fake reviews in the early period of a publication. Hence, the existence of a time-varying systematic difference that is not captured by the market fixed effects could be driving the results.

The market-specific trends analysis we discussed in Section 5.1.1 and the fact that all included products have reviews in all periods presented in Figure 4 are encouraging first steps that any time-varying systematic differences in publication dates do not drive our results.

Yet, by looking into our data, we observe a systematic bias in terms of review arrivals: 50% of U.S. books receive reviews 33.5 (or more) days earlier than U.K. books. To test whether this systematic difference drives our results, we randomly undersampled books from the United States so that the average difference between publication days of the United States and the United Kingdom is zero. The results in Table S8 align with our main results, suggesting that this systematic difference is likely captured in the market fixed effects and does not drive our observed results.

5.2.4 | Change in fake reviews

An alternative explanation argues that fake reviewers adapted their behavior after the introduction of OPV, such that they started providing longer and hence more helpful reviews. Figures S1B and S1C show that this is not the case: there is no significant difference between the length and the helpfulness of predicted fake reviews before and after the introduction of OPV.

5.2.5 | Amazon restricts the number of posted nonverified reviews per week

After the introduction of OPV, Amazon started restricting the number of nonverified reviews that a person can post on the platform to five per week—it is unclear whether Amazon implemented this policy before or after 2016.⁵ To test whether this decision affects our results, we check (1) whether there is a significant change in the average number of weekly nonverified reviews that a reviewer posts before and after the introduction of OPV, and (2) whether we have reviewers in our data that have posted more than five reviews after the introduction of OPV. Figure S1D shows the results: There is no statistically significant change in the number of posted nonverified reviews per week. Interestingly, multiple reviewers in our data posted more than five nonverified reviews after the introduction of OPV, suggesting that the aforementioned policy was likely not in place during our data collection period.

5.2.6 | Change of regular-reviewer behavior

Finally, an alternative mechanism argues that OPV changed reviewer behavior, such that they started writing fewer, lengthier, and more critical reviews. As a result, the observed effects are not driven exclusively by the reduction of positive fake reviews.

To test this possibility, we create a data set that captures reviewer behavior before and after the introduction of OPV. Practical constraints do not allow studying the behavior of all reviewers in our data set before and after the introduction of purchase verification. Such an analysis would require us to collect ~ 29,432,096 reviews. Instead, to investigate this alternative mechanism, we randomly choose 1000 reviewers from our data set who have posted reviews both before and after the introduction of OPV. Because they post multiple reviews over time, we name these reviewers as “regular reviewers.”

By collecting the complete reviewing histories of these regular reviewers, we create a separate data set of 20,462 reviews posted between September 2008 and September 2010 (1 year before and 1 year after the introduction of OPV in the United States). We then generate biweekly aggregations of reviewer behavior regarding the number of reviews posted, average product ratings, and the average length of posted reviews. Finally, we check whether there are significant differences in regular-reviewer behavior before and after the introduction of purchase verification.

Figure S2 shows the results. In Figure S2A, we find no evidence that reviewers post more reviews after the introduction of OPV. Similarly, in Figure S2B, there is a statistically insignificant *increase* in the average product ratings ($p > 0.05$). Finally, Figure S2C shows that reviewers do not post longer reviews after the introduction of OPV ($p > 0.05$). Overall, empirical evidence suggests that regular-reviewer behavior remains unchanged after the introduction of purchase verification. Combined with the fact that the volume of reviews decreases after the introduction of OPV (Section 4.2.6), this unchanged reviewer behavior suggests that the observed reduction in reviews comes indeed from our hypothesized decreased rate of fake reviews.

6 | CONTRIBUTIONS TO RESEARCH AND PRACTICE

We explored the effects of introducing OPV on the reputation ecosystem of an e-commerce platform. Theoretically, we argued that, despite the nonobvious additional cost of posting fake reviews, introducing OPV will likely reduce fake reviews because of indirect costs related to disconfirmation between expected and realized product quality. Using a quasi-experimental setup and a large set of product ratings and reviews from Amazon, we found that, indeed, introducing OPV reduces the amount of fake reviews, hence resulting in lower product ratings and longer

and more helpful reviews written by more experienced reviewers.

6.1 | Contributions to research in the OM-IS interface

By investigating these novel effects of OPV, our research extends the work on user-generated content and digital platforms in the OM-IS interface (Ba et al., 2020; Cui et al., 2018; Guan et al., 2020; Kokkodis et al., 2020; Kumar et al., 2018c, 2018a; Lau et al., 2018; Mallipeddi et al., 2021; Mejia et al., 2021), and, in particular, OM-IS research on response bias and fake reviews (Chen et al., 2016; Kumar et al., 2018b, 2019; Lappas et al., 2016; Lee et al., 2018). Closely relevant research to ours compared platforms that require verified reviews (e.g., Expedia) with platforms that do not offer any type of purchase verification (e.g., TripAdvisor) to find that fake reviews are more prevalent in platforms without purchase verification (Mayzlin et al., 2014). Other relevant works focused only on the period after the introduction of OPV and compared nonverified with verified reviews to find that the latter tend to be more positive (Anderson & Simester, 2014) and to have a stronger effect on sales (He et al., 2020; Kim et al., 2018a, 2018b). Our study is unique, as it is the first to investigate how *introducing* OPV (i.e., studying both the before and after periods) alters product ratings and review characteristics in e-commerce platforms.

Furthermore, no prior study has connected OPV with product ratings, review length, reviewer experience, and review helpfulness. Our research is the first to suggest that introducing purchase verification associates with a decrease in product ratings and an increase in review length, review helpfulness, and reviewer experience. Through a comprehensive empirical analysis, our work shows that these observations are a direct result of reducing fake reviews, the majority of which are positive. This new mechanism of action is relatively counterintuitive, as OPV does not exclude nonverified reviews and hence it does not add a monetary cost in posting fake reviews. However, it creates an environment where nonverified reviews have less value than verified ones. This implicit cost discourages fake reviewers from posting nonverified reviews, as such nonverified reviews might trigger negative verified reviews due to expectation disconfirmation. As a result, our work is the first to explain why OPV will likely benefit all aspects of the reputation ecosystem of an e-commerce platform.

Aside from these main research contributions, our analyses on identifying fake reviews provides new empirical evidence that, indeed, the majority of fake reviews in platforms where products do not have immediate substitutes are positive. This observation complements previous findings that fake reviews can be both positive and negative in environments where local competition exists (Mayzlin et al., 2014), and provides additional context on when review manipulation tends to be positive (Lee et al., 2018).

Interestingly, our work also shows that the most significant impact of introducing OPV applies to reviewers who eventually do not post a review (discouraged fake reviewers). This result can inform and guide future work in e-commerce and other types of online platforms to consider not only how design choices affect people who use the platform but also how they affect those who may choose not to use the platform to begin with.

In terms of empirical methodology, our study uses a new quasi-experimental setup that allows the comparison of the state of the two platforms before and after the introduction of OPV. This setup helps toward isolating the hypothesized effects, while it also overcomes a significant limitation that hinders efforts to quantify the effect of defense mechanisms against fake reviews: the lack of a deterministic way of establishing if a given review is fake or not. Finally, by using deep learning and predictive modeling analytics to identify fake reviews and estimate product ratings and review rankings (Sections 4.2 and 5, Supporting Information E, F, and G), we also contribute to the call for applications of machine learning and artificial intelligence (AI) in business operations (Cui et al., 2018; Geva & Saar-Tsechansky, 2021; Kumar et al., 2018c, 2018b).

6.2 | Contributions to practice

Given the documented benefits of allowing more customers to provide reviews (Wang et al., 2019), the results of this work have practical implications for the business operations of e-commerce platforms that either already host both verified and nonverified online reviews or are considering introducing OPV. First, we showed for the first time that OPV is still an effective mechanism for reducing fraudulent reviews, even though it does not require purchase verification to review. Because of this reduction of overly positive fake reviews, the average product rating decreases and likely becomes more representative of the true product quality. In addition, OPV resulted in higher quality textual reviews, without significantly reducing the total number of good quality (truthful) reviews, which is an unavoidable cost of platforms that introduce required purchase verification.

The positive effects of OPV extend beyond the ones studied in this paper. A better reviewing environment that includes OPV allows for efficient ranking algorithms to increase attention on high-quality reviews. For instance, Amazon now uses OPV as a feature to algorithmically rank reviews and products.⁶ As a result, by explicitly (through the reduction of fake reviews) and implicitly (through ranking algorithms that use OPV as a feature) eliminating information overload, OPV cannot only allow platforms to develop higher quality reputation systems, but also generate fruitful conditions for platform operations to provide relevant incentives and attract new high-quality reviewers.

Finally, our work can guide platforms that currently do not offer any type of purchase verification (e.g., Yelp or TripAdvisor). Even though in these platforms the effects of OPV

on product ratings might not be as negative—since these platforms on expectation receive more fake negative reviews than an e-commerce platform (Mayzlin et al., 2014)—the overall reputation ecosystem will likely improve by hosting more truthful, representative reviews. Hence, by showing that requiring purchase verification is not necessarily the only way to curate high-quality reputation systems, our work can motivate these platforms to allow OPV (e.g., through electronic receipt submission). More importantly, our work is the first show that such an action will succeed in reducing fake reviews through the discussed disconfirmation-induced cost. Even if very few users end up submitting verified reviews, the disconfirmation-induced effects can have significant positive spillovers in the overall reputation ecosystems of these platforms.

6.3 | Generalizability, limitations, and future directions

Generalizability

Our analysis focuses on book reviews. We chose books because a book's lifetime is significantly longer than products from other categories such as electronics. For instance, it would have been extremely hard—if not impossible—to collect reviews that span 6 years in both markets for a camera, as a camera's lifetime can be as short as a couple of years. However, the hypothesized mechanism (Figure 1) is product-type independent. Hence, we expect the mechanism to generalize to any type of product that attracts relatively more positive than negative fake reviews (Lee et al., 2018; Luca & Zervas, 2016).

Limitations

Despite the quasi-experimental setup, and although we have taken careful steps to provide empirical support to our theoretical arguments, our study relies on observational data and, as a result, has empirical limitations. For example, unobserved events that could be interfering with the results might not be fully captured by our time fixed effects and other control variables. Yet, the strong evidence (Sections 4.2 and 5) supporting our theory suggests that, at least, it is one of potentially multiple causal paths that drive the observed effects.

For instance, we argue that the introduction of OPV reduced the value of fake reviews. But other mechanisms could also be contributing to the decreased value of nonverified reviews:

- Increased ability to identify fake reviews: The introduction of OPV could have improved Amazon's ability to identify and remove fake reviews because OPV could be a useful predictive feature in Amazon's fake-review detection algorithms. Even though generally possible, we argue that this mechanism likely does not affect our data and results. Specifically, our data collection happened between 2010 and 2013, while the first evidence that Amazon started removing fake reviews comes from

2016 (Wehner, 2016). In fact, our analysis in Section 4.2.2 identified thousands of removed fake reviews. Together, these observations suggest that Amazon was not actively detecting and removing fake reviews during our collection period.

- Effects of the product ranking algorithm: Both Amazon's review and product ranking algorithms use OPV as a predictive feature.^{7,8} Hence, fake reviewers could have realized that their reviews are not as effective as they used to be since Amazon's algorithms started discounting them after the introduction of OPV. The variables "Likelihood of top-ranked review" and "Verified reviews" in our analysis capture a significant portion of such anticipated ranking effects. Furthermore, similar to before, it is unclear if (and to what extent) Amazon was using these algorithms during our collection window between 2010 and 2013.

Note that the catalyst in these two alternative scenarios is the same: The introduction of OPV improved the platform's ability to build better ranking models and restrict reviews and reviewers. A different type of study could examine whether these mechanisms interact with our main mechanism—a study that would survey businesses (perhaps anonymously) of how they alter their behavior in the presence of OPV.

A different type of limitation originates from the error propagation of our two predictive models described in Supporting Information E and F. We believe that the benefits of using these algorithms outweigh these prediction errors, as they allow us to control for portions of the effects of Amazon's algorithms that would have been impossible to control for otherwise. Besides, the fact that our results hold without these predictive quantities (columns A1, B1, C1, and D1 in Table 2) increases our confidence that these predictive errors do not drive the results.

Similarly, the prediction of fake reviews described in Supporting Information G is not perfect. Furthermore, the additional fake-detection approaches we use in Sections 4.2.1 and 4.2.3 also have limitations. However, the fact that all three independent and very different approaches support our hypothesized mechanism of Figure 1 increases our confidence that this mechanism indeed drives the observed results.

As we discussed in Section 5.1.3, some U.K. customers could have been treated by browsing reviews on the U.S. platform, which had already introduced OPV. Three different types of analysis in Section 5.1.3 provide encouraging evidence that this effect is limited and does not drive our observed results. It is also important to mention that even if this contamination exists in a way not captured by our analysis, then the observed effects are likely underestimated: Potential U.K. fake reviewers would have realized that U.K. customers can browse the U.S. website and get treated with OPV, and hence they would have likely been deterred from posting fake reviews, decreasing the amount of fake reviews in the pretreatment period of the U.K. platform.

6.4 | Conclusion

In conclusion, this paper is the first to study how introducing OPV affects the reputation system of an e-commerce platform. The presented empirical analysis of Amazon books showed evidence that introducing OPV reduces the number of fake reviews, most of which tend to be positive. This reduction of fake reviews results in lower, more representative product ratings, as well as in longer and more helpful reviews posted by more experienced reviewers. As a result, introducing OPV creates a better, more credible reviewing environment for both reviewers and consumers.

ORCID

Marios Kokkodis  <https://orcid.org/0000-0002-5037-6060>
Theodoros Lappas  <https://orcid.org/0000-0002-4669-4170>
Gerald C. Kane  <https://orcid.org/0000-0001-5239-1902>

ENDNOTES

- ¹ A similar setup that compares timing variations of different treatments across different Amazon markets was used in Adamopoulos et al. (2020).
- ² <https://www.amazon.com/gp/help/customer/display.html?nodeId=GQUXAMY73JFRVJHE>
- ³ Aggregations are useful in our context as they help decrease potential popularity effects of products that receive significantly larger-than-average amounts of reviews.
- ⁴ https://en.wikipedia.org/wiki/History_of_Amazon
- ⁵ <https://sellerengine.com/amazon-verified-product-reviews/>
- ⁶ <https://www.amazon.com/gp/help/customer/display.html?nodeId=GQUXAMY73JFRVJHE>
- ⁷ <https://sellerengine.com/amazon-verified-product-reviews/>
- ⁸ <https://amzn.to/2QApUPe>

REFERENCES

- Adamopoulos, P., Todri, V., & Ghose, A. (2020). Demand effects of the internet-of-things sales channel: Evidence from automating the purchase process. *Information Systems Research*, 32(1), 238–267.
- Akerlof, G. A. (1978). The market for lemons: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235–251). Elsevier.
- Amazon (2009). *Verified purchase reviews*. <https://bit.ly/339qViU>
- Amazon (2012). *Amazon verified purchase*. <https://bit.ly/2TSYvnp>
- Amazon (2018). *About Amazon verified purchase reviews*. <https://goo.gl/mcFnzb>
- Anderson, E. T., & Simester, D. I. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, 51(3), 249–269.
- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science*, 57(8), 1485–1509.
- Awad, N. F., & Ragowsky, A. (2008). Establishing trust in electronic commerce through online word of mouth: An examination across genders. *Journal of Management Information Systems*, 24(4), 101–121.
- Ba, S., Jin, Y., Li, X., & Lu, X. (2020). One size fits all? The differential impact of online reviews and coupons. *Production and Operations Management*, 29(10), 2403–2424.
- Bakos, Y. (2001). The emerging landscape for retail e-commerce. *Journal of Economic Perspectives*, 15(1), 69–80.
- Barney, J. (2014). *Amazon urges California referendum on online tax*. <https://on.ft.com/2TVE3nD>

- Bateman, P., Gray, P., & Butler, B. (2011). The impact of community commitment on participation in online communities. *Information Systems Research, 22*(4), 841–854.
- Berger, J., Sorensen, A. T., & Rasmussen, S. J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science, 29*(5), 815–827.
- Besley, T., & Burgess, R. (2004). Can labor regulation hinder economic performance? Evidence from India. *The Quarterly Journal of Economics, 119*(1), 91–134.
- Bolton, G. E., Katok, E., & Ockenfels, A. (2004). How effective are electronic reputation mechanisms? An experimental investigation. *Management Science, 50*(11), 1587–1602.
- Burtch, G., Hong, Y., Bapna, R., & Griskevicius, V. (2017). Stimulating online reviews by combining financial incentives and social norms. *Management Science, 64*(5), 2065–2082.
- Chan, J., & Ghose, A. (2014). Internet's dirty secret: Assessing the impact of online intermediaries on HIV transmission. *MIS Quarterly, 38*(4), 955–976.
- Chen, H., Zheng, Z., & Ceran, Y. (2016). De-biasing the reporting bias in social media analytics. *Production and Operations Management, 25*(5), 849–865.
- Chen, P.-Y., Hong, Y., & Liu, Y. (2018). The value of multidimensional rating systems: Evidence from a natural experiment and randomized experiments. *Management Science, 64*(10), 4629–4647.
- Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research, 50*(4), 463–476.
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345–354.
- Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science, 29*(5), 944–957.
- Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems, 23*(2), 149–171.
- Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The operational value of social media information. *Production and Operations Management, 27*(10), 1749–1769.
- Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science, 49*(10), 1407–1424.
- Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing, 21*(4), 23–45.
- Duan, W., Gu, B., & Whinston, A. B. (2008a). Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems, 45*(4), 1007–1016.
- Duan, W., Gu, B., & Whinston, A. B. (2008b). The dynamics of online word-of-mouth and product sales: An empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233–242.
- Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society, 20*(5), 325–344.
- Feng, S., Xing, L., Gogar, A., & Choi, Y. (2012). Distributional footprints of deceptive product reviews. In *Proceedings of the 6th international conference on web and social media* (pp. 98–105). AAAI.
- Geva, T., & Saar-Tsechansky, M. (2021). Who is a better decision maker? Data-driven expert ranking under unobserved quality. *Production and Operations Management, 30*(1), 127–144.
- Goes, P. B., Lin, M., & Au Yeung, C.-m (2014). “Popularity effect” in user-generated content: Evidence from online product reviews. *Information Systems Research, 25*(2), 222–238.
- Gu, B., Park, J., & Konana, P. (2012). The impact of external word-of-mouth sources on retailer sales of high-involvement products. *Information Systems Research, 23*(1), 182–196.
- Guan, X., Wang, Y., Yi, Z., & Chen, Y.-J. (2020). Inducing consumer online reviews via disclosure. *Production and Operations Management, 29*(8), 1956–1971.
- Han, X., Yaraghi, N., & Gopal, R. (2018). Winning at all costs: Analysis of inflation in nursing homes rating system. *Production and Operations Management, 27*(2), 215–233.
- Hansen, M. T., & Haas, M. R. (2001). Competing for attention in knowledge markets: Electronic document dissemination in a management consulting company. *Administrative Science Quarterly, 46*(1), 1–28.
- He, J., Wang, X., Vandenbosch, M. B., & Nault, B. R. (2020). Revealed preference in online reviews: Purchase verification in the tablet market. *Decision Support Systems, 132*, 113281.
- Ho, Y.-C., Wu, J., & Tan, Y. (2017). Disconfirmation effect on online rating behavior: A structural model. *Information Systems Research, 28*(3), 626–642.
- Hong, Y., Huang, N., Burtch, G., & Li, C. (2016). Culture, conformity, and emotional suppression in online reviews. *Journal of the Association for Information Systems, 17*(11), 2.
- Hu, N., Liu, L., & Zhang, J. J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management, 9*(3), 201–214.
- Hu, N., Pavlou, P. A., & Zhang, J. (2017). On self-selection biases in online product reviews. *MIS Quarterly, 41*(2), 449–471.
- Huang, N., Hong, Y., & Burtch, G. (2017). Social network integration and user content generation: Evidence from natural experiments. *MIS Quarterly, 41*(4), 1035–1058.
- Kim, D., Park, K., & Ahn, J.-H. (2018a). Do verified consumer reviews affect sales? An empirical analysis of mixed review systems in the film industry. In *Proceedings of the European conference on information systems (ECIS)*. AIS.
- Kim, S. J., Maslowska, E., & Malthouse, E. C. (2018b). Understanding the effects of different review features on purchase probability. *International Journal of Advertising, 37*(1), 29–53.
- Kim, S.-M., Pantel, P., Chklovski, T., & Pennacchiotti, M. (2006). Automatically assessing review helpfulness. *Proceedings of the 2006 conference on empirical methods in natural language processing* (pp. 423–430). Association for Computational Linguistics, Sydney, Australia.
- Kokkodis, M. (2012). Learning from positive and unlabeled amazon reviews: Towards identifying trustworthy reviewers. *Proceedings of the 21st international conference on world wide web* (pp. 545–546). Association for Computing Machinery, Lyon, France. <https://doi.org/10.1145/2187980.2188119>
- Kokkodis, M., & Ipeiritos, P. G. (2016). Reputation transferability in online labor markets. *Management Science, 62*(6), 1687–1706.
- Kokkodis, M., & Lappas, T. (2020). Your hometown matters: Popularity-difference bias in online reputation platforms. *Information Systems Research, 31*(2), 412–430.
- Kokkodis, M., Lappas, T., & Ransbotham, S. (2020). From lurkers to workers: Predicting voluntary contribution and community welfare. *Information Systems Research, 31*(2), 607–626.
- Kumar, N., Qiu, L., & Kumar, S. (2018a). Exit, voice, and response on digital platforms: An empirical investigation of online management response strategies. *Information Systems Research, 29*(4), 849–870.
- Kumar, N., Venugopal, D., Qiu, L., & Kumar, S. (2018b). Detecting review manipulation on online platforms with hierarchical supervised learning. *Journal of Management Information Systems, 35*(1), 350–380.
- Kumar, N., Venugopal, D., Qiu, L., & Kumar, S. (2019). Detecting anomalous online reviewers: An unsupervised approach using mixture models. *Journal of Management Information Systems, 36*(4), 1313–1346.
- Kumar, S., Mookerjee, V., & Shubham, A. (2018c). Research in operations management and information systems interface. *Production and Operations Management, 27*(11), 1893–1905.
- Lappas, T., Sabnis, G., & Valkanas, G. (2016). The impact of fake reviews on online visibility: A vulnerability assessment of the hotel industry. *Information Systems Research, 27*(4), 940–961.
- Lau, R. Y. K., Zhang, W., & Xu, W. (2018). Parallel aspect-oriented sentiment analysis for sales forecasting with big data. *Production and Operations Management, 27*(10), 1775–1794.

- Lee, S.-Y., Qiu, L., & Whinston, A. (2018). Sentiment manipulation in online platforms: An analysis of movie tweets. *Production and Operations Management*, 27(3), 393–416.
- Li, H., Chen, Z., Liu, B., Wei, X., & Shao, J. (2014). Spotting fake reviews via collective positive-unlabeled learning. In *International conference on data mining* (pp. 899–904). IEEE.
- Li, X., Hitt, L. M., & Zhang, Z. J. (2011). Product reviews and competition in markets for repeat purchase products. *Journal of Management Information Systems*, 27(4), 9–42.
- Liu, Q. B., & Karahanna, E. (2017). The dark side of reviews: The swaying effects of online product reviews on attribute preference construction. *MIS Quarterly*, 41(2), 427–448.
- LIWC (2018). *Linguistic inquiry and word count*. <https://www.liwc.app/>
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62(12), 3412–3427.
- Mallipeddi, R. R., Janakiraman, R., Kumar, S., & Gupta, S. (2021). The effects of social media content created by human brands on engagement: Evidence from Indian general election 2014. *Information Systems Research*, 32(1), 212–237.
- Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8), 2421–55.
- McCallum, A., & Nigam, K. (1998). A comparison of event models for naive Bayes text classification. In *Workshop on learning for text categorization* (vol. 752, pp. 41–48). AAAI.
- Mejia, J., Mankad, S., & Anandasivam, G. (2021). Service quality using text mining: Measurement and consequences. *Manufacturing & Service Operations Management*, 23(6), 1354–1372.
- Moe, W. W., & Trusov, M. (2011). The value of social dynamics in online product ratings forums. *Journal of Marketing Research*, 48(3), 444–456.
- Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on amazon. com. *MIS Quarterly*, 34(1), 185–200.
- Mukherjee, A., Venkataraman, V., Liu, B., & Gance, N. (2013). *Fake review detection: Classification and analysis of real and pseudo reviews*. Tech. Rep. UIC-CS-2013–03, University of Illinois at Chicago.
- Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7(3), 101–134.
- Pavlou, P. A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research*, 15(1), 37–59.
- Reinstein, D. A., & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The Journal of Industrial Economics*, 53(1), 27–51.
- Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *MIS Quarterly*, 36(3), 841–864.
- Saxena, A. (2018). *Deception detection on Amazon reviews dataset*. <https://bit.ly/2W5Cw15>
- Schneiderman (2014). *Attorney General Schneiderman announces agreement with 19 companies to stop writing fake online reviews and pay more than \$350,000 in fines*. <https://on.ny.gov/38L4a6t>
- Scrapehero (2018). *How many products does Amazon sell?* <https://bit.ly/39GzDI8>
- Smith, D. (2004). *Amazon reviewers brought to book*. The Guardian. <https://www.theguardian.com/technology/2004/feb/15/books.booksnews>
- Sridhar, S., & Srinivasan, R. (2012). Social influence effects in online product ratings. *Journal of Marketing*, 76(5), 70–88.
- Statista (2018). *Quarterly share of e-commerce sales of total U.S. retail sales from 1st quarter 2010 to 2nd quarter 2018*. <https://bit.ly/2Q1cCHR>
- Sussin, J., & Thompson, E. (2012). *The consequences of fake fans, “likes” and reviews on social networks*. <https://gtrn.it/2lyOUP5>
- Wang, Y., Goes, P., Wei, Z., & Zeng, D. (2019). Production of Online word-of-mouth: Peer effects and the moderation of user characteristics. *Production and Operations Management*, 28(7), 1621–1640.
- Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 29(1), 35–57.
- Wehner, M. (2016). *Amazon deletes over 500,000 sketchy reviews in massive cleanup effort*. BGR. https://finance.yahoo.com/news/amazon-deletes-over-500-000-sketchy-reviews-massive-211241936.html?guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAH0A3u-uuImj5Tn5DNG01zalj79COd9rQyiPgVDxBUxY30JdJDPt06J_FHcMrjajvEMwQtMtmJ5JX6E6GE0_YRhxuMGkYYrmoi78YwaLRVAum-Vcw7ddo_Qmc2aegQaR3t0YhLakqSdcJulg_ZA93ig9vKf62rOmQ4w6uYAa6aWl
- Wooldridge, M. J. (2010). *Econometric analysis of cross section and panel data*. MIT Press.
- Wu, C., Che, H., Chan, T. Y., & Lu, X. (2015). The economic value of online reviews. *Marketing Science*, 34(5), 739–754.
- Yin, D., Bond, S., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, 38(2), 539–560.
- Yuan, W., Hong, Y., & Pavlou, P. (2012). Do consumers trust online product reviews? An experimental study of biases in online product reviews. In *Americas conference on information systems*. AMCIS 2012 in Seattle.
- Zhang, D., Zhou, L., Kehoe, J. L., & Kilic, I. Y. (2016). What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. *Journal of Management Information Systems*, 33(2), 456–481.
- Zhao, Y., Yang, S., Narayan, V., & Zhao, Y. (2013). Modeling consumer learning from online product reviews. *Marketing Science*, 32(1), 153–169.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kokkodis, M., Lappas, T., & Kane, G. C. (2022). Optional purchase verification in e-commerce platforms: More representative product ratings and higher quality reviews. *Production and Operations Management*, 31, 2943–2961. <https://doi.org/10.1111/poms.13731>