JOURNAL OF MANAGEMENT INFORMATION SYSTEMS
2026, VOL. 43, NO. 2, 654–679
https://doi.org/10.1080/07421222.2026.2647471

Routledge
Taylor & Francis Group

Fewer Reviews but Better Content: When an Online Review Platform Disables Downvotes

Yinan Yu[a], Dezhi Yin[b], and Warut Khern-am-nuai[c]

[a]University of Oklahoma, Norman, OK, USA; [b]University of South Florida, Tampa, FL, USA; [c]McGill University, Montreal, QC, Canada

ABSTRACT

Most user-generated content platforms utilize peer evaluation systems that incorporate upvotes and downvotes to distinguish high-quality from low-quality content. However, downvotes are often misused to target others and discourage content creators, and some platforms, such as Amazon and TripAdvisor, have removed the downvote option altogether. Meanwhile, the real-world impact of this platform-level intervention on user contributions remains underexplored, and we fill this gap by analyzing a restaurant review platform that disabled downvotes. We apply the regression discontinuity in time (RDiT) method and find that removing downvotes reduces the quantity of reviews but increases their quality. Further mechanism analysis suggests that the former may stem from perceived unfairness, and the latter from a motivation of continuing reviewers to stand out. Additionally, our heterogeneity analysis shows that the decline in review quantity is stronger among long-tenure reviewers, whereas the improvements in multiple quality dimensions are concentrated among reviewers with fewer followers. Our study reveals the nuanced behavioral effects of a platform-level design change and offers practical insights for designing peer evaluation systems that balance fairness, motivation, and content value.

KEYWORDS

User-generated content; online reviews; peer evaluation; downvote removal; review quantity; review quality; fairness perception; attention-seeking behavior

Introduction

User-generated content (UGC) platforms commonly deploy peer evaluation systems that enable content consumers to upvote and downvote different pieces of content. By relying on "the wisdom of crowds" [49], these voting systems allow a platform to assess the quality of the vast amount of UGC and distinguish helpful pieces of content from unhelpful ones. This distinction is valuable for content consumers who prefer higher-quality content amid information overload. In particular, the option of casting unhelpful votes provides an effective avenue for identifying and demoting low-quality or fake content.

At the same time, however, this option can be abused by fellow content creators and UGC-related stakeholders to benefit themselves. In the context of our interest, online review platforms, malicious reviewers intending to get ahead (becoming a top reviewer) can downvote the helpfulness of others' reviews, and malicious businesses can downvote their competitors' positive reviews or authentic top reviews.F[1] In

CONTACT Yinan Yu yinan.yu@ou.edu University of Oklahoma, 307 West Brooks, Norman, OK 73019, USA
Supplemental data for this article can be accessed at www.tandfonline.com/mmis.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/07421222.2026.2647471

addition, many reviewers feel frustrated and demotivated after receiving unhelpful votes for their reviews, leading to a decline in their content contribution.1[2] This issue prompted a debate among researchers and practitioners regarding the potential implications of the negative evaluation option on online platforms. However, while several online platforms, such as Amazon and TripAdvisor, have preemptively removed the downvote option from their system in recent years, empirical evidence regarding the implications of such an important platform-level design choice is still limited in the literature. As such, platform managers are left in the dark on the appropriate and sustainable design of the peer evaluation system where the platform's welfare is at stake.

While the removal of downvotes may help address abuse of review evaluation systems, it can also produce unintended consequences for user-generated content. A growing body of research has examined factors that influence content generation on UGC platforms, including how *receipt* of positive or negative peer evaluations affects the behavior of feedback recipients, and how changes in the *visibility* of such feedback alter participation (e.g., [15, 27, 52, 63]). However, this work largely focuses on settings where the negative evaluation option remains available. Little research has empirically investigated the impact of platform-level design choices that eliminate this option altogether.

We address this gap by examining a major platform-level design change (i.e., the removal of the downvote option) and asking how this change affects two core outcomes of user-generated content: content quantity and multi-dimensional quality, both of which have direct implications for platform operations. Removing the downvote option at the platform level is not simply the opposite of receiving downvotes at a user level, and it also differs from merely manipulating the visibility of downvote counts. Instead, it represents a structural change that eliminates the *existence/capacity* for negative evaluation altogether and reconfigures the evaluation environment. By isolating this structural change and exploring its impacts on both content quantity and quality in a field setting, we clarify users' behavioral responses that platforms may anticipate when redesigning peer-evaluation systems.

Furthermore, this design change activates competing theoretical mechanisms. On the one hand, eliminating the downvote option could reduce reviewers' fear of negative evaluation and associated social anxiety [29], thus encouraging content contributions. On the other hand, the absence of the downvote option could create perceptions of procedural unfairness by preventing users from providing negative evaluations, thus discouraging contributions [1]. Because existing studies focus on individual-level feedback, their findings cannot be extended to predict the net effect of a platform-level design change. Our study addresses this gap by empirically testing the impact of downvote removal using a unique dataset.

The empirical context of our study is an online review platform. We collaborated with a prominent online platform in Asia to obtain a unique dataset, which allows us to investigate the impact of removing the downvote option at the platform level. Specifically, we exploit a natural experiment on a review platform that initially permitted both upvotes and downvotes but abruptly deactivated the downvote option on September 1, 2016. The platform provided us with a comprehensive dataset encompassing the information of each review and its reviewer from June to November 2016, allowing us to estimate the effects of removing the downvote option on the quantity and quality of reviews. Our analysis uses the regression

discontinuity in time (RDiT) technique as the primary identification strategy, with a difference-in-differences (DiD) design as a supplementary approach primarily for examining mechanisms and heterogeneous effects.

We find that the elimination of the downvote option is associated with a decrease in the quantity of reviews submitted by existing users, but an improvement in various review quality measures. We conducted additional analyses to better understand the dominant mechanisms underlying these results. Our findings suggest that removing the downvote option may lead reviewers to perceive the platform as less fair, which in turn decreases overall review contributions. With only upvotes remaining, writing reviews on the platform has inadvertently become a "popularity contest." Therefore, those who continue contributing tend to produce more detailed, diverse, and novel reviews to stand out. We also uncover meaningful heterogeneity in these effects: the decline in review quantity is stronger among long-tenure reviewers, while the improvements in effort-intensive dimensions of quality are concentrated among reviewers with fewer followers.

This work is among the first to empirically document the nuanced impacts of a critical design change to the review evaluation system on reviewers' content generation behavior. Our results complement the recent research on the effects of peer evaluation on UGC production. As highlighted by recent editorials [20-22], the key strength of our study lies in its phenomenon-driven orientation and the unique dataset, which enables us to address a distinctive yet important research question. Particularly, previous research on the effects of receiving downvotes or manipulating its visibility has relied on data where the downvote option was present [9, 15, 25, 62], limiting their ability to infer how user behavior may change when a platform implements a significant design change to remove this option. Importantly, there are fundamental differences in what users respond to in these two scenarios: the presence of negative peer evaluations versus the *ability* to leave a negative evaluation. Our dataset uniquely addresses this gap, offering valuable insights into the effects of removing the downvote option. Conceptually, we distinguish three design dimensions examined in prior literature: (i) the receipt of negative feedback at the user level, (ii) the visibility of negative feedback (counts or salience), and (iii) the existence or capacity of negative evaluation. Our context falls into the third category, providing novel insights into how structural changes to peer evaluation systems shape content creation dynamics.

Our findings carry important practical implications for platforms seeking to optimize their peer-evaluation systems. Although intuition might suggest that removing the downvote option would reduce contributors' fear of negative feedback and thereby increase participation, we find the opposite: review volume declines. At the same time, the reviews that remain become longer, more diverse, and more novel, indicating a quality gain. This quantity–quality trade-off matters because fewer reviews can reduce coverage and diversity of perspectives, and the removal of downvotes also eliminates a diagnostic crowd-sourced signal of usefulness, potentially lowering the platform's overall informativeness. Managers considering similar interventions should therefore weigh not only the direct behavioral effects but also how complementary options, such as badges, peer awards, expert curation, or reputation scores, might preserve fairness and evaluative signals while still protecting contributors from abusive or malicious voting.

Literature Review

Social and Peer-Evaluation Drivers of UGC Generation

Numerous studies have examined a wide range of factors influencing content generation on UGC platforms, including extrinsic incentives such as monetary rewards (e.g., [7]) and social-environmental factors such as social network and evaluation contexts. Because our study centers on peer-based rather than financial mechanisms, we next focus on the literature most directly relevant to social and peer-evaluation factors.

Early work shows that exposure to others' ratings and opinions systematically alters a focal reviewer's decisions about whether and how to contribute, documenting phenomena such as self-presentation, herding, and strategic conformity [31, 32]. Later studies based on structural and econometric models deepen this view. For example, Ho et al. [25] develop a Bayesian learning framework of expectation–disconfirmation and show that the discrepancy between expected and experienced quality predicts both the likelihood of posting and the polarity of the rating, with occasional posters being especially susceptible to peer signals. This stream collectively underscores that peer feedback mechanisms, not just intrinsic motivation, are critical governance devices on UGC platforms.

Complementing this line of research, studies also show that social technologies and network structures can meaningfully shape content generation. Platforms that enable "following" relationships facilitate social diffusion and visibility of contributions. Increases in a user's follower base tend to be associated with greater contribution quantity and sometimes quality [19, 34, 41, 56]. Furthermore, Wei et al. [54] even uncover an asymmetric effect: a larger follower network boosts the quantity of content generation, whereas a larger followee network dampens it. Together, these studies highlight that the broader social environment shapes users' content generation behavior.

Another important factor that influences users' content generation behavior is feedback from managers and other users. Existing studies suggest that managerial response to customer reviews generally increases review quantity, but its stimulating effects on different types of reviews are more nuanced. For example, Proserpio and Zervas [40] find that hotel management responses lead to more reviews on average, but they also result in fewer (but longer) negative reviews. For hotels that adopt managerial responses, Chevalier et al. [11] observe an increase in review quantity and length, but a decrease in rating. Furthermore, Chen et al. [10] show managerial responses create externalities by encouraging more customers to post future reviews while also lowering subsequent ratings, highlighting the broader systemic impact of managerial engagement on content generation.

Our work contributes to this stream of literature by examining the impact of a social factor, particularly a platform design change, that affects the content-generating behavior of contributors on an online review platform. Next, we survey prior works that specifically study peer evaluation systems on UGC platforms.

Peer Evaluation Systems on UGC Platforms

The second pertinent literature relevant to our study concerns peer evaluation within UGC platforms, predominantly in the form of voting systems. The first sub-stream of this literature has scrutinized a set of determinants that shape peer evaluations. This includes examining the intrinsic attributes of the content, such as the volume of information,

linguistic characteristics like readability, and the emotions contained within the content (e.g., [28, 37, 57]). For example, the seminal work of Mudambi and Schuff [37] shows that the amount of information (operationalized as the length of review content) has a positive effect on review helpfulness. Another set of studies focuses on the attributes of the content creators, demonstrating that content creators' expertise, credibility, and identity disclosure are positively associated with the content evaluation [3, 13, 17]. A further subset of research has probed into how various social dynamics, such as social influence and social interaction [36, 58], may distort the peer evaluations.

Another sub-stream closely related to our work focuses on how content evaluations influence the content-generating behavior of those evaluation subjects [8, 27, 50, 52, 62, 63]. For example, Burtch et al. [8] find that positive recognition in the form of a peer award leads to longer and more frequent content generation on Reddit. Zhu et al. [63] empirically study the changes in content contribution behavior of users who receive different types of peer evaluation votes. They find that compared with receiving no feedback at all, receiving negative feedback improves review quantity and quality. Turel and Qahri-Saremi [52] investigated how receiving likes and dislikes affects individuals' own risky online behaviors. Collectively, these studies investigate how receiving peer feedback shapes user behavior. Adjacent work has examined the impact of visibility. For instance, Kim et al. [27] explore how manipulating the public visibility of dislike counts affects both viewer and creator engagement.

Our work complements and extends this stream of literature, especially the latter sub-stream of research studying how upvotes or downvotes impact the behavior of evaluation subjects. Specifically, our work explores how the platform-level change in the peer evaluation system, where the downvote option is completely removed, impacts the review-generating behavior. The collective results from our work and prior literature offer a holistic view of how the peer evaluation system impacts relevant stakeholders, allowing platform managers to develop policies and interventions to further enhance the viability and sustainability of the UGC platform.

Hypothesis Development

The removal of the downvote option on consumer review sites may reduce reviewers' fear of negative evaluations, but it may alternatively influence reviewers' perception of the voting system. These different mechanisms predict diverging outcomes, as we elaborate below.

First, the downvote option is directly associated with negative evaluations and critiques that can cause fear. According to the Fear of Negative Evaluation (FNE) model, FNE or fear of failure refers to one's apprehension or distress over the possibility of others' negative evaluations [53]. People high on FNE are seeking others' approval and avoiding disapproval. As a result, they are more likely to avoid the possibility of being negatively evaluated [29]. Whereas FNE is conceptualized as a personal trait, additional research has revealed compelling evidence for the close relation of situation-activated fear or anxiety with avoidance behaviors [60].

At the review platforms where both upvote and downvote options are available, reviewers are likely to be negatively evaluated, and they may experience fear or social anxiety. With the elimination of the downvote feature, reviewers can no longer be criticized, and they are expected to experience a reduced sense of concern regarding the potential for

negative evaluations. Accordingly, reviewers should be more likely to engage in review-writing activities and contribute more reviews. In addition, as reviewers are relieved of the fear of negative evaluations and the accompanying social scrutiny or disapproval, they are less likely to hold their writing to a high standard and may instead exhibit a laxer approach to review writing. Taken together, we propose the first set of hypotheses based on FNE.

Hypothesis 1A: *The removal of the downvote option leads reviewers to write more reviews.*

Hypothesis 2A: *The removal of the downvote option leads reviewers to write lower-quality reviews.*

Other than personal fear of others' evaluations, reviewers' contribution behavior may be alternatively influenced by their moral evaluation of the entire review system. Moral judgments are emotional and intuitive in nature, and they can directly and "automatically" drive actions [23]. In particular, the (good or bad) valence of affect-laden intuitions can lead people to approach or avoid a situation accordingly [48].

In the context of online reviews, perceived fairness is a critical form of moral evaluation that has been explored at the individual review level. For example, Allard et al. [1] find that when consumers perceive negative reviews as unfair, they act to restore fairness by increasing their purchase intention to support the firm. Extending this concept, we shift the focus from the perceived fairness of individual reviews to the perceived fairness of the platform, and specifically examine how the removal of the downvote option can influence the peer evaluation process (and reviewers' reactions).

Organizational justice theory provides a relevant framework by categorizing perceived fairness into three dimensions: distributive fairness (i.e., fairness of the received outcomes), procedural fairness (i.e., fairness of the procedures used to determine those outcomes), and interactional fairness (i.e., fairness of interpersonal treatment during the enactment of procedures) [12]. This framework has been utilized in the context of online reviews. For instance, Ravichandran and Deng [43] study how managers should respond to negative reviews that highlight those three types of unfairness related to the product and service. In our study, we focus on perceived procedural fairness within the peer evaluation system, as the removal of the downvote option represents a procedural change in how reviews are evaluated. When the downvotes option is removed, the review voting system is no longer capable of separating good reviews from bad ones because negative evaluations are no longer possible. Thus, reviewers may intuitively feel that the voting system has been compromised by being one-sided and unfair from a procedural standpoint. Perceptions of procedural injustice have been shown to activate people's negative emotions [38] and reduce people's intrinsic motivation toward a given task [59]. Driven by this negative intuitive feeling and lowered motivation, reviewers should have an avoidance tendency, and they are less likely to contribute reviews in such an unjust system. Thus, we propose a competing hypothesis (compared to H1A).

Hypothesis 1B: *The removal of the downvote option leads reviewers to write fewer reviews.*

Next, we argue that the removal of the downvote option may also influence review quality through a distinct mechanism. Once the platform disables the downvote option, the review environment would shift toward a "popularity contest," in which reviewers compete

primarily for favorable attention from readers rather than responding to both positive and negative feedback. Because the downvote option is no longer available, the voting system also becomes less diagnostic (i.e., users lose a quick signal of which reviews are relatively weak) so reviewers seeking visibility must rely more heavily on the intrinsic appeal of their own content to stand out [18, 46]. This change raises the incentives to invest greater effort per review, for example, by writing longer and more detailed text, including more photos to increase multi-modality, and expanding topic coverage or offering novel insights to appeal to diverse reader interests.

Viewed through a complementary resource-allocation lens, the removal of downvotes reduces the diagnosticity of crowd feedback, under limited time and cognitive resources, contributors may reallocate effort from posting a greater number of reviews to enriching each review to maintain visibility and impact [47]. Together, these dynamics form the basis for a competing hypothesis to H2A.

Hypothesis 2B: *The removal of the downvote option leads reviewers to write higher-quality reviews.*

Context, Research Design, and Data

Research Context and Research Design

To address our research question and empirically test the competing hypotheses, we have partnered with a leading restaurant review platform in Asia. Since its establishment in 2010, this platform has become a prominent source for restaurant reviews in its respective country. It has featured over 350,000 restaurants and accumulated a user base of over three million registered users. Its wide reach and thorough coverage make it a standout and typical example among restaurant review sites, much like the popular global review platforms, such as Yelp and Google Reviews. This platform enables registered users to post textual reviews, assign ratings on a scale of 1 to 5 stars, and attach photos of food and restaurants. The platform also employs a review helpfulness voting system to manage the issue of information overload. This system enables review readers to evaluate the helpfulness of reviews by casting votes that are either helpful or unhelpful. The platform then aggregates these votes and calculates a helpfulness ratio, which users can utilize to sort and select top-rated reviews to read.

On September 1, 2016, the platform implemented a change to the review evaluation system without prior notification, by disabling the option for users to cast unhelpful votes and leaving only the option to cast helpful votes. This decision was driven by a strategic intent to curb the misuse of unhelpful votes, which some individuals exploited to target and undermine others' reviews inappropriately and to manipulate the review ecosystem. By retaining only the option for users to vote reviews as helpful, the platform aimed to foster a more supportive and constructive environment for feedback. The removal of the downvote option represents an inter-temporal change that provides us with a quasi-experimental setting where the system change is considered a treatment. Meanwhile, the platform confirms that there were no system updates, feature implementation, or platform-wide campaign around the same time.

Data and Variables

Our primary dataset, obtained through a collaboration with the platform, includes all the reviews posted between June 2016 and November 2016 (i.e., three months before and three months after the policy change). The dataset contains information such as the review text, rating, number of attached photos, and writing date, among others. Additionally, it includes the characteristics of the reviewers, such as their join date and gender. To ensure tracking of how reviewers' behavior was affected by the change in the voting system, we focused on reviewers who wrote at least one review during the three-month pre-treatment period. In total, our dataset consists of 3,251 reviewers.

Our research question focuses on studying both the quantity and quality aspects of review writing behavior. To measure the quantity of reviews, we track the number of reviews written by reviewers. To assess review quality, we utilize multiple measures. First, we proxy for review depth and richness using the length of the review (e.g., [37]) and the number of photos attached to the review [55], respectively. Both length and photo inclusion also serve as effort-related indicators of the reviewer's investment in producing higher-quality content.

Apart from these quantitative measures, we utilize several natural language processing techniques to extract semantics from the review texts. Specifically, we calculate the Gunning-Fog (GF) index, which estimates the years of formal education a person needs to understand the text on the first reading, to gauge a review's reading difficulty (e.g., [19]). A higher GF score indicates greater linguistic complexity and lower readability, meaning that more education is required to comprehend the text (i.e., clarity-related measure).

Furthermore, we construct measures to capture reviews' diversity and novelty (i.e., breadth-related measures). We perform the Latent Dirichlet Allocation (LDA) model to uncover latent topics discussed in the review content [4]. LDA is an unsupervised clustering technique, designed to discern abstract topics across a document collection and subsequently generate a specified number of these topics. Within the LDA framework, each review is conceptualized as a composite of multiple topics, with each topic characterized by a specific term distribution. For our textual analysis, we utilized the *scikit-learn* package for Python [39]. We execute the model with configurations of 3 to 10 topics and inspect the term distributions. An optimal configuration is identified with 8 topics, as evidenced by the minimal perplexity, which is a metric that measures the efficacy of topic models [4]. We manually inspect the topics generated by LDA and assign labels to them. Table 1 lists the top ten terms of the eight topics labeled by us.

After obtaining topic distribution for each review, we calculate the Gini coefficient to measure review diversity (e.g., [6]). This coefficient quantifies the heterogeneity in topics addressed within a review. Its value ranges between 0 and 1. A score of 0 indicates perfect equality, implying that a review discusses various topics with a balanced emphasis across a variety of topics, thereby reflecting the maximum level of diversity in our analysis. Conversely, a score of 1 indicates perfect inequality, denoting that a review exclusively focuses on a singular topic, which corresponds to the minimal level of diversity.

Finally, to evaluate a review's level of novelty, or conversely, its similarity to other reviews, we compute the cosine similarity between the focal review and previously posted reviews for the same restaurant [16]. First, we calculate the restaurant's topic vector by averaging the topic distributions of reviews posted within the 30 days

Table 1. Top terms of topics in reviews.

Topic	Top Terms
1. Price and menu	thi eat price veri good set salmon menu food
2. Atmosphere	shop food good veri thi restaur atmospher come sit price
3. Service	shop thi wa becaus order quot time like wait staff
4. Noodle and rice dishes	pork shop noodl thi eat rice restaur delicious tast soup
5. Dessert and drink	shop ice cream tea sweet chocol coffe veri cake menu
6. Taste	eat taste time like veri order delicious come shop tri
7. Western cuisine	thi eat pizza like spaghetti sauc steak delicious salad chees
8. Seafood cuisine	fri veri fish thi delicious food shrimp rice salad chicken

Note: Basic special character replacement and stemming are applied to the raw text before performing Latent Dirichlet Allocation.

Table 2. Summary statistics.

	Mean	SD	Min	Max
$ReviewCnt_{it}$	4.62	7.03	1.00	131
$Length_{it}$	531.84	589.46	2.00	10,205
$Photo_{it}$	5.53	4.77	0	85
$Readability_{it}$	8.57	2.87	0.40	66.14
$Concreteness_{it}$	2.81	0.19	1.64	4.92
$TopicGini_{it}$	0.66	0.23	0.01	1.00
$Similarity_{it}$	0.72	0.08	0.10	0.87

Notes: 1) we excluded weeks during which no reviews were written since we were unable to calculate the review quality measure, 2) the number of observations is 15,211.

preceding the focal review's publication. Next, we compute the cosine of the angle between the focal review's topic vector and the restaurant's topic vector. This cosine similarity score, which ranges from 0 to 1, represents the degree of similarity between the two vectors in terms of the topics covered. A higher score indicates greater similarity and, therefore, a less novel focal review.

We create a reviewer weekly panel dataset that spans 12 weeks prior to and following the removal of the downvote option. We calculate the weekly averages for each variable, corresponding to each of the reviewers. The summary statistics are shown in Table 2.

Empirical Analyses

As our analyses rely on observational data, they are naturally subject to endogeneity concerns. To address this, we employ a RDiT design as the primary identification strategy, exploiting the sharp timing of the platform policy change to compare outcomes immediately before and after the cutoff. RDiT is widely recognized for its strong internal validity in drawing causal inferences from external shocks such as policy changes [24], making it well-suited for our setting. To further probe mechanisms and heterogeneous impacts, we complement the RDiT with a supplementary DiD analysis on a reviewer-weekly panel, including moderating effect tests to explore how reviewers' behaviors change and who changes.

Main Analysis: RDiT

In this subsection, we adopt the RDiT research design [53]. This approach is recognized within the academic community as a rigorous identification strategy that facilitates robust causal inferences in quasi-experimental settings where the treatment is administered over a temporal dimension [14, 35]. This strategy is pertinent to our empirical context and therefore offers a sound methodological foundation for examining the causal effects of the downvote option removal.

RDiT builds on the fundamental principle of the regression discontinuity research design and utilizes time as the running variable (i.e., the continuous variable along which observations are ordered to determine which units fall just before or just after the treatment threshold). The underlying rationale is that the observations immediately preceding the temporal cutoff can serve as suitable counterfactuals for those immediately following the cutoff. In our context, the RDiT approach estimates the counterfactual review writing behavior in the absence of the downvote option removal based on a narrow time window before the system design change. Then, by calculating the average differences in behavior between the pre- and post-treatment periods around the change, we can obtain an estimate of the average treatment effect of removing the downvote option. The data used in this analysis is organized on a daily basis.

We implement the augmented local linear version of RDiT in two steps [24]. First, we regress each outcome on a set of controls using the entire study period and save the residuals. Second, within a narrow symmetric window around the policy change, we estimate separate local linear regressions of these residuals on time for the pre- and post-policy periods and compute the jump at the cutoff. This augmentation improves the bias–variance trade-off by separating control adjustment from the local fit. Rather than estimating the full model within the narrow window (which can be noisy), we first partial out a set of controls over the entire study period and then apply the local linear regression to the residuals. This two-step procedure removes slow-moving differences unrelated to the policy change and yields a cleaner, more precise estimate of the discontinuity at Day 0. Figure 1.1 in Online Supplemental Appendix 1 visually illustrates this approach. Specifically, we use the following augmented local linear specification:

$$Y_t = \alpha_0 + \alpha_1 DownvoteRemoval_t + \alpha_2 t + \alpha_3 X_t + \varepsilon_t \quad (1)$$

where Y_t indicates dependent variables that capture the quantity and quality of reviews. The dummy variable $DownvoteRemove_t$ is equal to 1 for the time period after the platform disabled the downvote option, and 0 otherwise. t is a linear time trend. X_t denotes a vector of control variables, including indicator variables for the day of the week and month dummies. α_1 is the key coefficient of interest in this specification, which shows the effect of downvote option removal on review writing behavior. To ensure the robustness of the results, we estimate Eq. (1) with bandwidths of ±50, ±60, and ±70 days, respectively.

Table 3 reports the augmented local linear RDiT estimates. Across bandwidths of ±50, ±60, and ±70 days, the results are highly consistent in both sign and magnitude. We present the ±60-day specification as the preferred estimate because it strikes a balance between local validity and sample size. The estimates in Column (1) show that the coefficient on $Log(ReviewCnt_{it})$ is significantly negative across all bandwidths, indicating that the removal of the downvote option reduced review frequency. At the ±60-day bandwidth, the point

Table 3. Effect of downvote option removal on review writing behavior (RDiT).

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	$Ln(ReviewCnt_t)$	$Ln(Length_t)$	$Ln(Photo_t)$	$Readability_t$	$TopicGini_t$	$Similarity_t$
			50 days			
	-0.461*	0.159**	0.234**	-0.028	-0.009**	-0.014*
	(0.128)	(0.032)	(0.043)	(0.088)	(0.002)	(0.007)
			60 days			
$DownvoteRemoval_t$	-0.587**	0.175**	0.261**	-0.054	-0.010**	-0.013*
	(0.120)	(0.031)	(0.040)	(0.081)	(0.002)	(0.006)
			70 days			
	-0.604**	0.195**	0.273**	-0.239**	-0.009**	-0.009
	(0.116)	(0.030)	(0.041)	(0.093)	(0.002)	(0.006)

Notes: We report the coefficients of $DownvoteRemoval_t$ from eighteen regressions in this table. Standard errors in parentheses are robust and clustered by reviewers.
**$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

estimate of –0.587 implies a decline of roughly 44 percent ($(e^{-0.587}-1)\times100$ percent) in review quantity relative to the pre-policy baseline. This finding supports H1B.

Turning to review quality, Columns (2) and (3) show positive and statistically significant effects on both $Log(Length_{it})$ and $Log(Photo_{it})$. The estimates suggest that after the downvote removal, reviews became about 19.1 percent ($(e^{0.175}-1)\times100$ percent) longer and contained 29.8 percent ($(e^{0.261}-1)\times100$ percent) more photos, reflecting significant improvements in these effort-related quality measures. Column (4) shows negative coefficients for reading difficulty, significant only at the ±70-day window, suggesting a shift toward simpler, easier-to-understand language when reviewers adjusted their style. Finally, the coefficients in Columns (5) and (6) are negative, indicating greater topical diversity and novelty in reviews. Specifically, the downvote removal lowered the Topic Gini index by about 0.010 and cosine similarity by about 0.013. A lower Topic Gini index suggests that reviewers' discussions became more evenly distributed across topics rather than concentrated on a few recurring themes, while a decline in cosine similarity implies that individual reviews became less textually similar to existing ones. Together, these results indicate that reviewers produced more diverse and distinctive content, reflecting improvements in breadth-related quality measures.

In sum, while the downvote option's removal led to fewer reviews from existing reviewers overall, the remaining reviews became longer, richer, and more varied in content. These results provide strong support for H2B, which posits that reviewers who continue to post after the policy change produce higher-quality reviews.

To provide intuitive, visual evidence of the treatment effects identified by our augmented local linear RDiT analysis, we further visualize the estimates in Figure 1 using the actual review data. Each panel plots binned averages of residualized outcomes along with local linear fits on either side of the policy cutoff, allowing readers to directly observe the discontinuity associated with the removal of the downvote option. For review quantity, a clear downward discontinuity is observed at the policy change, consistent with a decline in participation following the removal of downvotes. For review quality, heterogeneous patterns emerge across dimensions: review length and photo posting exhibit upward shifts at the cutoff, whereas Topic Gini and content similarity display downward shifts, indicating greater topical breadth and differentiation. Together, these figures provide visual evidence

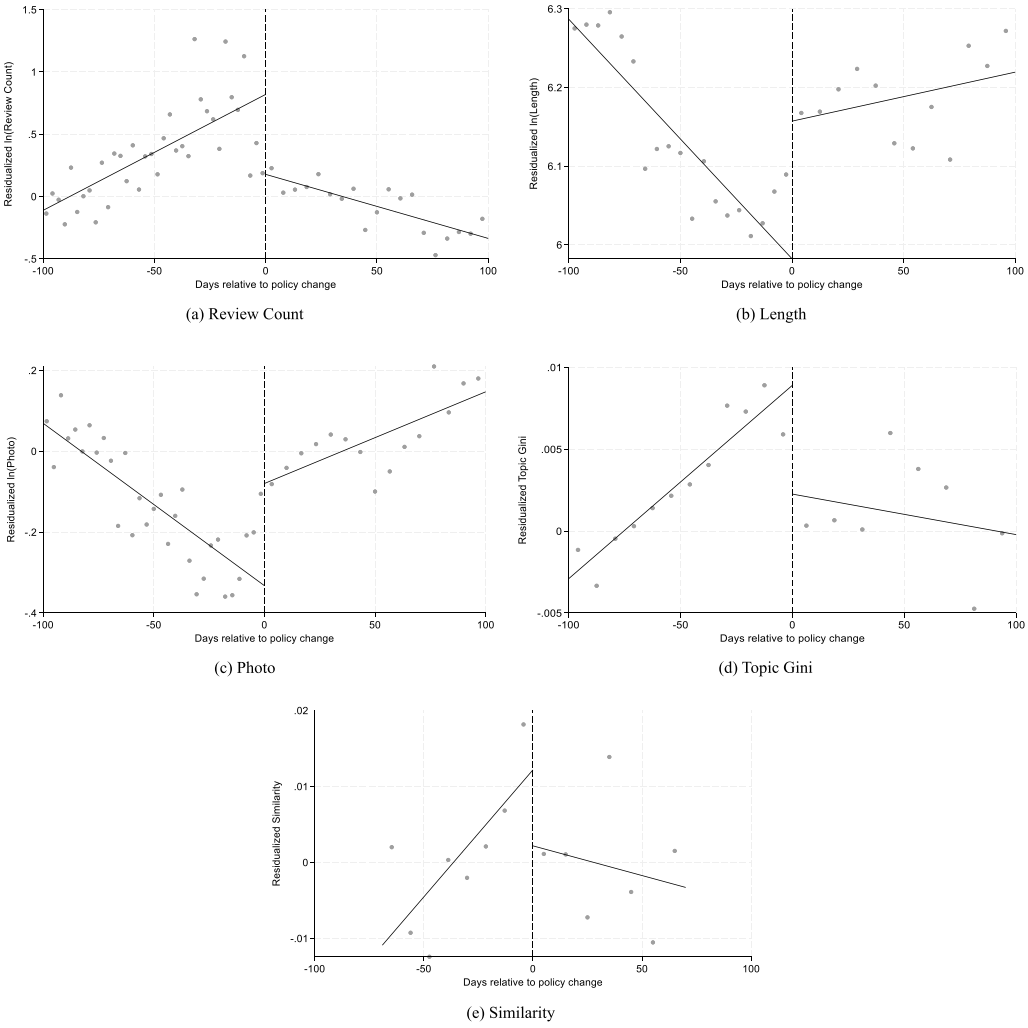

Figure 1. Augmented local linear RDiT: Visual evidence on review quantity and quality.

that the platform intervention reduced review quantity while simultaneously increasing review quality along multiple dimensions, consistent with the regression results.

While our main analysis focuses on existing reviewers who contributed in both the pre- and post-policy periods, we conduct a robustness check by extending the RDiT analysis to all reviewers, including new entrants. The results, reported in Table 1.1 of Online Supplemental Appendix 1, are statistically the same as our main estimates, indicating that our conclusions are not driven by compositional changes in the reviewer pool.

To further validate our RDiT estimates, we implement a donut specification. Although reviewers cannot directly "sort" on time (i.e., they cannot deliberately "choose" which side of the cutoff they fall on), behavior immediately around the cutoff may be temporarily noisy. For example, some users anticipated the change or needed time to adjust to the new interface. The donut approach mitigates this concern by removing a narrow window of observations around the cutoff and comparing outcomes in the more stable periods just

Table 4. Effect of downvote option removal on review writing behavior (donut RDiT).

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	$Ln(ReviewCnt_t)$	$Ln(Length_t)$	$Ln(Photo_i)$	$Readability_t$	$TopicGini_t$	$Similarity_t$
$DownvoteRemoval_t$	-0.953***	0.246***	0.426***	-0.154	0.050***	-0.015*
	(0.149)	(0.041)	(0.048)	(0.094)	(0.012)	(0.008)

Note: Estimates use a ±60-day bandwidth, excluding the 7 days on either side of the policy change. Standard errors in parentheses are robust and clustered by reviewers.
***$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

before and after the design change. Specifically, we exclude the 7 days on either side of the cutoff and re-estimate the model. The results shown in Table 4 remain statistically the same with our main RDiT estimates: the removal of the downvote option reduces review quantity while improving multiple dimensions of review quality.

Supplementary Analysis: DiD

While our main identification relies on RDiT, we also conduct a supplementary DiD analysis to examine mechanisms and heterogeneous effects. In this analysis, we organize outcomes at the reviewer-weekly level to smooth day-of-week seasonality, reduce zero inflation in reviewer-day observations, and yield enough within-reviewer variation to estimate interactions with precision. This DiD analysis serves as a complementary design that helps reveal *how* reviewers' behaviors change and *who* changes.

Again, we consider 3 months before the treatment (i.e., June 1, 2016, to August 31, 2016) as the pre-treatment period and 3 months after the treatment (i.e., September 1, 2016, to November 30, 2016) as the post-treatment period. Within this framework, we observe the change in user behavior in the post-treatment period versus the pre-treatment period. In addition, to further eliminate potential time-varying factors that may confound our empirical analysis, we follow an emerging approach used in several recent studies (e.g., [42, 61]) to obtain and use additional data from the year before the start of the pre-treatment period (i.e., June 1, 2015, to November 30, 2015) to construct a control group.[3] Figure 2.2 of Online Supplemental Appendix 2 provides a visual summary of the research design.

It is important to note that the control group does not necessarily consist of the same reviewers as those in the treatment group, nor do they review identical sets of restaurants. Therefore, to further enhance the validity of the control group (i.e., to ensure that users in the control group exhibit pre-treatment behavior that is similar to that of users in the treatment group), we employ the coarsened exact matching (CEM) to construct a matched sample wherein both groups exhibit parallel pre-treatment trends. Details of the matching are reported in Table 2.2 of Online Supplemental Appendix 2.

Before estimating the DiD specification, we first verify that the parallel trend assumption holds (i.e., that the behavior of treatment users and that of control users follow the parallel trend in the pre-treatment period). To test this assumption, we implement a relative time model [2]. Details are reported in the Online Supplemental Appendix 2 (Table 2.3). Once we successfully obtain a set of matched samples, we proceed with the DiD analysis with the following specification:

Table 5. Effect of downvote option removal on review writing behavior (DiD).

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	$Ln(ReviewCnt_{it})$	$Ln(Length_{it})$	$Ln(Photo_{it})$	$Readability_{it}$	$TopicGini_{it}$	$Similarity_{it}$
$DownvoteRemoval_{it}$	-0.132***	0.021*	0.043**	0.205**	-0.005*	-0.010*
	(0.013)	(0.010)	(0.011)	(0.068)	(0.002)	(0.005)
$Ln(Tenure_{it})$	-0.275***	0.068***	0.073***	0.095	0.002	-0.005
	(0.018)	(0.014)	(0.013)	(0.067)	(0.002)	(0.006)
Constant	3.122***	5.464***	1.077***	7.127**	0.616***	0.805***
	(0.113)	(0.088)	(0.081)	(0.414)	(0.014)	(0.035)
Reviewer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,856	21,856	21,856	21,856	21,856	21,856
R-squared	0.179	0.012	0.011	0.004	0.003	0.002

Notes: Standard errors in parentheses are robust and clustered by reviewers.
***$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

$$Y_{it} = \beta_0 + \beta_1 DownvoteRemoval_{it} + \beta_2 X_{it} + \lambda_i + \delta_t + \mu_{it} \quad (2)$$

where Y_{it} is the set of dependent variables that capture the quantity and quality of the reviews. $DownvoteRemoval_{it}$ is equal to 1 for treated reviewers in the periods following the platform's removal of the downvote option, and 0 otherwise. X_{it} are control variables (i.e., $Tenure_{it}$). λ_i is the reviewer fixed effect that controls for time-invariant reviewer characteristics that impact a reviewer's writing behavior. δ_t is the calendar week fixed effect that controls for unobservable time-specific factors that affect review generation. μ_{it} is the unexplained error term. β_1 is the parameter of interest, which represents the average treatment effect of downvote removal on the treated reviewers.

Table 5 reports the supplementary DiD estimates, which are broadly consistent with our main RDiT results. Namely, the removal of the downvote option reduces review quantity but enhances multiple dimensions of review quality among existing users. The convergence of results across distinct identification strategies reinforces the robustness of our causal inference regarding the impact of the system design change.

Mechanism Analysis

In this section, we examine the mechanisms underlying the observed effects after the removal of the downvote option. We first test whether changes in perceived procedural fairness help explain the decline in review quantity by examining heterogeneity based on reviewers' prior exposure to negative evaluation. Shifting our attention to review quality, we then assess whether reduced evaluative diagnosticity leads reviewers to invest more effort in differentiation by examining heterogeneous responses across reviewer segments and quality dimensions. Together, these analyses demonstrate the dominant mechanisms following the platform-level design change.

What Drives the Decrease in Content Quantity?

First, in H1B, we propose that the removal of the downvote option undermines the review voting system's ability to differentiate between high-quality and low-quality reviews, leading to a perception of unfairness and, consequently, a decrease in review contributions. To

indirectly measure reviewers' perceived (un)fairness, we rely on the number of downvotes received per review for each reviewer before the downvote removal. It is a common tendency for individuals to attribute poor performance to external factors, such as unfair evaluation by others (e.g., [5]). Therefore, the greater the number of downvotes received by a reviewer, the more likely they are to perceive the platform as unfair. This perception of unfairness may discourage them from continuing to contribute reviews. However, after the removal of the downvote option, reviewers who had previously received a higher number of downvotes tended to perceive the platform as being fairer. As a result, they were more likely to contribute reviews compared with reviewers who had received fewer downvotes in the past. Using the number of pre-policy downvotes as a proxy for exposure to negative evaluation also provides indirect insight into how reviewers who may have experienced more "toxic" treatment responded to the removal, thereby highlighting heterogeneous impacts of the design change. To test the moderating effect of the number of downvotes received, we use the following specification:

$$Y_{it} = \gamma_0 + \gamma_1 DownvoteRemoval_{it} + \gamma_2 After_t \times \mathrm{In}(DownvoteRcvd_i) + \gamma_3 DownvoteRemoval_{it} \times \mathrm{In}(DownvoteRcvd_i) + \gamma_4 X_{it} + \lambda_i + \delta_t + \mu_{it} \quad (3)$$

where $After_t$ is a dummy variable equal to 1 for observations in the post-treatment period, and 0 for those in the pre-treatment period. The variable $DownvoteRcvd_i$ represents the average number of downvotes per review received by reviewer i in the pre-treatment period. Our primary focus is on the coefficient γ_3, associated with the interaction term between $DownvoteRemoval_{it}$ and $\mathrm{Ln}(DownvoteRcvd_i)$, which captures the moderating effect.

The result of this moderating effect analysis is reported in Table 6. We obtain a positive and statistically significant coefficient on the interaction term between $DownvoteRemove_t$ and $Ln(DownvoteRcvd_i)$. This suggests that reviewers who received more downvotes before the policy change reduced their review quantity less after the downvote removal, likely because they perceived the evaluation environment to be fairer under the new system.

Table 6. The moderating effect of pre-policy downvote received on review quantity.

VARIABLES	$Ln(ReviewCnt_{it})$
$DownvoteRemoval_{it}$	-0.094**
	(0.026)
$After_t \times Ln(DownvoteRcvd_i)$	-0.075**
	(0.008)
$DownvoteRemoval_{it} \times Ln(DownvoteRcvd_i)$	0.041*
	(0.018)
$Ln(Tenure_{it})$	-0.239**
	(0.018)
Constant	2.912***
	(0.111)
Reviewer fixed effects	Yes
Week fixed effects	Yes
Observations	21,856
R-squared	0.179

Notes: Standard errors in parentheses are robust and clustered by reviewers.
***$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

What Drives the Increase in Content Quality?

Second, we propose in H2B that after the platform disables the downvote option, review writing becomes more of a "popularity contest," in which reviewers compete primarily for favorable attention from readers. Because the voting system becomes less diagnostic without the possibility of downvotes, reviewers seeking visibility may have had to exert greater effort to stand out and capture readers' attention, for example, by writing longer reviews, including more photos, or expanding the diversity and novelty of their content.

To test this mechanism, we turn to another moderation analysis. We use the ratio of first-person pronouns (I, me, my) to total pronouns as a proxy for the degree of self-focus in reviewers' writing. Prior research suggests that first-person pronoun use reflects a self-centric perspective, whereas lower reliance on first-person pronouns indicates a greater orientation toward the audience (e.g., [51]). Studies on impression management further show that self-focused or self-conscious individuals are especially motivated to manage impressions when evaluative environments shift or stakes are high (e.g., [30]). Building on this insight, we argue that reviewers with a stronger self-centric orientation are more motivated to stand out in the absence of diagnostic downvote signals, and thus should be especially likely to increase the quality of their reviews once the system shifts toward a popularity contest.

We therefore interact $DownvoteRemoval_{it}$ with the reviewer's pre-policy first-person pronoun ratio ($IRatio_i$). We use the specification described in Eq. (4). If the mechanism operates as expected, the quality improvements observed after the removal should be stronger among reviewers with higher first-person pronoun ratios.

$$Y_{it} = \zeta_0 + \zeta_1 DownvoteRemoval_{it} + \zeta_2 After_t \times IRatio_i + \zeta_3 DownvoteRemoval_{it} \times IRatio_i + \zeta_4 X_{it} + \lambda_i + \delta_t + \mu_{it} \quad (4)$$

The results in Table 7 support this prediction: the larger the first-person pronoun ratio, the stronger the observed improvements in effort-related quality measures (i.e., length and photos) and topic breadth-related measures (i.e., diversity and novelty). These findings provide consistent evidence that the quality effects of the downvote removal are driven, at least in part, by reviewers' increased motivation to differentiate their contributions in a less

Table 7. The moderating effect of pre-policy first person pronoun ratio on review quality.

	(1)	(2)	(3)	(5)	(6)
Variables	$Ln(Length_{it})$	$Ln(Photo_{it})$	$Readability_{it}$	$TopicGini_{it}$	$Similarity_{it}$
$DownvoteRemoval_t$	-0.117**	-0.106**	0.419	0.017***	0.046**
	(0.048)	(0.048)	(0.276)	(0.007)	(0.021)
$After_t \times IRatio_i$	0.171***	0.175***	-0.239*	-0.010**	-0.031**
	(0.031)	(0.031)	(0.122)	(0.004)	(0.013)
$DownvoteRemoval_{it} \times IRatio_i$	0.116*	0.144*	-0.291	-0.025***	-0.0583*
	(0.050)	(0.061)	(0.365)	(0.008)	(0.026)
$Ln(Tenure_{it})$	0.039***	0.054***	0.136**	-0.001*	-0.001
	(0.015)	(0.013)	(0.064)	(0.001)	(0.006)
Constant	5.690***	1.215***	6.786***	0.640***	0.687***
	(0.091)	(0.084)	(0.401)	(0.005)	(0.038)
Observations	20,841	20,841	20,841	20,841	20,841
R-squared	0.006	0.009	0.002	0.001	0.001

Notes: Standard errors in parentheses are robust and clustered by reviewers.
***$p < 0.00$. **$p < 0.01$. *$p < 0.05$.

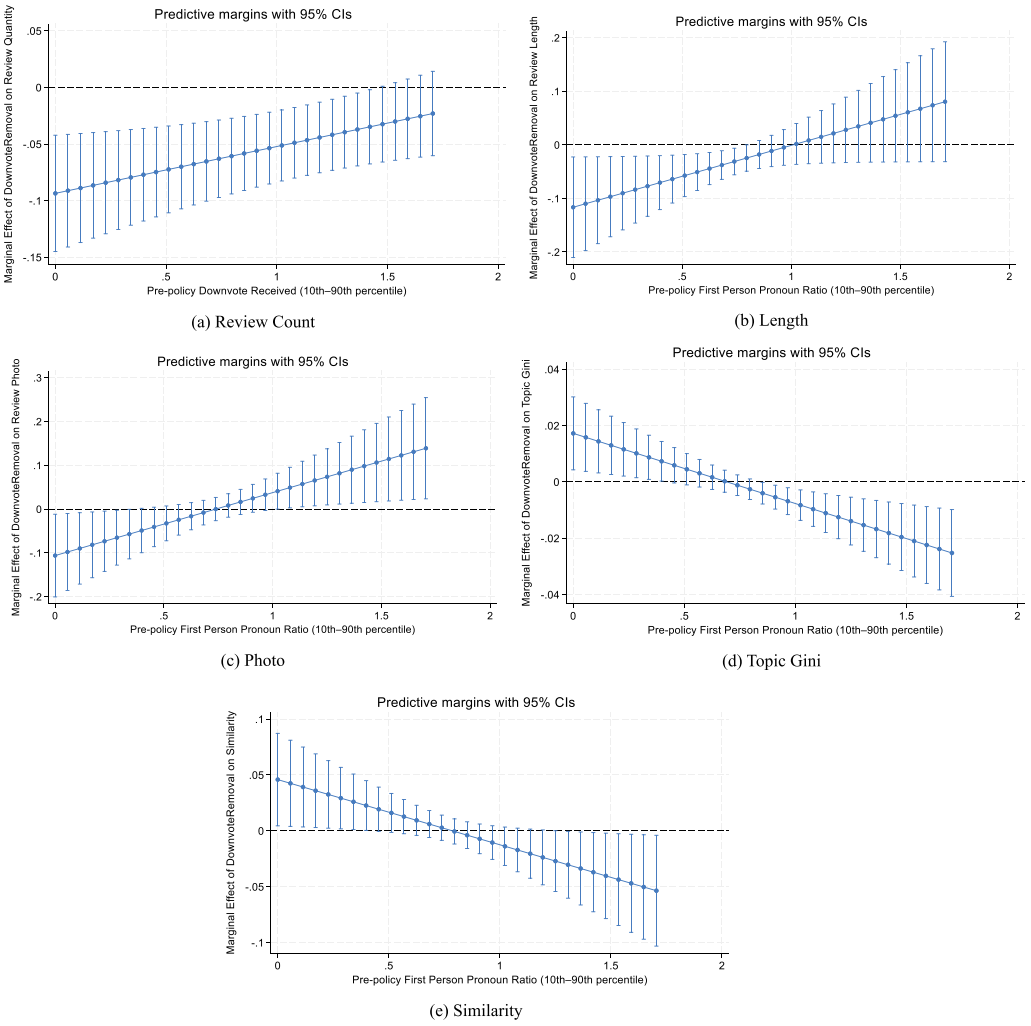

Figure 2. Marginal effects of downvote removal on review quantity and quality.

diagnostic voting environment. The moderation effect is insignificant on reading difficulty, probably because of its ambiguous signaling role. Lower reading difficulty can boost comprehension and persuasiveness by making content easier to process [44, 45], yet higher reading difficulty can also signal sophistication or expertise [33]. Because contributors may pursue either strategy, their adjustments may offset each other.

Furthermore, we conduct marginal effect analysis. Figure 2 plots the marginal effect of downvote removal across the distribution of the moderators. For review quantity, the negative effect is concentrated among reviewers with low prior exposure to negative evaluation and attenuates as the moderator increases, becoming statistically indistinguishable from zero at higher levels. For review quality, heterogeneous patterns emerge across dimensions: the negative effect on review length weakens and disappears, while for photo posting, the marginal effect becomes positive and statistically significant. By contrast, for Topic Gini and content similarity, the marginal effects decline with the moderator,

indicating greater topical breadth and differentiation. Together, these patterns show that the aggregate quantity–quality trade-off arises from heterogeneous behavioral responses to the platform design change.

We acknowledge that the platform intervention may activate multiple behavioral mechanisms simultaneously, potentially operating in different directions for different reviewers. Accordingly, our empirical strategy is not intended to establish the exclusive presence of a single mechanism, but rather to assess which mechanisms dominate in shaping observed outcomes. The heterogeneity and marginal-effect analyses help rule out purely uniform or mechanical platform effects, such as generic attention shifts after the interface change, which would predict similar directional responses across reviewer segments and outcome dimensions. Instead, we observe systematic heterogeneity and sign reversals: participation declines for some reviewers, while effort-intensive quality dimensions (e.g., photo posting) increase for others, alongside divergent patterns in topical concentration and content similarity. These patterns are difficult to reconcile with a single uniform explanation and suggest that multiple mechanisms may coexist and vary in their relative strength across review contributors.

Heterogeneity Analysis

To enrich our understanding of the downvote removal and to provide more practically useful insights, we complement the main estimates with a heterogeneity analysis. Examining how effects differ across meaningful user segments helps identify boundary conditions for the observed quantity–quality trade-off and indicates which groups of contributors drive the aggregate patterns.

We focus on two reviewer characteristics: tenure on the platform ($PreTenure_i$) and pre-policy follower counts ($PreFollower_i$). These variables capture distinct but important aspects of reviewer status. Tenure reflects how long a contributor has been active, which is associated with deeper embeddedness in platform norms and accumulated habits.

Table 8. Heterogeneous effect of reviewer tenure.

	(1)
Variables	$Ln(ReviewCnt_{it})$
$DownvoteRemoval_{it}$	0.338**
	(0.099)
$After_t \times Ln(PreTenure_i)$	-0.023***
	(0.002)
$DownvoteRemoval_{it} \times Ln(PreTenure_i)$	-0.049**
	(0.015)
$Ln(Tenure_{it})$	-0.260***
	(0.020)
Constant	3.052**
	(0.124)
Observations	21,856
R-squared	0.046

Note: Standard errors in parentheses are robust and clustered by reviewers.
***$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

Table 9. Heterogeneous effect of follower size.

	(1)	(2)	(3)
Variables	$Ln(Length_{it})$	$Ln(Photo_{it})$	$Similarity_{it}$
$DownvoteRemoval_{it}$	0.047**	0.0716***	-0.016*
	(0.015)	(0.017)	(0.008)
$After_t \times Ln(PreFollower_i)$	0.011**	0.009**	-0.001
	(0.003)	(0.003)	(0.001)
$DownvoteRemoval_{it} \times Ln(PreFollower_i)$	-0.019**	-0.019**	0.005*
	(0.005)	(0.005)	(0.002)
$Ln(Tenure_{it})$	0.056**	0.062***	-0.001
	(0.014)	(0.013)	(0.006)
Constant	5.534***	1.143***	0.737***
	(0.089)	(0.082)	(0.036)
Observations	21,856	21,856	21,856
R-squared	0.005	0.006	0.005

Note: Standard errors in parentheses are robust and clustered by reviewers.
***$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

Follower count reflects a reviewer's social reach on the platform, as well as their potential return on investing effort in writing reviews.

The results show clear heterogeneity consistent with our proposed mechanisms. According to Table 8, the decline in review quantity is significantly stronger among long-tenure reviewers, suggesting that more embedded users (those most accustomed to the prior evaluative system) reduce their contributions more when the downvote option is removed.

Table 9 shows that quality improvements on effort-related and breadth-related measures (e.g., review length, photos, and novelty) are disproportionately concentrated among low-follower reviewers. This pattern aligns with our main argument: when voting becomes less diagnostic and the system turns into more of a popularity contest, lower-reach reviewers have stronger incentives to invest additional effort per review to stand out, while established reviewers can rely on their existing reach and show attenuated improvements.

Taken together, the heterogeneity patterns in Tables 8 and 9 suggest that the effects of downvote removal are not uniform across review contributors, but instead reflect systematic differences in experience and visibility. The decline in review quantity is driven primarily by long-tenure reviewers, while improvements in effort-intensive and novelty-related quality dimensions are concentrated among low-follower reviewers. From a platform-governance perspective, these findings indicate that removing negative evaluation may disproportionately discourage established contributors while simultaneously motivating less visible reviewers to invest more effort in differentiation. This asymmetric response highlights the importance of targeted governance designs that balance retention of experienced contributors with incentives that encourage high-quality contributions from less prominent users.

Discussions and Implications

This study examines the effects of removing the downvote option from an online restaurant review platform in Asia. By analyzing user behavior before and after the policy change, we find that eliminating downvotes led to fewer but higher-quality reviews, reflected in

increased depth, modality, diversity, and novelty. The decline in review quantity appears driven by perceptions of reduced fairness, whereas the quality gains likely reflect reviewers investing more effort to stand out in the ensuing "popularity contest."

Theoretical Implications

First, our study contributes to the extensive UGC literature by examining a platform-level design change rather than user-level exposure to feedback. A growing body of research has documented how the reception of positive or negative peer evaluations affects the behavior of feedback recipients, and how changes in the visibility of such feedback alter participation (e.g., [15, 27, 52, 63]). However, this work largely focuses on settings where the negative evaluation option remains available. A minimal amount of research has empirically investigated the impact of platform-level design choices that eliminate the option altogether. We address this critical gap by analyzing the effects of removing the downvote option on two core outcomes of user-generated content (i.e., review quantity and multi-dimensional quality), both of which have direct implications for platform operations. Removing the downvote option is not simply the opposite of receiving downvotes, nor is it merely a manipulation of visibility, it represents a structural change that eliminates the capacity for negative evaluation and reconfigures the evaluation environment. By isolating this platform-level design change in a field setting, we document a managerially salient quantity–quality trade-off and clarify users' behavioral responses that platforms may anticipate when redesigning peer-evaluation systems.

Second, we provide initial evidence for the dominant mechanisms by which the removal of the downvote option can impact the quantity and quality of reviews authored by existing reviewers. Most relevant to our work, Zhu et al. [63] find that reviewers who receive downvotes tend to produce more and better content compared to those who receive no feedback. A natural extension of this finding is that getting rid of the downvote option at the platform level should reduce the quantity and quality of generated reviews, but this extension is inadequate because the receipt of negative evaluations (versus no feedback) for a review would orient the reviewer's attention to the review itself, while the removal (versus retainment) of the downvote option may influence the reviewer's perception of the entire review system and lead to different kinds of motivations. Our findings suggest that the removal of unhelpful votes reduces review quantity, and a likely reason is an unfair perception of the review system. In addition, we find the removal of unhelpful votes increases (not decreases) the quality of reviews, likely because of reviewers' need to stand out in such a less diagnostic environment. Understanding these underlying mechanisms is essential for researchers seeking to optimize user engagement and content quality via design features of UGC platforms.

Third, our findings also bear implications for the broader notion of "the wisdom of crowds." Customer reviews and their evaluations aggregate many users' opinions to assess product and review quality. However, different processes may influence each other, leading to unintended consequences that may undermine the wisdom of crowds. For instance, the removal of the downvote option (reflecting social disapproval) represents a tweak to the review evaluation system, but this simple tweak may have a powerful impact on review generation. While tremendous evidence has been accumulated on how customer reviews are evaluated [26], our findings suggest a greater need for future research to study the reverse, for example, how changes to the review evaluation system may influence customer reviews.

Practical Implications

Our study also has practical implications for review platforms seeking to optimize the peer evaluation system on their sites. Intuition suggests that getting rid of the downvote option should get rid of the fear of review writers, thus encouraging reviewers to make more contributions. However, we find evidence against this intuition and reveal an unanticipated cost of this design change. The quantity of reviews on a review platform is typically a strategic advantage, so platforms that wish to cultivate this advantage should think twice before getting rid of the downvote option.

On the other hand, removing the downvote option can encourage existing users to write longer, more diverse, and more novel reviews. Thus, review platforms that get rid of unhelpful votes may stand to benefit from better quality of the generated reviews. Moreover, this quantity-quality trade-off should be carefully considered. Platform administrators should balance the need for more reviews with the need for high-quality reviews. As an example, for reviewers who have rarely received negative peer evaluations and may therefore be more discouraged by such a change in the peer evaluation system, platforms could offer alternative incentives, such as badges and peer awards, to demonstrate that the platform remains fair and values their contributions.

Our findings highlight a broader design tension between reducing abuse and preserving evaluative informativeness. Removing the downvote option can limit malicious or abusive voting and reduce social-evaluation anxiety, but it also leads to a decline in review quantity. Fewer reviews may narrow coverage and reduce perspective diversity, while the absence of negative signals can make it harder for users and algorithms to identify less helpful content. Managers considering similar interventions should therefore pair downvote removal with complementary interventions, such as verified flags, expert curation, or reputation scores, to preserve informational value while protecting contributors.

Future Research

The paper also opens up opportunities for future research. First, although our empirical setting is a large restaurant review platform, which is inherently hedonic and context-specific, the theoretical mechanisms we highlight (such as perceived procedural fairness and shifts in effort to stand out under a less diagnostic voting system) may or may not generalize to other types of UGC environments. For example, on utilitarian or knowledge-sharing platforms such as Stack Overflow, contributors are often motivated by reputation systems, expertise signaling, and instrumental reciprocity. In such contexts, removing the negative feedback option might attenuate or alter the quantity–quality trade-off, as fairness perceptions may be less salient and the content itself more objectively verifiable. Future research could examine whether the same governance change produces similar or different patterns on more utilitarian platforms, thereby extending the theoretical reach of our findings.

Second, our analysis primarily tracks the content generation behavior of existing reviewers both before and after the peer evaluation design modification. Future research could explore how this adjustment in platform design impacts the engagement of new users, potentially either attracting or discouraging their participation. These avenues for future investigation underscore the need for a more comprehensive understanding of the broader implications of platform design changes.

Third, although our quasi-experimental design allows us to identify the effects of a major platform-level design change, it does not permit direct manipulation of specific mechanisms or direct observation of contributors' subjective perceptions, time allocation, or effort investment. Such data limitations therefore prevent us from fully disentangling coexisting mechanisms or tracing how contributors consciously reallocate effort across quantity and quality dimensions. We accordingly view experimental work, especially field experiments conducted in collaboration with platforms, as an important avenue for future research. Such studies could directly measure perceived justice and evaluation anxiety, causally isolate individual mechanisms, test alternative interface designs, and trace behavioral adaptation over time, thereby complementing and extending our observational findings.

Finally, our study focuses on the consequences of removing the downvote option, but future work could examine how to design interventions that deter malicious downvoting behaviors while still preserving the evaluative informativeness of negative feedback mechanisms. Such research could help platforms strike a better balance between fairness, abuse prevention, and the diagnostic signals that support content curation and decision making.

Conclusions

This study examines the consequences of disabling downvotes on an online restaurant review platform. Using an augmented local linear RDiT design, complemented by DiD analyses, we show that removing downvotes generates a clear quantity–quality trade-off: reviewers contribute fewer reviews, but the remaining reviews become longer and more diverse and novel in content.

We further provide evidence consistent with the dominant mechanisms underlying these effects. The decline in review quantity is concentrated among reviewers with limited prior exposure to negative evaluation, consistent with increased perceived unfairness after the removal of downvotes. In contrast, improvements in review quality reflect greater effort and differentiation in a less diagnostic evaluative environment. Heterogeneity analyses further indicate that quantity reductions are driven primarily by long-tenure reviewers, whereas quality improvements are concentrated among reviewers with fewer followers.

Overall, our findings highlight that platform governance interventions aimed at reducing toxicity can reshape both participation and informational value in systematic and heterogeneous ways. As platforms continue to redesign peer evaluation systems, our study underscores the importance of considering both behavioral margins and contributor segments when evaluating the welfare implications of design changes.

Notes

1.. See https://www.reviewercollective.com/the-downvoting-game-on-amazon-reviews/.
2.. See https://www.tripadvisor.com/ShowTopic-g1-i12104-k3265805-Non_Helpful_Button_on_Reviews-Help_us_make_Tripadvisor_better.html,
3.. Our LDA analysis encompasses the entire dataset, incorporating reviews written by both the treated group of reviewers and those in the control group.

Acknowledgments

We thank the Editor and the anonymous reviewers for their constructive feedback and guidance, which greatly improved this manuscript. Generative artificial intelligence tools (Grammarly and ChatGPT) were used during the preparation of this manuscript solely for proofreading and sentence refinement. These tools were employed exclusively for editorial purposes to identify grammatical errors and enhance clarity of expression. All research concepts, study design, data analyses, interpretations, and conclusions are entirely the original contributions of the authors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Yinan Yu (yinan.yu@ou.edu; corresponding author is Assistant Professor in the Division of Management Information Systems at the Michael F. Price College of Business, University of Oklahoma. She researches digital platforms, social media, user generated content, multi-channel management, and economics of AI. Dr. Yu's work has been published in *MIS Quarterly, Journal of Management Information Systems, Production and Operations Management*, and other academic journals.

Dezhi Yin (dezhiyin@usf.edu) is Associate Professor of Information Systems and Muma Fellow at Muma College of Business, University of South Florida. He received his PhD in IT management from Georgia Institute of Technology. His research interests include emotional expressions in online environments, user-generated content (e.g., online word-of-mouth), and AI-generated content (e.g., chatbots). Dr. Yin has published in *MIS Quarterly, Information Systems Research, Journal of the Association for Information Systems*, and other academic journals.

Warut Khern-am-nuai (warut.khern-am-nuai@mcgill.ca) is the FRQ–IVADO Chair in Economics and Ethics of Artificial Intelligence and an Associate Professor of Information Systems at the Desautels Faculty of Management, McGill University, Canada. He earned his PhD in Management Information Systems from the Krannert School of Management, Purdue University. Dr. Khern-am-nuai's research interests include the economics and ethics of artificial intelligence, platform for online marketplaces, and management of information security.

ORCID

Yinan Yu http://orcid.org/0000-0002-4581-8469

References

1. Allard, T.; Dunn, L.H.; and White, K. Negative reviews, positive impact: Consumer empathetic responding to unfair word of mouth. *Journal of Marketing*, *84*, 4 (2020), 86–108.
2. Angrist, J.; and Pischke, J. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press, 2008.
3. Baek, H.; Ahn, J.; and Choi, Y. Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, *17*, 2 (2012), 99–126.
4. Blei, D.M.; Ng, A.Y.; and Jordan, M.I. Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3* (2003), 993–1022.
5. Bradley, G.W. Self-serving biases in the attribution process: A reexamination of the fact or fiction question. *Journal of Personality and Social Psychology*, *36*, 1 (1978), 56–71.

6. Brynjolfsson, E.; Hu, Y.; and Simester, D. Goodbye Pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, 58, *7* (2011), 1373–1386.
7. Burtch, G.; Hong, Y.; Bapna, R.; and Griskevicius, V. Stimulating online reviews by combining financial incentives and social norms. *Management Science*, 64, *5* (2018), 2065–2082.
8. Burtch, G.; He, Q.; Hong, Y.; and Lee, D. How do peer awards motivate creative content? Experimental evidence from Reddit. *Management Science*, *68*, 5 (2022), 3488–3506.
9. Chan, K.W.; Li, S.Y.; Ni, J.; and Zhu, J.J. What feedback matters? The role of experience in motivating crowdsourcing innovation. *Production and Operations Management*, *30*, 1 (2021), 103–126.
10. Chen, W.; Gu, B.; Ye, Q.; and Zhu, K.X. Measuring and managing the externality of managerial responses to online customer reviews. *Information Systems Research*, *30*, 1 (2019), 81–96.
11. Chevalier, J.A.; Dover, Y.; and Mayzlin, D. Channels of impact: User reviews when quality is dynamic and managers respond. *Marketing Science*, *37*, 5 (2018), 688–709.
12. Cohen-Charash, Y.; and Spector, P.E. The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*, 2 (2001), 278–321.
13. Connors, L.; Mudambi, S.; and Schuff, D. Is it the review or the reviewer? A multi-method approach to determine the antecedents of online review helpfulness. In *Proceedings of the 44th Hawaii International Conference on System Sciences*. Piscataway, NJ, USA: IEEE, 2011, pp. 1–10.
14. Cooper, H. Evaluating and interpreting research syntheses in adult learning and literacy. *NCSALL Occasional Paper*. National Center for the Study of Adult Learning and Literacy, 2007.
15. Deolankar, V.; Fong, J.; and Sriram, S. Content generation on social media: The role of negative peer feedback. Working paper, Available at SSRN 4522092 (2024).
16. Deng, Y.; Zheng, J.; Khern-am-nuai, W.; and Kannan, K. More than the quantity: The value of editorial reviews for a user-generated content platform. *Management Science*, *68*, 9 (2022), 6865–6888.
17. Forman, C.; Ghose, A.; and Wiesenfeld, B. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*, 3 (2008), 291–313.
18. Germano, F.; Gómez, V.; and Le Mens, G. The few-get-richer: A surprising consequence of popularity-based rankings? In *Proceedings of the World Wide Web Conference*. 2019, pp. 2764–2770.
19. Goes, P.B.; Lin, M.; and Au Yeung, C.M. "Popularity effect" in user-generated content: Evidence from online product reviews. *Information Systems Research*, *25*, 2 (2014), 222–238.
20. Golder, P.N.; Dekimpe, M.D.; An, J.T.; van Heerde, H.J.; Kim, D.S.U.; and Alba, J.W. Learning from data: An empirics-first approach to relevant knowledge generation. *Journal of Marketing*, *87*, 3 (2023), 319–336.
21. Gopal, A.; Chen, P.Y.; Oh, W.; Xu, S.X.; and Sarker, S. On crafting effective theoretical contributions for empirical papers in economics of information systems: Some editorial reflections. *Information Systems Research*, *35*, 3 (2024), 917–935.
22. Graebner, M.E.; Knott, A.M.; Lieberman, M.B.; and Mitchell, W. Empirical inquiry without hypotheses: A question-driven, phenomenon-based approach to strategic management research. *Strategic Management Journal*, *44*, 1 (2023), 3–10.
23. Greene, J.; and Haidt, J. How (and where) does moral judgment work? *Trends in Cognitive Sciences*, *6*, 12 (2002), 517–523.
24. Hausman, C.; and Rapson, D.S. Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, *10* (2018), 533–552.
25. Ho, Y.C.; Wu, J.J.; and Tan, Y. Disconfirmation effect on online rating behavior: A structural model. *Information Systems Research*, *28*, 3 (2017), 626–642.
26. Hong, H.; Xu, D.; Wang, G.; and Fan, W. Understanding the determinants of online review helpfulness. *Decision Support Systems*, *102* (2017), 1–11.
27. Kim, A.; Lu, Y.; Ma, T.; and Tafti, A.R. The impact of YouTube's hiding dislike count on viewer and creator engagement. Working paper, Available at SSRN 4967594 (2024).

28. Kuan, K.; Hui, K.; Prasarnphanich, P.; and Lai, H. What makes a review voted? An empirical investigation of review voting in online review systems. *Journal of the Association for Information Systems*, *16*, 1 (2015), 48–71.
29. Leary, M.R. A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin*, *9*, 3 (1983), 371–375.
30. Leary, M.R.; and Kowalski, R.M. Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*, 1 (1990), 34–47.
31. Lee, Y.J.; Hosanagar, K.; and Tan, Y. Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Science*, *61*, 9 (2015), 2241–2258.
32. Moe, W.W.; and Schweidel, D.A. Online product opinions: Incidence, evaluation, and evolution. *Marketing Science*, *31*, 3 (2012), 372–386.
33. Moldovan, A. Technical language as evidence of expertise. *Languages*, *7*, 1 (2022), 41.
34. Moqri, M.; Mei, X.W.; Qiu, L.F.; and Bandyopadhyay, S. Effect of "following" on contributions to open source communities. *Journal of Management Information Systems*, *35*, 4 (2018), 1188–1217.
35. Moss, B.G.; and Yeaton, W.H. Shaping policies related to developmental education: An evaluation using the regression-discontinuity design. *Educational Evaluation and Policy Analysis*, *28*, 3 (2006), 215–229.
36. Muchnik, L.; Aral, S.; and Taylor, S.J. Social influence bias: A randomized experiment. *Science*, *341*, 6146 (2013), 647–651.
37. Mudambi, S.; and Schuff, D. Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*, 1 (2010), 185–200.
38. Murphy, K.; and Tyler, T. Procedural justice and compliance behaviour: The mediating role of emotions. *European Journal of Social Psychology*, *38*, 4 (2008), 652–668.
39. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; and Dubourg, V. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12* (2011), 2825–2830.
40. Proserpio, D.; and Zervas, G. Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, *36*, 5 (2017), 645–665.
41. Qiu, L.F.; and Kumar, S. Understanding voluntary knowledge provision and content contribution through a social-media-based prediction market: A field experiment. *Information Systems Research*, 28, 3 (2017), 529–546.
42. Quinn, M.; and Gutt, D. Heterogeneous effects of generative AI on knowledge seeking in online communities. *Journal of Management Information Systems*, *42*, 2 (2025), 370–399.
43. Ravichandran, T.; and Deng, C.Q. Effects of managerial response to negative reviews on future review valence and complaints. *Information Systems Research*, *34*, 1 (2023), 319–341.
44. Reber, R.; Winkielman, P.; and Schwarz, N. Effects of perceptual fluency on affective judgments. *Psychological Science*, 9, 1 (1998), 45–48.
45. Schwarz, N.; Newman, E.; and Leach, W. Making the truth stick & the myths fade: Lessons from cognitive psychology. *Behavioral Science & Policy*, *2*, 1 (2016), 85–95.
46. Shen, W.Q.; Hu, Y.J.; and Ulmer, J.R. Competing for attention: An empirical study of online reviewers' strategic behavior. *MIS Quarterly*, *39*, 3 (2015), 683–696.
47. Song, T.T.; and Ho, Y.C. Relational resonance and content creation. *Management Science*, *71*, 9 (2025), 7366–7389.
48. Sonenshein, S. The role of construction, intuition, and justification in responding to ethical issues at work: The sensemaking-intuition model. *Academy of Management Review*, *32*, 4 (2007), 1022–1040.
49. Surowiecki, J. *The Wisdom of Crowds*. New York: Anchor, 2005.
50. Tan, X.; Jin, F.J.; and Dennis, A.R. How appreciation and attention affect contributions to electronic networks of practice. *Journal of Management Information Systems*, *39*, 4 (2022), 1037–1063.
51. Tausczik, Y.R.; and Pennebaker, J.W. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 1 (2010), 24–54.

52. Turel, O.; and Qahri-Saremi, H. Role of "likes" and "dislikes" in influencing user behaviors on social media. *Journal of Management Information Systems*, *41*, 2 (2024), 515–545.
53. Watson, D.; and Friend, R. Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, *33*, 4 (1969), 448–457.
54. Wei, Z.; Xiao, M.; and Rong, R. Network size and content generation on social media platforms. *Production and Operations Management*, *30*, 5 (2021), 1406–1426.
55. Wu, R.; Wu, H.H.; and Wang, C.L. Why is a picture "worth a thousand words"? Pictures as information in perceived helpfulness of online reviews. *International Journal of Consumer Studies*, *45*, 3 (2021), 364–378.
56. Yoo, E.; Gu, B.; and Rabinovich, E. Diffusion on social media platforms: A point process model for interaction among similar content. *Journal of Management Information Systems*, *36*, 4 (2019), 1105–1141.
57. Yin, D.; Bond, S.; and Zhang, H. Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, *38*, 2 (2014), 539–560.
58. Yu, Y.; Khern-am-nuai, W.; Pinsonneault, A.; and Wei, Z. The impacts of social interactions and peer evaluations on online review platforms. *Journal of Management Information Systems*, *40*, 4 (2023), 1271–1300.
59. Zapata-Phelan, C.P.; Colquitt, J.A.; Scott, B.A.; and Livingston, B. Procedural justice, interactional justice, and task performance: The mediating role of intrinsic motivation. *Organizational Behavior and Human Decision Processes*, *108*, 1 (2009), 93–105.
60. Zeidner, M.; and Matthews, G. Evaluation anxiety. In *Handbook of Competence and Motivation*. New York: Guilford Press, 2005.
61. Zhang, X.; Xie, K.; Gu, B.; and Guo, X. Engaging newly enrolled physicians with introductory incentives: The role of online and offline references. *MIS Quarterly*, *49*, 3 (2025), 917-952.
62. Zheng, J.; Yin, G.; Yong, T.; and Din, J. Does help help? An empirical analysis of social desirability bias in ratings. *Information Systems Research*, *35*, 3 (2024), 1052–1073.
63. Zhu, K.; Khern-am-nuai, W.; and Yu, Y.N. Negative peer feedback and user content generation: Evidence from a restaurant review platform. *Production and Operations Management*, *34*, 12 (2024), 3814–3829.