Journal of the Association for Information Systems (2026) **27**(4), 1015-1045
doi: 10.17705/1jais.00991

RESEARCH ARTICLE

ISSN 1536-9323

Helpful or Unhelpful: The Hidden Influence of Review Voters and the Wisdom of Crowds

Yinan Yu,[1] Dezhi Yin,[2] Warut Khern-am-nuai,[3] Alain Pinsonneault,[4] Zaiyan Wei[5]

[1]University of Oklahoma, USA, yinan.yu@ou.edu
[2]University of South Florida, USA, dezhiyin@usf.edu
[3]McGill University, Canada, warut.khern-am-nuai@mcgill.ca
[4]McGill University, Canada, alain.pinsonneault@mcgill.ca
[5]Purdue University, USA, zaiyan@purdue.edu

Abstract

Online review platforms use voting systems to identify and promote helpful reviews, aiming to alleviate information overload for potential consumers. Existing research has primarily focused on how review and reviewer characteristics affect review helpfulness but has largely overlooked the role of voters. Drawing on cognitive consistency theories, we propose that review voters cast votes in ways that align with their historical content-related activities, perceptions of review context, and social relationships with reviewers. Using a unique, fine-grained vote-level dataset, we find that voter-related factors have a more pronounced impact on review helpfulness than review or reviewer characteristics do. These findings suggest that past research examining antecedents of review helpfulness may have overestimated the effects of review and reviewer characteristics and that more attention to voters is warranted.

Keywords: Online Reviews, Review Helpfulness, Voter, Cognitive Consistency

Traci Carte was the accepting senior editor. This research article was submitted on August 22, 2024, and underwent three revisions.

1 Introduction

Online consumer reviews significantly influence purchase decisions, but their sheer volume creates information overload and overwhelms consumers (Pan & Wu, 2020; Seutter et al., 2023; Yin et al., 2014). To mitigate this issue, many review platforms utilize a voting system that enables consumers to rate reviews as either *helpful* or *unhelpful*. Such a system facilitates the ranking of reviews based on their helpfulness (e.g., the ratio of *helpful* votes), thereby elevating the reviews deemed most helpful to prominent positions. These top reviews can sway consumers' purchase decisions (Lei et al., 2022) and influence their tendency to seek more reviews (Yin et al., 2023).

Due to the importance of review helpfulness for platforms, retailers, and service providers, a growing body of literature has examined factors that affect review helpfulness. According to two large-scale meta-analyses (Hong et al., 2017; Qahri-Saremi & Montazemi, 2019) and an extensive literature review from Liu et al. (2020), the existing literature (including those since 2020) typically classifies the antecedents into two broad categories: review-related, based on factors such as review length, readability, and emotions (e.g., Mudambi & Schuff, 2010; Yin et al., 2014, 2021), and reviewer-related, based on factors such as reviewer expertise, self-disclosure, and follower base (e.g., Choi et al., 2022; Forman et al., 2008; Kuan et al., 2015).

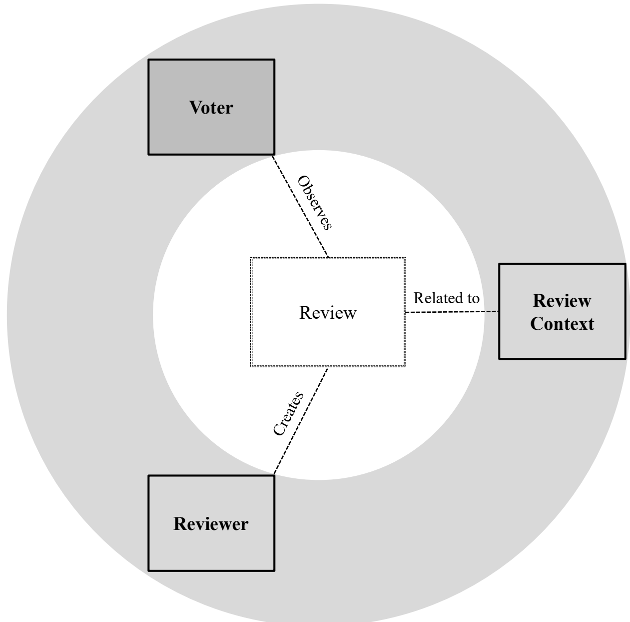


Figure 1. Triadic Voting Context

However, it is worth noting that *helpful* or *unhelpful* votes—the foundation of review helpfulness—are determined by a triad: the review itself, the reviewer, and the voter casting the vote. Despite the importance of voters in this process, their characteristics have received limited attention, most likely due to the scarcity of granular data at the individual vote level. Moreover, existing research often assumes that voters rationally evaluate reviews based solely on their quality, which is determined by review and reviewer attributes. This assumption also underpins the notion that the aggregated votes embody the "wisdom of crowds" and accurately reflect the helpfulness of reviews. It is crucial to explore the validity of this assumption and the premise of "wisdom of crowds" because *helpful* reviews not only enhance information-seeking efficiency and reduce purchase deferral (Yin et al., 2023) but are also usually more visible and repeatedly exposed to consumers. Such repeated exposure amplifies their influence over consumer decisions (Lei et al., 2022, 2025). Additionally, this line of inquiry could help businesses anticipate how different review voters (with diverse backgrounds and situations) may react differently to the same piece of online word-of-mouth and tailor their strategies to influence customer engagement and ultimate sales.

In this paper, we revisit the "wisdom of crowds" assumption by exploring the impacts of voter-specific factors and comparing them to the impacts of review- and reviewer-related antecedents of review helpfulness. Specifically, we ask the following research questions: *How do voter-related factors influence a voter's decision to cast a helpful or unhelpful vote? What are the primary drivers of review helpfulness: voter-related, review-related, or reviewer-related factors?* We argue that when voters cast a vote, they may bring their own experiences, preferences, and biases, which can influence their perceived review helpfulness above and beyond the impact of review and reviewer characteristics. Building on theories of cognitive consistency (e.g., Abelson et al., 1968; Kim et al., 2019; Zhu et al., 2024b), we propose that voters may exhibit a preference for consistency across three domains: their historical content-related behavior (e.g., past helpfulness vote patterns), their perceptions of contextual cues (e.g., a review's deviation from the product's average rating), and their social relationships with the reviewer (e.g., whether the voter follows the reviewer). We provide the first empirical test of these predictions using a unique, fine-grained vote-level dataset. Our findings indicate that when casting votes, voters prefer consistency with their past behavior, tend to vote consistently with contextual factors, and strive for consistency and alignment in social relationships. We also explored the relative impacts of voter-related factors compared to review-related and reviewer-related factors, which have been the predominant focus of prior research, finding that the former trumps the latter.

Our findings have important theoretical implications. First, we extend the existing literature on antecedents of review helpfulness by demonstrating that voter-related factors play a critical role and should not be overlooked. Second, we extend cognitive consistency theories by applying them to a real-world, triadic voting context—involving voter, reviewer, and review context (see Figure 1)—and distinguishing three forms of consistency—behavioral, contextual, and social. As such, this research expands the generalizability of consistency theories from traditional intrapersonal attitude settings to a multi-part, crowd-based environment. Third, we contribute to the literature on the wisdom of crowds, revealing that merely implementing a voting system does not necessarily lead to "unbiased" wisdom. Instead, because individual cognitive biases (such as consistency seeking across past behavior, contextual cues, and social relationships) can *systematically* skew votes, the resulting aggregate scores may reflect psychological alignment more than true review helpfulness. Finally, our findings complement recent research distinguishing the impacts of *helpful* and *unhelpful* votes (Yu et al., 2024; Zhu et al., 2024a) by introducing a voter-centric lens to better understand the origin and impact of vote casting.

Our findings also have crucial practical implications for managers of online review platforms regarding the design of review evaluation systems that rely on crowd voting. Specifically, the insights from our work demonstrate the need to improve and fine-tune crowd-based evaluation systems so that review helpfulness scores can better reflect review helpfulness and help prospective consumers make purchase decisions.

2 Literature Review and Hypotheses Development

2.1 Review Helpfulness Evaluation

The helpfulness evaluation of online reviews has attracted significant scholarly attention. Prior research, primarily based on review-level secondary data, has extensively examined a range of review- and reviewer-related determinants. As summarized in Table A1 in Appendix A, however, these studies have largely overlooked the potential influence of voters, leaving voter-related factors underexplored.

Based on the premise that the helpfulness of a review is determined primarily by the review itself, most prior studies have focused on how ratings and content characteristics, such as information amount and readability, impact review helpfulness (e.g., Kuan et al., 2015; Mudambi & Schuff, 2010). Star ratings of reviews have attracted tremendous attention in the literature, but empirical findings are mixed (Hong et al., 2017). Regarding review content characteristics, the seminal work of Mudambi and Schuff (2010) showed that the amount of information (operationalized as the length of review content) has a positive effect on review helpfulness. Kuan et al. (2015) further showed that longer reviews with more information are generally perceived to be more helpful until the review becomes too long for consumers to read. Regarding readability, the evidence of its impact is also mixed (e.g., Kuan et al., 2015; Yin et al., 2014); Hong et al.'s (2017) meta-analysis showed that its effect on review helpfulness is not significant. Assuming that the characteristics of review text are the primary drivers of review helpfulness, a separate stream of studies has applied technical methods to predict review helpfulness based on text characteristics (e.g., Kim et al., 2006).

Studies in a related stream have explored how review helpfulness is influenced by the characteristics of the reviewers writing the reviews. Examples of reviewer factors considered in this literature include reviewer expertise, reviewer credibility, and identity disclosure (Hong et al., 2017). Research has shown that reviewers' expertise and credibility are useful predictors of review helpfulness (e.g., Baek et al., 2012; Connors et al., 2011). Similarly, the social network circle of reviewers can also influence review helpfulness (e.g., Yu et al., 2023).

To the best of our knowledge, the empirical investigations in these two streams conducted data analyses at the review level, and most of them quantified review helpfulness using measures such as the ratio of *helpful* votes over all votes. In addition, individual voters—the third entity of the triad—have generally remained in the background of these empirical analyses, and the impact of voter-related factors has received little attention. For example, the review-level findings of Yin et al. (2016) implied a confirmation bias among voters, but their study did not provide direct evidence for this bias at the voter level (without vote-level or voter-level data), nor did they examine other tendencies of voters or compare the relative potency of voter-related factors versus review- and reviewer-level factors. The absence of voter-level investigation can be attributed to the lack of granular data on voters and their cast votes. Our study aims to address this important research gap by exploring how voter-related factors affect the casting of *helpful* vs. *unhelpful* votes.

2.2 Theories of Cognitive Consistency

Theories of cognitive consistency provide the overarching lens for our study. These theories posit that people experience psychological tension when their attitudes, beliefs, or actions are misaligned and they are motivated to reduce the tension and restore harmony (Abelson et al., 1968; Festinger, 1962; Osgood & Tannenbaum, 1955). Two well-established mechanisms are particularly relevant to our research: *cognitive dissonance theory*, which stresses the internal discomfort following incongruent behavior (Festinger, 1962), and *congruity (balance) theory*, which emphasizes maintaining evaluative harmony in triadic relationships (Osgood & Tannenbaum, 1955). Originally developed to explain changes in intrapersonal attitudes, these theories have since been applied across consumer and social media contexts, showing that even subtle inconsistencies can influence evaluations and behaviors (Kim et al., 2019; Yin et al., 2016; Zhu et al., 2024b). For instance, the concept of congruence has been used in user-generated content (UGC) research to examine how content-context alignment affects perceptions of content helpfulness (Peng et al., 2020).

Building on this integrative view, we propose three sets of hypotheses and map each set to the most relevant theory. Hypotheses 1 and 2 (*voters'* historical behavior) draw on *dissonance reduction*: Voters align new votes with prior actions to reduce internal tension. Hypotheses 3 and 4 (*contextual cues* of the review) likewise reflect *dissonance-driven* alignment: Voters conform to average ratings or accumulated votes to maintain coherence with situational norms. Hypotheses 5 and 6 (*voter-reviewer* social ties) invoke *congruity/balance*: Voters upvote reviews from reviewers whom they follow or who follow them to preserve harmony across the triad of voter, reviewer, and review. In applying these mechanisms to a real-world voting system, we extend cognitive consistency theories beyond traditional intrapersonal settings to a multi-part crowd-evaluation environment, and we also link individual alignment (in vote casting) to platform-level bias (when individual biases manifest at an aggregate level).

2.2.1 Consistency With Past Behavior

First, review voters may seek behavioral consistency with their past actions. Given that review evaluation and generation are two major content activities in the context of online reviews, we focus on voters' prior votes on others' reviews and their own review-writing experiences.

From the perspective of cognitive consistency, individuals strive to align their current behaviors with their past actions to avoid cognitive dissonance, enhance efficiency, and maintain psychological coherence (Festinger, 1962). As cognitive misers, consumers seek to make judgments efficiently by leveraging previously established patterns. For example, research on "cognitive inertia" has shown that the satisfaction judgments of repeat shoppers for new consumption experiences are heavily influenced by their prior satisfaction judgments (Mattila, 2003).

In the context of online reviews, a voter may have a tendency to evaluate reviews favorably or unfavorably, and this tendency or voting pattern reflects the voter's internal preferences and evaluation standards. To preserve consistency with their established voting history, voters may rely on these patterns when they evaluate new reviews. This drive for consistency implies that review voters who have cast favorable (unfavorable) votes in the past are more likely to cast favorable (unfavorable) votes for future reviews. Therefore, we propose:

H1: Review voters who have previously cast more *favorable votes* are more likely to cast *helpful* (vs. *unhelpful*) votes on reviews.

Apart from voting patterns, voters' past review-writing experience can also shape their voting behaviors. Voters with more review-writing experience are more likely to consider themselves to be reviewers (instead of mere review readers), and they are more likely to treat other reviewers as part of their "in-group." People generally prefer in-group members over out-group members because in-group members are likely to be more consistent with their self-identity (Castano et al., 2002; Fein & Spencer, 1997; Turner et al., 1987). Consequently, voters who are experienced reviewers are likely to view themselves as members of this group and will therefore be motivated to maintain consistency in their evaluations and treat other reviewers as peers whose efforts they value. As a result, they are more likely to evaluate others' reviews favorably, as doing so aligns with their "reviewer" identity.

We acknowledge that experienced experts may, in some cases, apply higher standards and evaluate others' work more critically (Lane et al., 2020). However, in the context of online reviews—where participation is voluntary and often motivated by reputation—we argue that the self-centered identity-consistency motive likely outweighs the critical scrutiny motive. Supporting peers through *helpful* votes allows experienced reviewers to express in-group solidarity and reinforce their self-image as constructive community members. Therefore, we propose:

H2: Review voters with more experience writing reviews themselves are more likely to cast *helpful* (vs. *unhelpful*) votes on reviews.

2.2.2 Consistency With Contextual Cues

Besides seeking consistency with their past behavior, voters are likely to also strive for contextual consistency by aligning their judgments with environmental cues that are cognitively processed during decision-making. Empirical evidence suggests that the helpfulness evaluation of user-generated content is not context-free (e.g., Peng et al., 2020; Yin et al., 2016). Instead, it is shaped by task-related contextual cues—signals from the surrounding environment that influence beliefs and mindsets (Murnane et al., 1999). Based on our triadic framework (involving voter, reviewer, and review context), we treat these review-related cues as part of the voter's evaluative context because they inform the perception and interpretation of review content at the moment of judging its helpfulness.

One such contextual cue is the average rating of a product or service. Prominently displayed on review platforms, the average rating shapes consumers' initial beliefs about the product or service (De Langhe et al., 2016). These preconceived beliefs foster confirmation bias, where individuals tend to favor information that aligns with their initial impressions, while dismissing contradictory information to minimize cognitive discomfort or dissonance (Chaiken et al., 1996; Nickerson, 1998).

Reviews that deviate significantly from the average rating, the basis of consumers' preexisting beliefs, challenge those beliefs and can trigger dissonance. To alleviate this dissonance and restore consistency, voters are more likely to dismiss or make counterarguments for such inconsistent reviews, favoring those that align more closely with the average rating.

Yin et al. (2016) offered review-level empirical evidence that the helpfulness of a review is not determined by the review's rating per se but by its deviation from the product's average rating. Note that this finding of confirmation bias *implies* a voter bias rather than directly testing it at the voter level. We propose that at the level of an individual vote, a voter will be more likely to cast a *helpful* (vs. *unhelpful*) vote on a review if the review deviates less from the average rating *at the time of voting*—not because of the review's content itself but because this alignment reinforces the voter's expectations or beliefs that are activated during evaluation.

While some prior work suggests that reviews with more extreme ratings may be seen as more informative due to their distinctiveness (e.g., Filieri et al., 2018), our focus on deviation from the average reflects a different mechanism rooted in confirmation bias. These contrasting perspectives underscore the need for empirical investigation into how rating alignment shapes individual voting decisions. Therefore, we propose:

H3: Review voters are more likely to cast *helpful* (vs. *unhelpful*) votes on reviews that deviate less from the average rating.

In addition, review voters may prefer contextual consistency and may thus align their helpfulness votes with prior voting patterns on a review, reflecting the bandwagon effect. The bandwagon effect refers to individuals' tendency to follow others' actions or judgments, often aligning their behavior with contextual cues that reflect collective opinions, regardless of their own judgment (Schmitt-Beck, 2015). This effect arises from either a preference to conform to social norms or the reliance on those norms as a convenient and credible heuristic for decision-making (van Herpen et al., 2009). The bandwagon effect manifests itself in *reviewer* behavior: When an individual writes a review, their rating of a product is heavily influenced by the average of prior ratings for that product (Ma et al., 2013; Zhou & Duan, 2016).

We propose that the bandwagon effect can also apply to *review voters* who vote on the helpfulness of a given review. Before reading a review, voters may observe the existing *helpful* vote ratio, a contextual signal that can influence their helpfulness judgement, independent of review content. This ratio reflects the "wisdom of crowds" and offers an accessible input for assessing the review's helpfulness. Aligning with prior votes enables voters to maintain consistency with the perceived social consensus, reducing cognitive effort and reinforcing coherence with the established voting pattern or "wisdom." In this sense, prior votes can act as informational cues that shape how voters resolve uncertainty at the moment of voting. Thus, we propose:

H4: Review voters are more likely to cast *helpful* (vs. *unhelpful*) votes on reviews that have previously received more *helpful* votes.

2.2.3 Consistency With Social Relationships

Drawing on congruity theory, we argue that review voters' behavior can also reflect social consistency or congruity, as they are likely to align their actions with their social relationships to maintain evaluative harmony and positive connections with others (Osgood & Tannenbaum, 1955).

Many online review platforms enable users to follow one another. As with other types of social media (Lee et al., 2015), this feature allows followers to receive notifications whenever the user being followed leaves a new review. A social media user with a large number of followers is often called an influencer,[1] and their opinions can significantly influence followers because of followers' trust in the influencer underlying the "following" relationship (Alibakhshi & Srivastava, 2022; Qahri-Saremi & Turel, 2023). Followers' trust in an influencer also explains the propagation and rapid diffusion of social media messages (e.g., Harvey et al., 2011). Followers also need to maintain the relationship with the influencer by conforming to the view of the influencer, especially when the influencer's opinions are known, because "agreement leads to liking" (Chaiken et al., 1996).

In the context of online reviews, a reviewer who consistently posts high-quality reviews may gain followers and develop a reputation as an influencer. When a review voter follows a reviewer, the voter is likely to trust that reviewer's opinions and feel motivated to favorably evaluate the reviewer's content to maintain consistency in the social relationship and uphold the social bond created through the "following" relationship. Trust has been shown to enhance attitudes and persuasion (Legal et al., 2012), further reinforcing this alignment. Therefore, we propose:

H5: Review voters who follow a reviewer are more likely to cast *helpful* (vs. *unhelpful*) votes on that reviewer's reviews.

[1] Influencers could be regular members who are simply trusted by others on social media platforms. These individuals are not necessarily celebrities or public figures in a traditional sense.

A voter may follow a reviewer, but the reverse can also be true: A reviewer may follow a voter, signaling interest, trust, or endorsement. Although this relationship may seem asymmetrical, balance (congruity) theory posits that individuals strive for harmonious sentiment across triads involving two people and an object (Osgood & Tannenbaum, 1955)—in this case, the voter, the reviewer, and the review. When a reviewer follows a voter, the reviewer is expressing a positive sentiment toward that voter. Because the reviewer presumably also has an implicitly positive opinion about their own review (having authored it), a *helpful* vote from the voter preserves balance in the triad. In contrast, if the voter were to rate the review as *unhelpful*, the structure would become imbalanced—a positive reviewer-voter tie, a positive reviewer-review tie, but a negative voter-review tie. To restore harmony, the voter could cast a *helpful* vote, aligning their evaluation of the review with both the reviewer's implicit stance and the social signal implied by the reviewer's decision to follow the voter. In this way, the *helpful* vote can serve as a low-cost means of maintaining evaluative balance and affirming mutual goodwill.

This prediction is also consistent with the norm of reciprocity, understood not as a transactional obligation but as a means of sustaining a coherent self-image aligned with fairness and social responsiveness (Gouldner, 1960; Matook et al., 2015). When a reviewer follows a voter, it typically signals trust, appreciation, or a willingness to engage socially—elements that contribute to the perception of a positive relationship. In line with congruity theory, the voter may experience a desire to respond in a way that upholds this positive tie and affirms their own identity as a responsive, reciprocal community member. Casting a *helpful* vote becomes a simple and visible way to reinforce mutual regard and maintain social alignment. In doing so, the voter's behavior reflects an attempt to create harmony between their social environment and their internal expectations of how such relationships ought to be managed. This alignment not only contributes to perceived consistency in social interactions but also strengthens relational cohesion within the platform's reviewer community. Thus, we propose:

H6: Review voters who are followed by a reviewer are more likely to cast *helpful* (vs. *unhelpful*) votes on that reviewer's reviews.

3 Research Context and Data

3.1 Research Context

To test our research hypotheses, we used a large-scale platform of restaurant reviews from Asia. The platform was launched in 2010. A decade after its launch, more than 3 million registered users created 1.4 million reviews of around 380,000 restaurants. The platform, which is a counterpart of Yelp in the hosting country, provided a rich dataset that included information on (1) all reviews generated during the period from January 2016 to May 2016 (e.g., review ID, reviewer ID, restaurant ID, posting date of review, textual content of review, associated star rating); (2) historical review-related activities of active users (i.e., review writers, voters, readers) since their first recorded activity during our sample period; (3) time-stamped review reading logs (i.e., review ID, reader ID, date and time of reading) [2]; (4) time-stamped vote casting activities (review ID, voter ID, date and time of a vote) and vote type (i.e., *helpful* or *unhelpful*; note that the platform requires users to log in before casting votes, and a voter can only cast one vote per review); (5) time-stamped users' following activities (i.e., follower ID, followee ID,[3] date and time of following); and (6) restaurant information (e.g., restaurant ID, category). During our sample period, 46,450 reviews were generated for 21,784 restaurants; these reviews received a total of 331,286 votes from 4,486 users.[4]

3.2 Variables and Summary Statistics

In the variable descriptions that follow, *i* indexes a vote, *j* indexes the review that receives vote *i*, *k* indexes the voter who casts vote *i, m* indexes the reviewer who writes review *j*, *n* indexes the restaurant that receives review *j*, and *t* indexes the time of vote *i*. We constructed a vote-level dataset, with each observation being a vote cast on a review generated during our study period.

Our dependent variable was $VoteHelpful_i$, indicating whether the vote was an upvote (*helpful*) or a downvote (*unhelpful*). We included three sets of main independent variables that reflect the influence of voter-related factors *at the time of voting*. The first set captured voters' historical voting tendency (i.e., the ratio of upvotes to total votes cast by a voter $VoterUpvoteRatio_{kt}$), and past review-writing experience (i.e., the number of reviews that have been written $VoterReviewCnt_{kt}$). The second set captured voters' perception of the contextual cues, where we assessed confirmation bias by measuring the deviation of a review's rating from the related restaurant's

[2] A shortened version of each review is listed on the restaurant page. A reading log is generated when a reader clicks on the review and goes to the review page to read the full version.

[3] As we discuss in our analysis section, in the case of a following relationship, the user who follows the other party is named the "follower," and the user who is being followed by the other party is called the "followee."

[4] Since we conducted a text analysis to measure reviews' readability and valence, we excluded reviews without text content (i.e., reviews with only star ratings).

average rating (which is a salient contextual cue) at the time of voting ($ReviewRatingDiff_{jnt}$) (Yin et al., 2016); we used the ratio of upvotes to total votes for a specific review ($ReviewUpvoteRatio_{jt}$) to test the bandwagon effect. The third set explored voters' social relationship with reviewers through two variables: whether the voter was following the reviewer ($VoterIsFollower_{kmt}$) and vice versa ($VoterIsFollowee_{kmt}$). We calculated all voter-related factors using data right prior to the focal vote. We also controlled for the impact of other voter characteristics, including the number of days the voter has been on the platform ($VoterTenure_{kt}$) and the voter's gender ($VoterGender_k$).

Apart from these voter-related explanatory variables, we also introduced control variables that captured the characteristics of the review, the reviewer, and the restaurant. For reviews, we considered review length ($ReviewLength_j$) and its quadratic term to capture the impact of information amount (Kuan et al., 2015; Mudambi & Schuff, 2010; Yu et al., 2022), as well as review age ($ReviewAge_{jt}$, i.e., the number of days elapsed since a review was published until it received the focal vote). Next, we applied a set of natural language processing techniques to extract textual features from the review content, including readability analysis, sentiment analysis, and topic modeling. We calculated the Gunning-Fog (GF) index, which estimates the years of formal education a person needs to comprehend a text after an initial reading, to measure readability ($ReviewReadability_j$) (e.g., Goes et al., 2014; Yin et al., 2017). A higher GF index indicates that a review is more difficult to comprehend. We derived our review valence score based on the sentiment dictionary developed by Warriner et al. (2013),[5] which is widely used in the text mining literature (e.g., Nian et al., 2021). We used the ratio of positive and negative words to total words (i.e., positive, negative, and neutral words) to measure positive valence and negative valence ($ReviewPosValence_j$, $ReviewNegValence_j$), respectively. We adopted latent Dirichlet allocation (LDA) (Blei et al., 2003) as our topic model to identify the latent topics mentioned in the review. We identified four latent topics, including food and meal, drink and dessert, service, and infrastructure and hardware. Each review was then represented by four topic proportions ($Topic1_j$, $Topic2_j$, $Topic3_j$, and $Topic4_j$), reflecting the degree to which each topic is present in the review. We report more detail on topic modeling in Appendix B.

Regarding reviewer characteristics, we incorporated variables that reflected reviewers' historical review-writing activities to control for their potential influence, including the number of reviews ($ReviewerReviewCnt_{mt}$), average review rating ($ReviewerAvgRating_{mt}$), tenure on the platform ($ReviewerTenure_{mt}$), and gender ($ReviewerGender_m$). Lastly, we accounted for restaurant characteristics, including the number of reviews ($RestReviewCnt_{nt}$), average rating ($RestAvgRating_{nt}$), and restaurant category dummies ($RestCategoryDummies_n$, indicating if a restaurant was a dine-in restaurant, express restaurant, drink and dessert, or other category). We present our variable descriptions and summary statistics in Table C1 in Appendix C.

4 Analysis and Results

4.1 Main Regression Specification and Results

We estimated a mixed-effects logistic regression by using variables described above as our independent variables and a random intercept for voters. In the following equation, Ø is the logit function, and θ_i is the probability parameter such that $E(VoteHelpful) = \theta$. δ_k is the random intercept for the voter, which considers correlated error terms for votes of the same voter (Wooldridge, 2010). We log transformed the count variables to improve the model fit.

$$
\begin{aligned}
\emptyset(\theta_i) = \alpha_0 + {} & \alpha_1 VoterUpvoteRatio_{kt} \\
& + \alpha_2 ln(VoterReviewCnt_{kt}) \\
& + \alpha_3 ReviewRatingDiff_{jnt} \\
& + \alpha_4 ReviewUpvoteRatio_{jt} \\
& + \alpha_5 VoterIsFollower_{kmt} \\
& + \alpha_6 VoterIsFollowee_{kmt} \\
& + \alpha_7 \ln(VoterTenure_{kt}) \\
& + \alpha_8 VoterGender_k \\
& + \alpha_9 \ln(ReviewLength_j) \\
& + \alpha_{10} \ln(ReviewLength_j)^2 \\
& + \alpha_{11} \ln(ReviewAge_{jt}) \\
& + \alpha_{12} ReviewReadability_j \\
& + \alpha_{13} ReviewPosValence_j \\
& + \alpha_{14} ReviewNegValence_j + \alpha_{15} Topic1_j \\
& + \alpha_{16} Topic2_j + \alpha_{17} Topic3_j \\
& + \alpha_{18} ln(ReviewerReviewCnt_{mt}) \\
& + \alpha_{19} ReviewerAvgRating_{mt} \\
& + \alpha_{20} \ln(ReviewerTenure_{mt}) \\
& + \alpha_{21} ReviewerGender_m \\
& + \alpha_{22} ln(RestReviewCnt_{nt}) \\
& + \alpha_{23} RestAvgRating_{nt} \\
& + RestCategoryDummies_n + \delta_k + \epsilon_i \quad (1)
\end{aligned}
$$

In Table 1, we present the results of our logit regressions. Column 1 incorporates review characteristics and other control variables. Consistent with previous findings (e.g., Kuan et al., 2015), we

[5] The dictionary contains 13,915 words with valence scores ranging from 1.26 to 8.48. We used this dictionary to compute the sentiment score of each word in a review. A word is defined as a positive word if the score is higher than 6, a negative word if the score is lower than 4, or a neutral word if the score is between 4 and 6. We also tried a different cutoff of the value 5 (i.e., we defined a word as positive (negative) if the score is higher (lower) than 5). When using this different cutoff value, our results stayed statistically the same.

observed a negative and statistically significant coefficient for the quadratic term of $ReviewLength_j$, suggesting that reviews of moderate length were deemed more helpful than those that were excessively brief or lengthy. We found that older reviews were not perceived as more helpful, likely due to their lack of timeliness; furthermore, difficult-to-read reviews were also less likely to be considered helpful, and review valence did not significantly impact the helpfulness vote. Regarding review topics, compared with reviews covering restaurants' infrastructure and hardware, reviews discussing food, drink and dessert, and service were more likely to receive *helpful* votes. Column 2 incorporates reviewer characteristics and other control variables. We found that reviews generated by frequent reviewers were more likely to receive *helpful* (vs. *unhelpful*) votes, likely because reviewing experience signals reviewers' credibility. Reviews created by writers who gave more favorable ratings and had joined more recently tended to receive more *helpful* votes as well. We incorporate review characteristics, reviewer characteristics, and other controls together in Column 3.

Table 1. Impacts of Voter-Related Factors on Helpfulness Vote

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: $\boldsymbol{VoteHelpful_i}$				
$VoterUpvoteRatio_{kt}$				5.141*** (0.279)	5.210*** (0.281)
$ln(VoterReviewCnt_{kt})$				0.263*** (0.038)	0.181*** (0.038)
$ReviewRatingDiff_{jnt}$				-0.325*** (0.066)	-0.142** (0.072)
$ReviewUpvoteRatio_{jt}$				2.814*** (0.449)	1.878*** (0.489)
$VoterIsFollower_{kmt}$				1.287*** (0.113)	0.990*** (0.123)
$VoterIsFollowee_{kmt}$				1.556*** (0.106)	1.218*** (0.110)
$ln(VoterTenure_{kt})$				-0.448*** (0.058)	-0.317*** (0.055)
$VoterGender_k$				0.149 (0.152)	0.162 (0.152)
$ln(ReviewLength_j)$	1.418*** (0.109)		0.737*** (0.114)		0.857*** (0.153)
$ln(ReviewLength_j)^2$	-0.294*** (0.030)		-0.203*** (0.031)		-0.226*** (0.039)
$ln(ReviewAge_{jt})$	-0.260*** (0.013)		-0.244*** (0.013)		-0.234*** (0.020)
$ReviewReadability_j$	-0.034*** (0.008)		-0.031*** (0.008)		-0.035*** (0.008)
$ReviewPosValence_j$	-0.221 (0.331)		0.616* (0.331)		0.286 (0.433)
$ReviewNegValence_j$	-1.241 (0.860)		-0.198 (0.846)		1.624 (1.151)
$Topic1_j$	0.423** (0.135)		0.107*** (0.135)		-0.067 (0.171)
$Topic2_j$	0.775*** (0.108)		0.418*** (0.111)		0.229* (0.139)
$Topic3_j$	0.896*** (0.151)		0.560*** (0.152)		0.456** (0.186)
$ln(ReviewerReviewCnt_{mt})$		0.440*** (0.015)	0.378*** (0.018)		0.219*** (0.025)
$ReviewerAvgRating_{mt}$		0.364*** (0.056)	0.265*** (0.058)		0.278*** (0.084)
$ln(ReviewerTenure_{mt})$		-0.079*** (0.020)	-0.081*** (0.020)		-0.066** (0.027)
$ReviewerGender_m$		-0.032 (0.062)	-0.067 (0.063)		-0.223*** (0.079)
$ln(RestReviewCnt_{nt})$	0.132*** (0.020)	0.030 (0.019)	0.141*** (0.020)	0.012 (0.024)	0.036 (0.026)
$RestAvgRating_{nt}$	-0.309*** (0.057)	-0.192*** (0.062)	-0.164*** (0.062)	-0.003 (0.066)	0.049 (0.074)
$RestCategoryDummies_n$	Yes	Yes	Yes	Yes	Yes
Constant	6.739*** (0.318)	3.497*** (0.328)	4.221*** (0.360)	-0.190 (0.704)	-1.674** (0.771)
Pseudo-R^2	0.052	0.053	0.077	0.173	0.196
Observations	331,286	331,286	331,286	331,286	331,286

Note: We included a random intercept at the voter level. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Column 4 shows the result of the model that incorporated voter-related factors and control variables. First, we observed a positive correlation between a voter's past upvote ratio ($VoterUpvoteRatio_{kt}$) and their likelihood of casting an upvote, indicating that a voter's history of favorable evaluations is positively associated with their current upvoting behavior; thus, H1 is supported. Specifically, the coefficient of 5.141 indicates that a 1% increase in $VoterUpvoteRatio_{kt}$ was associated with an increase of 5.3% ($e^{5.141/100}$ - 1) in the odds of casting a *helpful* (vs. *unhelpful*) vote. We also found that users who had written many reviews in the past were more likely to cast *helpful* (vs. *unhelpful*) votes on others' reviews, likely because they consider other reviewers to be part of their in-group. The coefficient indicates that a 1% increase in $VoterReviewCnt_{kt}$ was associated with around a 0.26% increase in the odds of upvoting, supporting H2. These results suggest that voters tended to seek consistency with their past actions to avoid cognitive dissonance.

Second, we observed a negative and significant coefficient for $ReviewRatingDiff_{jnt}$. Namely, voters were less likely to cast *helpful* (vs. *unhelpful*) votes on reviews whose ratings deviated more from the restaurant's average rating, which would have presumably served as the basis of the voter's initial belief. Specifically, the odds ratio of casting a *helpful* vote for a one-unit increase in $ReviewRatingDiff_{jnt}$ is 0.72 times as large. Therefore, H3 is supported. We also found that voters were more likely to cast favorable votes on reviews with a higher ratio of past *helpful* votes (i.e., a 1% increase in $ReviewUpvoteRatio_{jt}$ was associated with an increase of 3% in the odds of casting a *helpful* vote), supporting the bandwagon effect proposed in H4. These results suggest that voters also sought contextual consistency by aligning their judgments with contextual cues present in the environment.

Lastly, we observed that both $VoterIsFollower_{kmt}$ and $VoterIsFollowee_{kmt}$ had statistically significant and positive impacts on $VoteHelpful_i$. The coefficients indicate that the tendency of a focal voter who was following (followed by) a reviewer to upvote that reviewer's review was 3.62 (4.74) times the odds of the tendency of a voter without a social relationship. The results lend support to H5 and H6. These results suggest that voters aligned their voting behavior with their social relationships to maintain coherence and positive connections with others.

In addition to observing the significant impacts of voter-related factors, we see that the pseudo-R^2 in Column 4, 0.173, is more than triple that in Columns 1 and 2, and more than double that in Column 3, indicating that voter-related factors may play a more important role in a review's helpfulness evaluation than review characteristics and reviewer characteristics. To further compare the effect sizes of different types of voter-related variables, we incorporated all the independent variables in Column 5 and conducted Wald tests (Sun et al., 2020). Our results show that variables capturing voters' past voting behavior ($VoterUpvoteRatio_{kt}$), bandwagon effect ($ReviewUpvoteRatio_{jt}$), and social relationship (i.e., $VoterIsFollower_{kmt}$ and $VoterIsFollowee_{kmt}$) had significantly larger effects on $VoteHelpful_i$ than review and reviewer characteristics. There was no significant difference in the effect size between variables $ReviewRatingDiff_{jnt}$, $VoterReviewCnt_{kt}$ and those reflecting review and reviewer characteristics.

4.2 Addressing Endogeneity Issues

Our main analysis may have suffered from endogeneity issues. In this subsection, we discuss distinct identification strategies for each of the three groups of voter-related factors, tailored to address the specific confounding factors that may obscure their effects.

4.2.1 Identification Strategy for Voters' Behavioral Tendencies

First, the two variables that capture voters' behavioral tendencies of voting and reviewing ($VoterUpvoteRatio_{kt}$ and $VoterReviewCnt_{kt}$) may have suffered from endogeneity issues caused by unobserved voter characteristics (e.g., their overall commitment to the platform and desire to build reputations). These unobserved characteristics could have simultaneously impacted voters' previous review-writing and voting experience, and their current voting behavior.

To address this issue, we adopted the following identification strategy. We exploited the fact that each voter's behavioral tendencies varied over time. For example, a voter's $VoterUpvoteRatio_{kt}$ (the fraction of their total votes that are *helpful* votes) may have started below the sample median but then may have risen above it if the voter continued casting many *helpful* votes. Similarly, a voter's $VoterReviewCnt_{kt}$ (the total number of reviews they have written) could have transitioned from below to above the median as they wrote more reviews.

Accordingly, we classified all reviews voted on by the same voter at different time points into *low* and *high* groups. The *low* group included reviews that the voter evaluated at a point when their behavioral tendency was below the sample median. The *high* group included reviews evaluated when that same voter's tendency exceeded the median. Then, we estimated the impact of voter behavioral tendencies by comparing the voter's likelihood of casting *helpful* (vs. *unhelpful*) votes for reviews in different groups. By comparing voting outcomes *within* a single voter across these two groups (i.e., *low* vs. *high*), we held constant any unobserved, time-invariant characteristics of that voter. Figure 2 illustrates the idea.

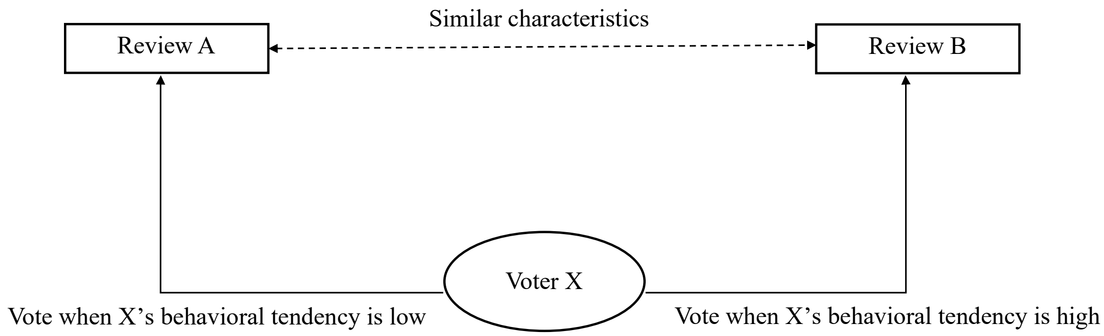


Figure 2. Illustration of Identification Strategy for Voters' Behavioral Tendencies

Table 2. Impact of *HighVoterUpvoteRatio* on Helpfulness Vote

	Dependent variable: $VoteHelpful_{kjt}$
$HighVoterUpvoteRatio_{kjt}$	2.467*** (0.265)
Voter fixed effects	Yes
Constant	5.380*** (0.088)
Observations	7,662
Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	

Table 3. Impact of *HighVoterReviewCnt* on Helpfulness Vote

	Dependent variable: $VoteHelpful_{kjt}$
$HighVoterReviewCnt_{kjt}$	0.281*** (0.094)
Voter fixed effects	Yes
Constant	6.074*** (0.061)
Observations	6,170
Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	

Specifically, we identified all voters who cast votes for reviews in the sample of the main analysis and tracked their voting behavior during our study period (i.e., January 2016 to May 2016). We split reviews that received votes from the same voter into two groups based on the median value of the two behavioral tendencies, respectively.

Next, to ensure that any observed difference in voters' likelihood to cast a *helpful* (vs. *unhelpful*) vote was not simply driven by the types of reviews in each group, we employed propensity score matching (PSM). Specifically, we matched each review in the *high* behavioral tendency group to a review from the *low* group based on a set of review characteristics (i.e., length, readability, positive valence, negative valence, age, and topic distribution of the review) (Caliendo & Kopeinig, 2008; Dehejia & Wahba, 2002). We conducted one-to-one matching with replacement and a caliper size of 0.001 times the standard deviation of the propensity scores. Table D1 and Table D2 in Appendix D present the results of the balance check after matching, which indicate that the two groups of reviews were successfully matched for both voter behavioral tendencies.

Lastly, we rearranged our dataset into voter-time panel datasets, and we estimated the impact of different levels of voter behavioral tendencies using logit models with voter fixed effects. This enabled the model to only use within-voter variation and control for the impact of unobserved voter characteristics. The dependent variable was $VoteHelpful_{kjt}$, indicating whether voter k cast a *helpful* (vs. *unhelpful*) vote on review j at time t.[6] The explanatory variables were dummy variables $HighVoterUpvoteRatio_{kjt}$ and $HighVoterReviewCnt_{kjt}$, indicating whether the voter was at the *high* level of the two behavioral tendencies when voting. Tables 2 and 3 show that the coefficients of both $HighVoterUpvoteRatio_{kjt}$ and $HighVoterReviewCnt_{kjt}$ are positive and significant (p < 0.01), implying that voters were more likely to cast *helpful* (vs. *unhelpful*) votes when they had cast more favorable votes previously or when they had more review-writing experience. These results lend further support to H1 and H2.

[6] For the identification analysis, we used kjt subscripts to emphasize within-voter variation across reviews and voting times; each kjt observation corresponds to a unique vote i.

4.2.2 Identification Strategy for Contextual Cues

Although we controlled for a set of observed review characteristics in the main analyses, some variables reflecting the review context (such as the average rating of the restaurant and the review's accumulated helpfulness votes) may still have suffered from endogeneity problems caused by unobserved review characteristics, such as review quality. Here we focused on two context-related variables: $ReviewUpvoteRatio_{jt}$ and $ReviewRatingDiff_{jnt}$. For instance, high-quality reviews that previously received a greater portion of *helpful* votes were more likely to receive another *helpful* vote from the focal voter, and it is possible that this relationship was not caused by the bandwagon effect.

To address this issue, we adopted the following identification strategy. For each review, we identified two groups of voters who both voted for the same review but at different time points so that they would have faced different levels of contextual cues at the time of voting. Since different groups of voters voted for the same review, the unobserved review characteristics would remain the same. Thus, if we then compared the likelihood of the two groups of voters to cast *helpful* votes for the focal review, the difference between them, if any, could be attributed to the contextual cues. Figure 3 illustrates this identification strategy.

Specifically, we observed voters' behavior within 30 days after the focal review was published. For each review, we split its voters into two groups based on the median value of the two contextual cues they were observing, respectively. Accordingly, one group of voters faced a low value (i.e., below the median) of a certain contextual cue, and the other group faced a high value (i.e., above the median) of that cue. To eliminate the effect driven by voter heterogeneity, we again employed PSM to match each voter in the high-value group with a voter from the low-value group based on a set of voter characteristics such that these two matched groups of voters were very similar to each other. We adopted voters' tenure on the platform, voters' gender, the number and average rating of reviews generated by the voters, the number of votes cast by the voters, and the number of followers and followees of the voters prior to the focal vote to generate the propensity score. We conducted one-to-one matching with replacement and a caliper size of 0.001 times the standard deviation of the propensity scores. Tables D3 and D4 in Appendix D show that the two groups of voters were successfully matched for both contextual cues.

Next, we rearranged our dataset into review-time panel datasets and then compared the likelihood of casting *helpful* (vs. *unhelpful*) votes on a focal review of the matched groups of voters. Specifically, we conducted logit regression with review fixed effects, which allowed us to use only the within-review variation in contextual cues and to control for unobserved review heterogeneity as confounders. The dependent variable was $VoteHelpful_{jkt}$, indicating whether review j received a *helpful* vote from voter k at time t. The explanatory variables were the dummy variables $HighReviewRatingDiff_{jknt}$ and $HighReviewUpvoteRatio_{jkt}$, which indicated whether the voter observed a higher level of those cues. As shown in Table 4, the coefficient of $HighReviewRatingDiff_{jknt}$ is negative and significant, indicating that when observing a high level of the absolute difference between the focal review rating and the average restaurant rating, voters were less likely to cast *helpful* votes. The coefficient of $HighReviewUpvoteRatio_{jkt}$ presented in Table 5 is positive and significant, indicating that voters were more likely to cast *helpful* (vs. *unhelpful*) votes if they observed a higher ratio of upvotes received by the review. These results provide additional evidence for H3 and H4.

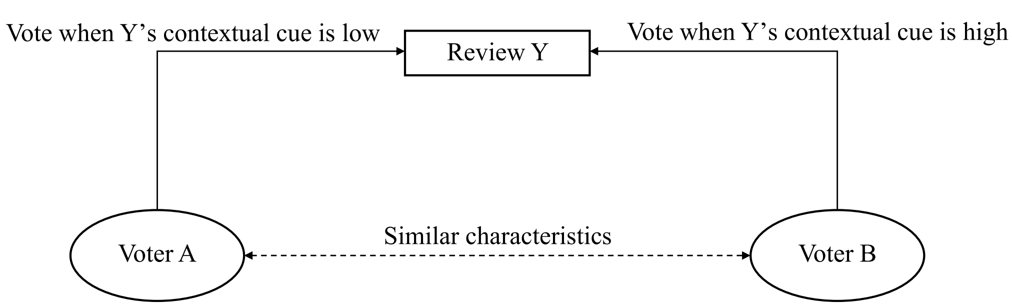


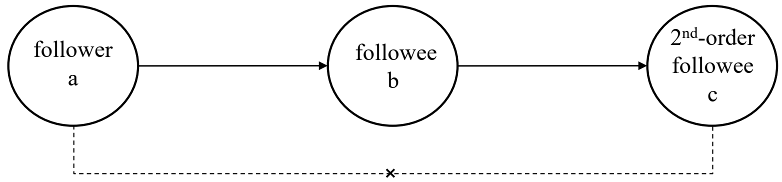
Figure 3. Illustration of Identification Strategy for Review Contextual Cues

Table 4. Impact of *HighReviewRatingDiff* on Helpfulness Vote

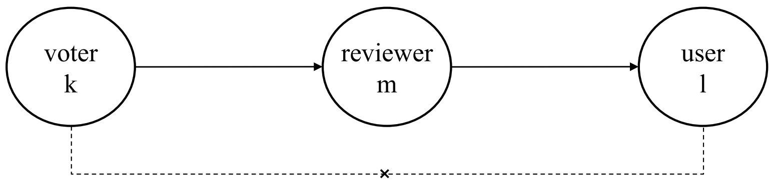
	Dependent variable: $VoteHelpful_{jkt}$
$HighReviewRatingDiff_{jkt}$	-3.341*** (1.018)
Review fixed effects	Yes
Constant	8.047*** (1.000)
Observations	6,248
Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	

Table 5. Impact of *HighReviewUpvoteRatio* on Helpfulness Vote

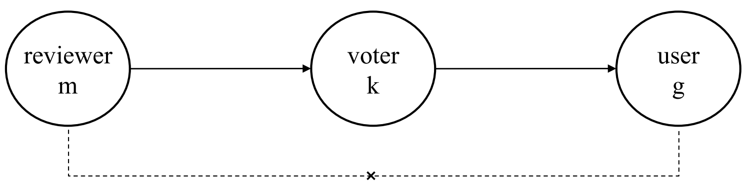
	Dependent variable: $\boldsymbol{VoteHelpful_{jkt}}$
$HighReviewUpvoteRatio_{jkt}$	1.428*** (0.422)
Review fixed effects	Yes
Constant	4.751*** (0.187)
Observations	6,766
Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	



Panel A General Illustration



Panel B Illustration for $\boldsymbol{VoterIsFollower_{kmt}}$



Panel C Illustration for $\boldsymbol{VoterIsFollowee_{kmt}}$

Figure 4. Illustration of the Second-Order Followee

4.2.3 Instrumental Variable Method for Follower-Followee Relationships

Since the follower-followee network is endogenously formed, the observed impact of social relationships on helpfulness vote casting may have been caused by unobserved characteristics that affect users' voting behavior and their following behavior at the same time. For example, users who share a similar taste in food or restaurants (therefore holding similar opinions towards certain restaurants) would presumably be more likely to form a follower-followee relationship. To address this issue, we adopted the instrumental variable method.

Consistent with previous literature, to construct instruments for the follower-followee relationships, we used the average characteristics of a follower's second-order followees who did not have any direct social relationship with the follower (Wei et al., 2021). Valid instruments were excluded variables in the voting equation (i.e., Equation 1) that correlated with the follower-followee relationships but were uncorrelated with the unobservable characteristics. To show why our instruments are valid, we illustrate a parsimonious network with only three users, *a*, *b*, and *c*, and their following relations in Panel A of Figure 4 (i.e., *a* is a follower of *b*, *b* is therefore a followee of *a*, and *c* is a second-order followee of *a*). Since *c* is *b*'s immediate followee, the characteristics of *c* will influence *b*'s review-related behavior, which in turn will determine *a*'s decision to follow *b*. This variable is therefore relevant. Meanwhile, since *c* is not *a*'s immediate followee, it is highly unlikely that *c*'s characteristics will directly influence *a*'s voting decision on *b*'s review. This variable, therefore, satisfies the instrument exogeneity condition. Accordingly, for variable $VoterIsFollower_{kmt}$, the instrument is the average characteristics of the second-order followee *l* (see Panel B, Figure 4), and for $VoterIsFollowee_{kmt}$, the instrument is the average characteristics of the second-order followee *g* (see Panel C, Figure 4).

The characteristics of a user's second-order followees included demographic characteristics (i.e., gender, tenure), review generating behavior (i.e., the number of reviews and the average of ratings generated), vote casting behavior (i.e., the number of upvotes and downvotes), and following behavior (i.e., the number of followers and followees). For each observation, we calculated the average of the second-order followees' characteristics.

Table 6. Results of Instrumental Variable Analysis

	Dependent variable: $VoteHelpful_i$
$VoterIsFollower_{kmt}$	0.285*** (0.036)
$VoterIsFollowee_{kmt}$	0.267*** (0.034)
Review characteristics	Yes
Reviewer characteristics	Yes
Restaurant characteristics	Yes
Constant	1.215*** (0.259)
Observations	263,405

Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7. Impacts of Voter-Related Factors on Helpfulness Vote (Second-Stage Estimates from the Two-Stage Model)

	Dependent variable: $VoteHelpful_i$
$VoterUpvoteRatio_{kt}$	6.375*** (0.829)
$ln(VoterReviewCnt_{kt})$	0.233*** (0.031)
$ReviewRatingDiff_{jnt}$	-0.242*** (0.063)
$ReviewUpvoteRatio_{jt}$	1.222*** (0.390)
$VoterIsFollower_{kmt}$	1.199*** (0.111)
$VoterIsFollowee_{kmt}$	0.825*** (0.103)
Control variables	Yes
Inverse Mills ratio	0.088 (0.087)
Constant	-4.584*** (1.047)
Observations	331,286

Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Since both the dependent variable and the endogenous variables were dummy variables, we adopted the multivariate probit model for our instrument analysis. The results reported in Table 6 show that both $VoterIsFollower_{kmt}$ and $VoterIsFollowee_{kmt}$ had positive impacts on the likelihood of the voter casting a *helpful* (vs. *unhelpful*) vote on the focal review. They also suggest that ignoring the endogeneity of a voter's following activities (Column 5, Table 1) would have led to an overestimation of the impact of engaging in social relationships, which was probably caused by a positive correlation between the endogenous follower-followee relationship and the error term.

4.2.4 Two-Stage Correction for Selection Bias

In addition to those above-mentioned omitted variable biases, our main analysis on review helpfulness may have also suffered from a selection bias. The selection problem could have arisen because some unobserved factors that affected the selection of a review to be voted on (vs. not voted on) may have also affected the review's likelihood of being voted *helpful* (vs. *unhelpful*), as reflected in the outcome variable. To address this issue, we modeled a two-stage process as a robustness check. In the first stage, we modeled the vote-or-not decision. Here the dependent variable was $Vote_{it}$, indicating if the review reader cast a vote after reading the review. We incorporated the same set of independent variables used in the main analysis and estimated a mixed-effects logit regression. We then calculated the inverse Mills ratio and incorporated it into the second-stage upvote or downvote decision (i.e., the dependent variable is $VoteHelpful_i$) to correct for the selection bias. The results shown in Table 7 are qualitatively the same as those from our main analysis.

4.2.5 Additional Robustness Checks

To ensure the robustness of our findings, we performed additional checks. These included addressing data imbalance (i.e., the dominance of upvotes in the dependent variable) through undersampling and oversampling strategies and employing an alternative approach to model voter decisions. The results of these analyses remained consistent with our main findings. Details are provided in Appendix E.

4.3 Mechanism Analysis

Next, we empirically tested whether the voter-related factors impacted helpfulness vote casting through the mechanisms proposed in the theory development. Specifically, we conducted a moderating effect analysis for each of the factors.

H1 posited that voters would align their voting behavior with their past voting patterns to maintain cognitive consistency and avoid dissonance. To test this mechanism, we used voters' vote-casting frequency, measured by the average voting time intervals between two consecutive upvotes ($AvgUpvotingTimeInt_{kt}$), as a moderator of the relationship between the previous upvote ratio and the current likelihood to cast a *helpful* (vs. *unhelpful*) vote. The underlying premise is that increased frequency in behavior strengthens habitual tendencies to maintain consistency, thereby enhancing the likelihood of replicating past actions (Ajzen & Fishbein, 2000). In this case, we propose that the impact of past voting will be greater among voters who cast upvotes more frequently (i.e., in smaller time intervals).

In H2, we posited that as voters gain more review-writing experience, their identity as reviewers will strengthen, making them more likely to vote favorably on reviews to maintain cognitive consistency with their evolving reviewer identity. The moderator that we used is the voter's tenure on the platform ($VoterTenure_{kt}$). Previous literature has shown that in the working context, job tenure is positively associated with employees' strength of identity with their team and organization (Barker & Tompkins, 1994; Hall et al., 1970). Similarly, in our context, the longer a user stays on the platform, the stronger their reviewer identity is. We, therefore, propose that voters' tenure will strengthen the effect of the review writing experience, namely, $VoterReviewCnt_{kt}$.

To test the mechanism of confirmation bias posited in H3, we introduced the dispersion of restaurant ratings as a moderator ($RestRatingDispersion_{nt}$, measured by the standard deviation of the ratings received by the focal restaurant). Confirmation bias arises when an individual's existing beliefs are challenged. Specifically, a rating that significantly diverges from the average, which informs initial perceptions about the restaurant, would tend to increase discomfort and the likelihood of unfavorable voting. However, a wider dispersion in ratings, indicating varied opinions, would likely undermine confidence in the average rating as a reliable quality indicator (Petrocelli et al., 2007). Hence, we propose that rating dispersion will attenuate confirmation bias.

We used the total number of *helpful* votes received by a review ($ReviewUpvoteRcvd_{jt}$) as a moderator to test the mechanism of the bandwagon effect posited in H4. The larger the *helpful* vote counts, the more confident the voter would be about the consensus of prior votes, which would thus make it even more likely that the voter would vote consistently with previous voters. Thus, we propose that the total number of *helpful* votes will strengthen the positive impact of the review's helpfulness ratio before a focal vote.

We posited that follower voters would tend to favorably evaluate their followees' reviews because of trust (H5), and followee voters would tend to favorably evaluate their followers' reviews due to reciprocity (H6). These behaviors reflect a desire for consistency in maintaining social relationships, as voters align their evaluations with the expectations and norms of their social connections. To measure people's trust toward a user, we used the number of performance-based badges obtained by the user ($ReviewerBadgeCnt_{mt}$ and $VoterBadgeCnt_{kt}$). The focal platform awards badges based on users' reviewing performance to showcase high quality contributions. Therefore, the higher the badge counts, the greater the level of trust from the follower to the followee. Note that our arguments for H5 rely on trust from the voter (as follower) to the reviewer, so we expected the reviewer's badge counts to moderate the effect of $VoterIsFollower_{kmt}$. However, our arguments for H6 are based on reciprocity, not trust per se, so we did not expect the voter's badge counts to moderate the effect of $VoterIsFollowee_{kmt}$.

The results of the mechanism analysis are shown in Table 8. We obtained the expected moderating effects for the first four hypotheses. Additionally, we observed a significant coefficient on the interaction term $VoterIsFollower_{kmt} \times ln(ReviewerBadgeCnt_{mt})$, but, as predicted, the interaction term $VoterIsFollowee_{kmt} \times ln(VoterBadgeCnt_{kt})$ did not attain significance. Together, these two results thus lend support to the mechanisms of H5 and H6.

4.4 Review-Level Analyses

Our vote-level analysis demonstrates that voter-related factors significantly influence the casting of *helpful* and *unhelpful* votes. This leads to a natural question: Do these factors, when aggregated at the review level, cancel each other out or continue to impact the overall helpfulness ratio? We speculate that the answer is complex. For instance, the bandwagon effect suggests that voting behavior is interdependent, with each vote following a trend of previous votes and reinforcing that trend. Therefore, we would expect this effect to persist at the review level. However, opposing tendencies, such as the favorable votes of experienced users versus the unfavorable votes of newcomers (H2), might neutralize each other when aggregated. To investigate this, we conducted supplementary analyses to determine whether the impact of voter-related factors is still evident at the aggregated review level.

4.4.1 Review-Level Cross-Sectional Analysis

First, we conducted a cross-sectional analysis in which we pooled reviews in our sample together and examined votes received within 30 days post-publication. Consistent with previous literature (e.g., Mudambi & Schuff, 2010), the dependent variable was the ratio of *helpful* to total votes. For each independent variable, we calculated values at the individual vote level using pre-vote data and then averaged these across all votes on a review (see Table F1, Panel A in Appendix F). The rationale for our variable operationalization is that each voter faces specific factors before voting; we therefore calculated those variables at the voter level and aggregated them to the review level to examine their collective impact.

Table 8. Mechanism Analysis

	Dependent variable: $VoteHelpful_i$
$VoterUpvoteRatio_{kt} \times ln(AvgUpvotingTimeInt_{kt})$	-0.377*** (0.107)
$ln(VoterReviewCnt_{kt}) \times ln(VoterTenure_{kt})$	0.113*** (0.020)
$ReviewRatingDiff_{jnt} \times RestRatingDispersion_{nt}$	0.500** (0.199)
$ReviewUpvoteRatio_{jt} \times ln(ReviewUpvoteRcvd_{jt})$	0.331*** (0.072)
$VoterIsFollower_{kmt} \times ln(ReviewerBadgeCnt_{mt})$	0.189** (0.094)
$VoterIsFollowee_{kmt} \times ln(VoterBadgeCnt_{kt})$	0.123 (0.100)
$VoterUpvoteRatio_{kt}$	5.971*** (0.336)
$ln(VoterReviewCnt_{kt})$	1.957*** (0.316)
$ReviewRatingDiff_{jnt}$	0.160 (0.194)
$ReviewUpvoteRatio_{jt}$	1.274*** (0.391)
$VoterIsFollower_{kmt}$	0.691*** (0.181)
$VoterIsFollowee_{kmt}$	1.143*** (0.216)
$ln(AvgUpvotingTimeInt_{kt})$	0.158 (0.102)
$ln(ReviewUpvoteRcvd_{jt})$	-0.014*** (0.004)
$RestRatingDispersion_{nt}$	0.312* (0.163)
$ln(ReviewerBadgeCnt_{mt})$	-0.193*** (0.046)
$ln(VoterBadgeCnt_{kt})$	0.018 (0.047)
Control variables	Yes
Constant	-2.280*** (0.764)
Observations	301,284
Note: We included a random intercept at the voter level. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.	

Since the value of the dependent variable, $Helpful\%_{j(30d)}$, is between 0 and 1, we estimated a fractional response model with a logit link function (Papke & Wooldridge, 1996, 2008) and the following conditional mean specification:

$$E(y|\boldsymbol{x}) = G(\boldsymbol{x\beta} + \beta_0), \tag{2}$$

in which y is the dependent variable $Helpful\%_{j(30d)}$; G(·) is the logit function; $\boldsymbol{x}$ is a vector of explanatory variables, including voter-related factors, review characteristics, reviewer characteristics, and restaurant characteristics; and $\boldsymbol{\beta}$ is a vector of parameters to be estimated. Note that our cross-sectional analysis could not capture the bandwagon effect by using *ReviewUpvoteRatio*, which is dynamically changing over time; we address this limitation in the next panel data analysis.

We first conducted a Ramsey regression equation specification error (RESET) test to check the functional form misspecification for our conditional mean model in Equation (2) and a generalized goodness of functional form (GGOFF) test to check the link function form (Ramalho et al., 2011, 2014). Neither the RESET test (RESET = 2.108, p = 0.147) nor the GGOFF test (GGOFF = 1.434, p = 0.231) was able to reject the null hypothesis at an alpha level of 5%, indicating that our model was correctly specified. We present the result of our fractional logit model in Table 9. In the first three columns, we show the results of regressions, which include review characteristics, reviewer characteristics, separately and together, and control variables.

Column 4 includes voter-related factors aggregated at the review level and control variables. We included all variables in our regression reported in Column 5. Consistent with our vote-level finding (Table 1), we observe significant positive coefficients for $VoterUpvoteRatio_{k(30d)}$ and $Ln(VoterReviewCnt_{k(30d)})$, indicating the tendency to vote in alignment with past behavior for consistency. The negative coefficient of $ReviewRatingDiff_{jn(30d)}$ highlights the significant influence of the tendency to vote in alignment with existing beliefs shaped by contextual cues. Additionally, we note that reviews voted on by users who engaged in more following activities were more likely to obtain higher helpfulness ratios, reflecting a tendency to maintain social consistency. Furthermore, we note that the pseudo-R^2 of the regression in Column 4 is much larger than that of the first three columns, suggesting that even at the review level, voter-related factors still played a more important role in determining reviews' helpfulness ratio.

4.4.2 Review-Level Panel Data Analysis

The cross-sectional analysis above provides initial evidence that the impacts of voter-related factors also exist at the review level, but it has some limitations. First, since we were unable to measure *ReviewUpvoteRatio* using cross-sectional data, our previous analysis could not examine the bandwagon effect. Second, there is a potential reverse causality issue between helpfulness ratios and following activities, as users may have followed reviewers with high helpfulness ratios. Third, while we considered several review characteristics, other unobserved, time-invariant factors might have affected helpfulness. To overcome these limitations, we next conducted a panel data analysis.

We selected reviews generated in January 2016 and constructed a balanced panel dataset spanning four months (from February 2016 to May 2016). The dependent variable was the helpfulness ratio obtained by the focal review up to week *w*. For time-varying characteristics, we first calculated each of the variables at the vote level using data prior to each vote received by the focal review. Then, we averaged these across votes cast until week *w*-1. We also calculated reviewer and restaurant characteristics using data until week *w*-1 (see Table F1, Panel B).

Table 9. Impact of Voter-Related Factors: Results From Review-Level Cross-Sectional Analyses

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: $Helpful\%_{j(30d)}$				
$VoterUpvoteRatio_{k(30d)}$				4.970*** (0.114)	5.067*** (0.123)
$ln(VoterReviewCnt_{k(30d)})$				0.073*** (0.008)	0.061*** (0.008)
$ReviewRatingDiff_{jn(30d)}$				-0.062*** (0.019)	-0.040** (0.019)
$ln(Follower_{m(30d)})$				0.031*** (0.004)	-0.019*** (0.006)
$ln(Followee_{m(30d)})$				0.030*** (0.006)	0.023*** (0.006)
$ln(VoterTenure_{k(30d)})$				0.052*** (0.020)	0.056*** (0.021)
$MaleVoterRatio_{k(30d)}$				0.184*** (0.025)	0.186*** (0.026)
$ln(ReviewLength_j)$	0.259*** (0.064)		0.169** (0.071)		-0.084 (0.081)
$ln(ReviewLength_j)^2$	-0.018*** (0.005)		-0.016*** (0.006)		0.002 (0.006)
$ln(ReviewAge_{j(30d)})$	-0.049*** (0.010)		-0.062*** (0.010)		-0.042*** (0.010)
$ReviewReadability_j$	-0.005** (0.003)		-0.004 (0.003)		-0.003 (0.002)
$ReviewPosValence_j$	-0.101 (0.102)		-0.107 (0.106)		0.080 (0.105)
$ReviewNegValence_j$	-0.501** (0.225)		-0.323 (0.228)		0.345 (0.264)
$Topic1_j$	0.110*** (0.037)		0.081** (0.038)		0.047 (0.041)
$Topic2_j$	0.159*** (0.030)		0.113*** (0.031)		0.038 (0.033)
$Topic3_j$	0.151*** (0.041)		0.110*** (0.042)		0.076* (0.043)
$ln(ReviewerReviewCnt_{m(30d)})$		0.071*** (0.005)	0.072*** (0.005)		0.100*** (0.008)
$ReviewerAvgRating_{m(30d)}$		0.066*** (0.017)	0.064*** (0.017)		0.047*** (0.017)
$ln(ReviewerTenure_{m(30d)})$		0.002 (0.007)	0.004 (0.007)		0.002 (0.007)
$ReviewerGender_m$		0.064*** (0.018)	0.061*** (0.018)		0.028 (0.019)
$ln(RestReviewCnt_{n(30d)})$	0.005 (0.006)	0.005 (0.006)	0.009 (0.006)	0.009 (0.006)	0.009 (0.006)
$RestAvgRating_{n(30d)}$	0.009 (0.016)	0.024 (0.016)	0.027* (0.016)	0.009 (0.016)	0.025 (0.017)
$RestCategoryDummies_n$	Yes	Yes	Yes	Yes	Yes
Constant	0.801*** (0.216)	1.025*** (0.099)	0.609** (0.244)	-4.266*** (0.184)	-4.364*** (0.340)
Pseudo-R^2	0.048	0.045	0.056	0.284	0.309
Observations	46,450	46,450	46,450	46,450	46,450

Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10. Impact of Voter-Related Factors: Results From Review-Level Panel Data Analysis

	Dependent variable: $\boldsymbol{Helpful\%_{j(w)}}$
$VoterUpvoteRatio_{k(w-1)}$	1.906*** (0.209)
$ln(VoterReviewCnt_{k(w-1)})$	0.009 (0.023)
$ReviewRatingDiff_{jn(w-1)}$	-0.146** (0.070)
$ReviewUpvoteRatio_{j(w-1)}$	4.935*** (0.303)
$ln(Follower_{m(w-1)})$	0.151** (0.068)
$ln(Followee_{m(w-1)})$	0.184** (0.083)
Voter-related control variables	Yes
Review characteristics	Yes
Reviewer characteristics	Yes
Restaurant characteristics	Yes
Time-invariant controls	Yes
Constant	-2.761** (1.149)
Observations	45,614

Note: We report only the coefficients of the explanatory variables that vary by reviews and weeks. Since we control for the impact of time-invariant variables, we do not include a separate column that incorporates only review characteristics and control variables. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Papke and Wooldridge (2008) show that time-invariant, unobserved effects correlated with explanatory variables can be controlled by adding the time averages of the explanatory variables to the conditional mean model. Therefore, we estimated a fractional logit regression for our balanced panel data with the following conditional mean specification:

$$E(y_{jw}|\boldsymbol{z_j}) = G(\boldsymbol{z}_{jw}\boldsymbol{\gamma} + \boldsymbol{\bar{z}}_j\boldsymbol{\xi} + \psi), \quad (3)$$

in which y_{jw} is the dependent variable $Helpful\%_{jw}$; G(·) again is the logit link function; $\boldsymbol{z}_{jw}$ is a vector of time-varying explanatory variables, including voter-related factors, review age, reviewer and restaurant characteristics; $\boldsymbol{\bar{z}}_j$ is a vector of time averages of the explanatory variables; and $\boldsymbol{\gamma}$ is a vector of parameters to be estimated. We estimated the model using a generalized estimating equation (GEE) approach (Papke & Wooldridge, 2008).

We present the results of our panel data fractional logit regression in Table 10. We note that after we controlled for the impact of time-invariant review characteristics, voter-related factors still exhibited significant influences on helpfulness ratios, except for voters' review writing experience, probably because consumers did not typically write reviews on a weekly basis. These results are largely consistent with our cross-sectional, review-level analysis (Table 9) and our main analysis at the vote level (Table 1). In summary, our analyses demonstrate the systematic impact of voters' tendency to maintain cognitive consistency on helpfulness vote casting at both the individual vote level and the aggregated review level.

5 Discussion

This study broadens the scope of online review helpfulness research by highlighting the largely overlooked role of voter-related factors, thereby extending beyond the review- and reviewer-centric approaches commonly found in the literature (Hong et al., 2017). Drawing on cognitive consistency theories, we analyzed a fine-grained, time-stamped vote-level dataset to show how voters' past behavior, contextual cues, and social relationships collectively shape the likelihood of casting *helpful* (vs. *unhelpful*) votes. In doing so, we extend existing insights on consistency-driven biases (e.g., Yin et al., 2016) by demonstrating that voter influences can outweigh the effects of either review or reviewer characteristics alone. This underscores the critical need to account for the voter perspective and the cognitive processes behind it when examining review helpfulness.

5.1 Theoretical Implications

First, our study advances online review research by highlighting the critical—but largely overlooked—role of voters and demonstrating that their characteristics can be even more predictive of review helpfulness than the review- or reviewer-level factors foregrounded in prior IS work. While previous studies have implied, based on review-level analyses, that voter-related factors might influence helpfulness, they have not directly tested this proposition (e.g., Yin et al., 2016). By addressing this omission with fine-grained vote-level data, our findings move the literature beyond its traditional two-part focus—review content and reviewer traits—and demonstrate that a comprehensive model of review helpfulness needs to account for three parts around review content: the voter, the reviewer, and the review's context.

Second, whereas we draw on cognitive consistency theories as our overarching lens (Chaiken et al., 1989; Festinger, 1962; Kunda, 1990; Osgood & Tannenbaum, 1955), we also extend this tradition in several nontrivial ways. We shift the unit of analysis from intrapersonal attitude adjustment to a multi-part crowd-evaluation setting, positioning the voter, the reviewer, and the review context as a triadic system. We also introduce a three-dimensional framework—behavioral, contextual, and social consistency—that specifies how distinct cues activate consistency-seeking in online review settings. By testing these mechanisms with fine-grained, behavioral vote-level data, we show how individual alignment motives aggregate into systematic crowd-level bias in review rankings. Collectively, these contributions extend cognitive consistency theories beyond their traditional attitude-change roots and illuminate their value for explaining platform dynamics. Our findings also call for additional research on how prospective consumers manage consistency when evaluating and seeking reviews (Liu et al., 2019; Yin et al., 2023).

Building on this three-dimensional framework, we highlight one dimension in particular—contextual cues—and show how prior helpfulness votes shape later evaluations. Whereas previous research (e.g., Tseng et al., 2023) has examined this linkage at the review level—emphasizing visibility or ranking bias—our study complements this research by analyzing the effect at the vote level and embedding it within a broader set of voter-related influences (historical behavior, contextual cues, and social ties). Our results suggest that voters align with existing vote signals as part of a broader consistency-seeking process, rather than simply due to herding. This more comprehensive perspective invites future research to disentangle, and ultimately integrate, review-level display effects with voter-level alignment mechanisms when modeling how prior evaluations cascade through platforms.

Third, our study contributes to the understanding of "the wisdom of crowds." The popularity of the helpfulness voting system on many review platforms speaks to the trust that platforms and platform users have in the power of "crowds." Both practitioners and scholars often assume that the collective input of many voters will neutralize individual biases, revealing the true quality of a review over time (Surowiecki, 2005). Additionally, in past research on helpfulness voting, the invisibility and untraceability of individual voters at the review level have led to the exclusion of voters' perspectives. As a result, voters in prior studies are often (implicitly or explicitly) treated as rational decision makers, and voting decisions are presumed to arise from review-related and reviewer-related factors (Hong et al., 2017; Liu et al., 2020). To our knowledge, no research has examined the validity of this implicit assumption.

Leveraging our unique vote-level dataset, we assessed the validity of this assumption at both the vote and review levels. Our analysis highlights that voter-related factors significantly enhance the explanatory power of our model and have a greater effect on both individual votes and overall review helpfulness than review and reviewer factors. These findings imply that prior research may have underestimated voter influence and overestimated review and reviewer impacts. Furthermore, our results question the reliability of "the wisdom of crowds" in review helpfulness evaluations. The aggregated helpfulness ratio at the review level may be biased by factors such as voters' content-generating habits, perceptions of the review's contextual cues, and interpersonal factors that may not represent the true value of the review's content. We call for more research examining when this principle is effective and how to derive more accurate measurements of review helpfulness at the review level.

As such, our results challenge the implicit "rational voter" assumption embedded in much of the existing IS literature on crowd-based scoring systems. Rather than reflecting unbiased aggregations of independent judgments, aggregated review helpfulness can encode systematic cognitive-consistency biases, such as dissonance avoidance and social alignment. This suggests that the limitations are not merely due to measurement error but also a lack of theoretical grounding in how individuals psychologically process an individual review and contribute to the collective evaluation of the review.

Lastly, our study complements a recent stream of literature investigating the relationship between *helpful* and *unhelpful* votes on a review. For instance, past work has found that *helpful* and *unhelpful* votes are fundamentally different (Yu et al., 2024) and have different impacts on review writers receiving the votes (Zhu et al., 2024a). By also accounting for the perspective of review readers, our work provides insights into how review-, reviewer-, and voter-related characteristics drive the differences in review helpfulness that have been empirically identified in prior research; we also illuminate the underlying mechanisms of the identified dynamics.

5.2 Practical Implications

This study's findings have practical implications for both review platforms and writers. The commonly used helpfulness voting system on review platforms might not always accurately reflect the most useful reviews due to the impact of voters' own characteristics and the influence of review and reviewer factors. This suggests that rankings based solely on helpfulness votes may not consistently serve consumers' needs. To enhance the system's effectiveness, platforms could incorporate data on voter-related factors to, for example, adjust for voters' past activities or relationships with reviewers. For instance, platforms might consider discounting votes from followers of reviewers. Additionally, to mitigate herd mentality effects like the bandwagon effect, platforms could hide the total vote count or helpfulness ratio or even disable downvotes, a strategy some platforms like Amazon already use. However, such interventions could make it harder for consumers to assess review helpfulness or might introduce unintended consequences (Yu et al., 2026). Other interventions, such as displaying factors related to intrinsic review quality (Hou & Ma, 2022), may aid consumers in making better decisions, ultimately improving their satisfaction and the platform's performance.

Our research provides valuable insights for reviewers seeking to enhance their status on review platforms. While writing guidelines often emphasize review content, our findings highlight the importance of considering voter-related factors. Reviewers with a large follower network may receive more favorable evaluations, while reviews deviating from the consensus might be viewed less positively. This does not suggest that reviewers should manipulate content unauthentically; rather, reviewers should simply be aware that review value is not solely determined by content. Reviewers need to recognize the influence of their social network and the general sentiment of prior reviews on the reception of their reviews.

Although our data predates the widespread use of generative AI, our findings remain relevant in the current review environment. Since most major platforms explicitly prohibit AI-generated reviews and remove detected content, helpfulness voting systems are still designed to evaluate ostensibly human-authored reviews. Moreover, prior research has shown that consumers cannot reliably distinguish AI-generated reviews from those written by humans (Kovács, 2024). Because the voting behaviors we document are driven by psychological processes (e.g., consistency seeking, confirmation bias, and social alignment) that depend on voters' perceptions rather than the objective origin of review content, these mechanisms should continue to shape helpfulness evaluations as long as reviews are perceived to be human generated. This suggests that even in the era of generative AI, improving review evaluation systems requires attention not only to content detection but also to voter-level biases embedded in crowd voting.

5.3 Future Research

This paper offers several directions for future research. While our study focused on the main effects of voter-related factors on helpfulness votes, we did not explore interactions among these factors due to the complexity of the theoretical framework. Future studies could systematically examine these interactions, as well as interactions between voter-related factors and review- or reviewer-related characteristics, to provide a more comprehensive understanding of voting behavior. Second, the focal platform requires users to log in before voting. Future research could explore whether voter behavior differs when anonymous votes are allowed. Lastly, recognizing the importance of "the wisdom of crowds" in evaluating online content, our study underscores the influence of voter-related factors on review helpfulness. However, the applicability of these factors may vary across different types of user-generated content, such as Q&A communities. Future research should assess the applicability of these factors in various contexts of user-generated content.

5.4 Conclusion

It is common practice for online review platforms to identify helpful information by using helpfulness voting systems, and the prior literature has examined review-related and reviewer-related factors as determinants of review helpfulness. Our study contributes to this literature by investigating the impact of voter-related factors that have been largely omitted in prior studies. Our empirical results suggest that voter-related factors play a nontrivial role in review readers' casting of *helpful* (vs. *unhelpful*) votes. Moreover, the "wisdom of crowds" may not necessarily arise from the mere implementation of a helpfulness voting system, and more research is needed to better understand individual consumers' consumption and evaluation of reviews.

Acknowledgments

We thank the senior editor and the anonymous reviewers for their constructive feedback and guidance, which greatly improved this manuscript. Generative artificial intelligence tools (Grammarly, ChatGPT, and Claude) 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 the clarity of expression. All research concepts, study design, data analyses, interpretations, and conclusions are entirely the original contributions of the authors.

References

Abelson, R. P., Aronson, E. E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (1968). *Theories of cognitive consistency: A sourcebook*. Academic Press.

Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology*, *11*(1), 1-33.

Alibakhshi, R., & Srivastava, S. C. (2022). Post-story: Influence of introducing story feature on social media posts. *Journal of Management Information Systems*, *39*(2), 573-601.

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, *17*(2), 99-126.

Barker, J. R., & Tompkins, P. K. (1994). Identification in the self-managing organization characteristics of target and tenure. *Human Communication Research*, *21*(2), 223-240.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993-1022.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31-72.

Castano, E., Yzerbyt, V., Bourguignon, D., & Seron, E. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology*, *38*(3), 315-322.

Chaiken, S., Giner-Sorolla, R., & Chen, S. (1996). Beyond accuracy: Defense and impression motives in heuristic and systematic information processing. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 553-578). Guilford.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. *Unintended thought*, 212-252.

Cheng, Y. H., & Ho, H. Y. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research, 68*(4), 883-887.

Choi, H. S., Oh, W., Kwak, C., Lee, J., & Lee, H. (2022). Effects of online crowds on self-disclosure behaviors in online reviews: A multidimensional examination. *Journal of Management Information Systems*, *39*(1), 218-246.

Connors, L., Mudambi, S., & Schuff, D. (2011). Is it the review or the reviewer? A multi-method approach to determine the antecedents of online review helpfulness. *Proceedings of the 44th Hawaii International Conference on System Sciences.*

De Langhe, B., Fernbach, P., & Lichtenstein, D. (2016). Navigating by the stars: Investigating the actual and perceived validity of online user ratings. *Journal of Consumer Research*, *42*(6), 817-833.

De Matos, C. A., & Rossi, C. A. V. (2008). Word-of-mouth communications in marketing: A meta-analytic review of the antecedents and moderators. *Journal of the Academy of Marketing Science, 36*(4), 578-596.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, *84*(1), 151-161.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, *73*(1), 31.

Festinger, L. (1962). Cognitive dissonance. *Scientific American*, *207*(4), 93-106.

Filieri, R., Raguseo, E., & Vitari, C. (2018). When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type. *Computers in Human Behavior*, *88*, 134-142.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291-313.

Ghose, A., & Ipeirotis, P. G. (2010). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering, 23*(10), 1498-1512.

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.

Gouldner, A. W. (1960). The norm of reciprocity: a preliminary statement. *American Sociological Review*, *25*, 161-178.

Hall, D. T., Schneider, B., & Nygren, H. T. (1970). Personal factors in organizational identification. *Administrative Science Quarterly*, *15*(2), 176-190.

Harvey, C., Stewart, D., & Ewing, M. (2011). Forward or delete: What drives peer-to-peer message propagation across social networks? *Journal of Consumer Behaviour*, *10*(6), 365-372.

Hong, H., Xu, D., Wang, G., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems*, *102*, 1-11.

Hou, J., & Ma, X. (2022). Space norms for constructing quality reviews on online consumer review sites. *Information Systems Research*, *33*(3), 1093-1112.

Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior, 48,* 17-27.

Kim, A., Moravec, P. L., & Dennis, A. R. (2019). Combating fake news on social media with source ratings: The effects of user and expert reputation ratings. *Journal of Management Information Systems*, *36*(3), 931-968.

Kim, S.-M., Pantel, P., Chklovski, T., & Pennacchiotti, M. (2006). Automatically assessing review helpfulness. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Kovács, B. (2024). The Turing test of online reviews: Can we tell the difference between human-written and GPT-4-written online reviews? *Marketing Letters*, *35*(4), 651-666.

Kuan, K., Hui, K., Prasarnphanich, P., & Lai, H. (2015). What makes a review voted? An empirical investigation of review voting in online review systems. *Journal of the Association for Information Systems*, *16*(1), 48-71.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480-498.

Lane, J. N., Teplitskiy, M., Gray, G., Ranu, H., Menietti, M., Guinan, E. C., & Lakhani, K. R. (2020). *When do experts listen to other experts? The role of negative information in expert evaluations for novel projects.* Harvard Business School.

Lee, J., Agrawal, M., & Rao, H. (2015). Message diffusion through social network service: The case of rumor and non-rumor related tweets during Boston bombing 2013. *Information Systems Frontiers*, *17*(5), 997-1005.

Legal, J., Chappe, J., Coiffard, V., & Villard-Forest, A. (2012). Don't you know that you want to trust me? Subliminal goal priming and persuasion. *Journal of Experimental Social Psychology*, *48*(1), 358-360.

Lei, Z., Yin, D., Mitra, S., & Zhang, H. (2022). Swayed by the reviews: Disentangling the effects of average ratings and individual reviews in online word-of-mouth. *Production and Operations Management*, *31*(6), 2393-2411.

Lei, Z., Yin, D., & Zhang, H. (2021). Focus within or on others: The impact of reviewers' attentional focus on review helpfulness. *Information Systems Research, 32*(3), 801-819.

Lei, Z., Yin, D., & Zhang, H. (2025). Deliberative or automatic: disentangling the dual processes behind the persuasive power of online word-of-mouth. *MIS Quarterly*, *49*(1), 331-346.

Liu, A. X., Li, Y., & Xu, S. X. (2021). Assessing the unacquainted: Inferred reviewer personality and review helpfulness. *MIS Quarterly, 45*(3), 1113-1148.

Liu, X., Lee, D., & Srinivasan, K. (2019). Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning. *Journal of Marketing Research*, *56*(6), 918-943.

Liu, X., Wang, G. A., Fan, W., & Zhang, Z. (2020). Finding useful solutions in online knowledge communities: A theory-driven design and multilevel analysis. *Information Systems Research*, *31*(3), 731-752.

Liu, Y., & Park, S. (2006). What makes a useful online review? Implication for travel product websites. *Tourism Management*, *27*(5), 687-698.

Ma, X., Khansa, L., Deng, Y., & Kim, S. (2013). Impact of prior reviews on the subsequent review process in reputation systems. *Journal of Management Information Systems*, *30*(3), 279-310.

Matook, S., Cummings, J., & Bala, H. (2015). Are you feeling lonely? The impact of relationship characteristics and online social network features on loneliness. *Journal of Management Information Systems*, *31*(4), 278-310.

Mattila, A. (2003). The impact of cognitive inertia on postconsumption evaluation processes. *Journal of the Academy of Marketing Science*, *31*(3), 287-299.

Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine learning with oversampling and undersampling techniques: Overview study and experimental results. *Proceedings of the 11th International Conference on Information and Communication Systems.*

Mudambi, S., & 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.

Murnane, K., Phelps, M., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General*, *128*(4), 403-415.

Nian, T., Hu, Y., & Chen, C. (2021). Examining the impact of television-program-induced emotions on online word-of-mouth toward television advertising. *Information Systems Research*, *32*(2), 605-632.

Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175-220.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*(1), 42-55.

Pan, Y., & Wu, D. (2020). A novel recommendation model for online-to-offline service based on the customer network and service location. *Journal of Management Information Systems*, *37*(2), 563-593.

Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing, 87*(4), 598-612.

Papke, L., & Wooldridge, J. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619-632.

Papke, L., & Wooldridge, J. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, *145*(1-2), 121-133.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Peng, C., Yin, D., & Zhang, H. (2020). More than words in medical question-and-answer sites: A content-context congruence perspective. *Information Systems Research*, *31*(3), 913-928.

Petrocelli, J. V., Tormala, Z. L., & Rucker, D. D. (2007). Unpacking attitude certainty: Attitude clarity and attitude correctness. *Journal of Personality and Social Psychology*, *92*(1), 30.

Pu, Z., Li, S., Wu, J., & Liu, Y. (2024). Social network integration and online review helpfulness: An empirical investigation. *Journal of the Association for Information Systems, 25*(6), 1563-1584.

Qahri-Saremi, H., & Montazemi, A. R. (2019). Factors affecting the adoption of an electronic word of mouth message: a meta-analysis. *Journal of Management Information Systems*, *36*(3), 969-1001.

Qahri-Saremi, H., & Turel, O. (2023). Situational contingencies in susceptibility of social media to phishing: A temptation and restraint model. *Journal of Management Information Systems*, *40*(2), 503-540.

Qin, J., Zheng, P., & Wang, X. (2022). Comprehensive helpfulness of online reviews: A dynamic strategy for ranking reviews by intrinsic and extrinsic helpfulness. *Decision Support Systems, 163,* Article 113859.

Racherla, P., & Friske, W. (2012). Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories. *Electronic Commerce Research and Applications, 11*(6), 548-559.

Ramalho, E., Ramalho, J., & Murteira, J. (2011). Alternative estimating and testing empirical strategies for fractional regression models. *Journal of Economic Surveys*, *25*(1), 19-68.

Ramalho, E., Ramalho, J., & Murteira, J. (2014). A generalized goodness-of-functional form test for binary and fractional regression models. *Manchester School*, *82*(4), 488-507.

Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems, 81*, 30-40.

Schmitt-Beck, R. (2015). Bandwagon effect. In G. Mazzoleni (Ed.), *The international encyclopedia of political communication* (n.p.). John Wiley & Sons.

Seutter, J., Kutzner, K., Stadtländer, M., Kundisch, D., & Knackstedt, R. (2023). "Sorry, too much information": Designing online review systems that support information search and processing. *Electronic Markets*, *33*, Article 47.

Sun, H., Fan, M., & Tan, Y. (2020). An empirical analysis of seller advertising strategies in an online marketplace. *Information Systems Research*, *31*(1), 37-56.

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.

Tseng, S.-L., Lu, S., Weathers, D., & Grover, V. (2023). How product review voting is influenced by existing votes, consumer involvement, review valence, and review diagnosticity. *Decision Support Systems*, *172*, Article 113981.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell.

van Herpen, E., Pieters, R., & Zeelenberg, M. (2009). When demand accelerates demand: Trailing the bandwagon. *Journal of Consumer Psychology*, *19*(3), 302-312.

Warriner, A., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.

Wei, Z., Xiao, M., & Rong, R. (2021). Network size and content generation on social media platforms. *Production and Operations Management*, *30*(5), 1406-1426.

Wooldridge, J. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

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.

Yin, D., Bond, S., & Zhang, H. (2017). Keep your cool or let it out: Nonlinear effects of expressed arousal on perceptions of consumer reviews. *Journal of Marketing Research*, *54*(3), 447-463.

Yin, D., de Vreede, T., Steele, L., & de Vreede, G. (2023). Decide now or later: Making sense of incoherence across online reviews. *Information Systems Research*, *34*(3), 1211-1227.

Yin, D., Mitra, S., & Zhang, H. (2016). When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. *Information Systems Research*, *27*(1), 131-144.

Yin, D., Bond, S. D., & Zhang, H. (2021). Anger in consumer reviews: Unhelpful but persuasive? *MIS Quarterly*, *45*(3), 1059-1086.

Yu, Y., Khern-am-nuai, W., & Pinsonneault, A. (2022). When paying for reviews pays off: The case of performance-contingent monetary rewards. *MIS Quarterly*, *46*(1), 609-626.

Yu, Y., Khern-am-nuai, W., Pinsonneault, A., & Wei, Z. (2023). The impacts of social interactions and peer evaluations on online review platforms. *Journal of Management Information Systems*, *40*(4), 1271-1300.

Yu, Y., Gutt, D., & Khern-am-nuai, W. (2024). Review helpfulness scores vs. review unhelpfulness scores: Two sides of the same coin or different coins? *IEEE Transactions on Engineering Management*, *71*, 8031-8044.

Yu, Y., Yin, D., & Khern-am-nuai, W. (2026). Fewer reviews but better content: When an online review platform disables downvotes. *Journal of Management Information Systems*, *43*(2), 654-679.

Zhou, W., & Duan, W. (2016). Do professional reviews affect online user choices through user reviews? An empirical study. *Journal of Management Information Systems*, *33*(1), 202-228.

Zhu, K., Khern-am-nuai, W., & Yu, Y. N. (2024a). Negative peer feedback and user content generation: Evidence from a restaurant review platform. *Production and Operations Management*, *34*(12), 3814-3829.

Zhu, Q. H., Sun, R., Lv, D., Cheng, R. X., Zuo, J. J., & Qin, S. K. (2024b). Research on cognitive neural mechanism of consumers convinced by intelligent recommendation platform eavesdropping. *Scientific Reports*, *14,* Article 28108.

Appendix A: Summary of Existing Studies on Review Helpfulness

Table A1. Summary of Major Determinants of Review Helpfulness Identified in Prior Research

Determinant	Representative studies
Review-related	Cheng & Ho, 2015; De Matos & Rossi, 2008; Filieri et al., 2018; Hong et al. , 2017; Huang et al., 2015; Kuan et al., 2015; Lei et al., 2021; Liu & Park, 2006; Mudambi & Schuff, 2010; Pan & Zhang, 2011; Qahri-Saremi & Montazemi, 2019; Qin et al., 2022; Salehan & Kim, 2016; Yin et al., 2014; Yin et al., 2016
Reviewer-related	Cheng & Ho, 2015; Ghose & Ipeirotis, 2010; Hong et al., 2017; Huang et al., 2015; Liu & Park, 2006; Liu et al., 2021; Pu et al., 2024; Qahri-Saremi & Montazemi, 2019; Racherla & Friske, 2012; Yu et al., 2023
Voter-related	Yin et al., 2016 (implied); this study

Note: This table presents a representative, not exhaustive, list of studies that examine the determinants of review helpfulness.

Appendix B: Topic Modeling

Our topic modeling was based on the latent Dirichlet allocation (LDA) model (Blei et al., 2003), which allowed us to discover semantic structures hidden in the textual content of reviews in our sample. LDA is an unsupervised clustering model for discovering abstract topics of a collection of documents and generating a predefined number of topics. In LDA, each review is modeled as a mixture of various topics, which, in turn, are modeled as term distributions. We used the *scikit-learn* package for Python (Pedregosa et al., 2011) to analyze the review text. We ran the model with three, four, five, and six topics, respectively, and inspected the distribution of terms. We found that the choice of four topics delivered the lowest perplexity, which commonly measures the evaluation of topic models (Blei et al., 2003), and delivered the most meaningful topic distribution. We manually inspected the topics generated by LDA and labeled them as food and meal, drink and dessert, service, and infrastructure and hardware, respectively. Table B1 lists the top ten terms of the four topics we labeled.

Table B1. Top Terms of Topics in Reviews

Topic	Top terms
1 Food and meal	Pork, fried, rice, delicious, eat, chicken, taste, noodle, sauce, soup
2 Drink and dessert	Like, ice, eat, cream, sweet, good, tea, coffee, delicious, milk
3 Service	Eat, food, good, order, like, come, time, restaurant, price, menu
4 Infrastructure and hardware	Shop, park, restaurant, eat, come, road, open, locate, lot, store

Appendix C: Variable Descriptions and Summary Statistics

Table C1. Variable Descriptions and Summary Statistics

Variable	Description	Mean	*SD*	Min	Max
Vote-related main variables:					
$VoterUpvoteRatio_{kt}$	The ratio of the number of upvotes to total votes cast by the focal voter	0.96	0.03	0	1
$VoterReviewCnt_{kt}$	The number of reviews generated by the focal voter	479.22	445.15	0	4,178
$ReviewRatingDiff_{jnt}$	The absolute difference between the rating of the focal review and the average rating obtained by the focal restaurant	0.43	0.45	0	3.64
$ReviewUpvoteRatio_{jt}$	The ratio of the number of upvotes to total votes received by the focal review	0.97	0.03	0	1
$VoterIsFollower_{kmt}$	A dummy variable indicating if the voter is following the reviewer of the focal review	0.45	0.50	0	1
$VoterIsFollowee_{kmt}$	A dummy variable indicating if the voter is followed by the reviewer of the focal review	0.43	0.50	0	1
Voter-related control variables					
$VoterTenure_{kt}$	The number of days elapsed since the voter joined the platform	1,016.23	469.27	0	3,167
$VoterGender_{k}$	A dummy variable indicating if the voter is a male (1) or female (0)	0.62	0.48	0	1
Review characteristics:					
$ReviewLength_{j}$	The number of words in the review	532.01	612.52	4	23,257
$ReviewReadability_{j}$	The educational grade level required to understand the review	8.39	3.20	0.40	109.80
$ReviewPosValence_{j}$	The ratio of positive words to total words	0.17	0.10	0	1
$ReviewNegValence_{j}$	The ratio of negative words to total words	0.03	0.03	0	0.74
$ReviewAge_{jt}$	The number of days elapsed since the focal review was published	36.24	130.06	0	1,257
$Topic1_{j}$	Percentage of Topic 1 (food and meal) in the review	0.22	0.27	0	1
$Topic2_{j}$	Percentage of Topic 2 (drink and dessert) in the review	0.31	0.36	0	1
$Topic3_{j}$	Percentage of Topic 3 (service) in the review	0.28	0.30	0	1
$Topic4_{j}$	Percentage of Topic 4 (infrastructure and hardware) in the review	0.19	0.24	0	1
Reviewer characteristics:					
$ReviewerReviewCnt_{mt}$	The number of reviews generated by the focal reviewer	494.63	408.95	1	6,230
$ReviewerAvgRating_{mt}$	The average rating of reviews generated by the focal reviewer	3.81	0.50	1	5
$ReviewerTenure_{mt}$	The number of days elapsed since the reviewer joined the platform	1,017.31	461.01	0	2,205
$ReviewerGender_{m}$	A dummy variable indicating if the reviewer is a male (1) or female (0)	0.63	0.48	0	1
Restaurant characteristics:					
$RestReviewCnt_{nt}$	The number of reviews received by the focal restaurant	33.23	58.30	1	2,529
$RestAvgRating_{nt}$	The average rating of reviews received by the focal restaurant	3.81	0.50	1	5
Note: The number of vote-level and review-level observations are 331,286 and 46,450, respectively. Variables with subscript *t* were calculated based on data before the focal vote. We excluded users without logins when reading the reviews since we are unable to capture their information. Among 21,784 restaurants, 5,492 were dine-in restaurants, 10,675 were express restaurants, 1,948 were drink and dessert shops, and 3,669 belonged to other categories.					

Appendix D: Balance Check after PSM

Table D1. Balance Check After PSM for *VoterUpvoteRatio*

Variable	Treatment mean	Control mean	*t*-value (*p*-value)
ReviewLength	894.40	887.20	1.006 (0.315)
ReviewReadability	8.51	8.49	0.754 (0.451)
ReviewPosValence	0.15	0.15	0.913 (0.361)
ReviewNegValence	0.03	0.03	0.820 (0.413)
ReviewAge	1.97	2.01	0.797 (0.426)
Topic1	0.24	0.24	0.238 (0.812)
Topic2	0.31	0.31	1.289 (0.197)
Topic3	0.23	0.23	-1.738 (0.082)

Table D2 Balance. Check After PSM for *VoterReviewCnt*

Variable	Treatment mean	Control mean	*t*-value (*p*-value)
ReviewLength	1,075.11	1,114.39	-1.560 (0.119)
ReviewReadability	8.26	8.37	-1.936 (0.053)
ReviewPosValence	0.14	0.14	1.579 (0.115)
ReviewNegValence	0.03	0.03	0.455 (0.649)
ReviewAge	1.04	1.13	-1.699 (0.089)
Topic1	0.24	0.23	0.723 (0.470)
Topic2	0.34	0.29	1.872 (0.062)
Topic3	0.20	0.21	-1.344 (0.179)

Table D3. Balance Check After PSM for *ReviewRatingDiff*

Variable	Treatment mean	Control mean	*t*-value (*p*-value)
VoterTenure	990.30	979.62	0.875 (0.382)
VoterGender	0.72	0.72	0.084 (0.933)
VoterReviewCnt	370.22	378.30	-0.809 (0.419)
VoterRating	3.82	3.81	0.877 (0.381)
VoterFollower	4,255.30	4,557.76	-1.161 (0.246)
VoterFollowee	79.81	81.88	-0.369 (0.712)
VoterVote	21,671.10	22,814.06	1.664 (0.096)

Table D4. Balance Check after PSM for *ReviewUpvoteRatio*

Variable	Treatment mean	Control mean	*t*-value (*p*-value)
VoterTenure	1,116.88	1,131.28	-1.368 (0.171)
VoterGender	0.70	0.68	1.705 (0.088)
VoterReviewCnt	422.37	416.76	0.643 (0.520)
VoterRating	3.66	3.67	-1.853 (0.064)
VoterFollower	5,590.52	5,350.85	0.855 (0.393)
VoterFollowee	83.25	84.83	-0.640 (0.522)
VoterVote	14,670.83	14,912.13	-0.630 (0.529)

Appendix E Robustness Checks and Additional Analysis

To validate the robustness of our findings, we conducted additional checks. These checks consistently confirmed our results.

Addressing Data Imbalance: Undersampling and Oversampling Strategies

In our dataset, the majority of votes were *helpful* votes. To address the challenge of an imbalanced dataset, we followed established methods and employed both undersampling and oversampling strategies (Mohammed et al., 2020). These techniques involved either reducing observations in the majority class or increasing observations in the minority class to achieve a more balanced dataset.

First, we applied undersampling, which involves randomly reducing the number of observations from the majority group to match the size of the smaller minority group. Specifically, our vote-level dataset contained 16,539 downvotes. To balance the dataset, we randomly selected an equal number of upvotes and re-ran our analysis. The results, reported in Table E1 of the appendix, Column 1, remained consistent with those from our main analysis (Table 1), demonstrating that voter-related factors significantly influenced voting behavior.

Next, we implemented oversampling, which increases the number of observations in the minority group. Specifically, we duplicated the subset of downvote observations until their count matches that of the upvote group. We then reran our analysis on this expanded dataset. The results, presented in Table E1, Column 2, are statistically consistent with our main findings, further reinforcing the robustness of the initial results. These additional analyses confirm that the impact of voter-related factors on voting behavior held under different data-balancing strategies.

Table E1. Impacts of Voter-Related Factors on Helpfulness Vote Using Undersampling and Oversampling Strategies

	(1)	(2)
	$\boldsymbol{VoteHelpful_i}$	
Variables	**Undersampling**	**Oversampling**
$VoterUpvoteRatio_{kt}$	63.97*** (4.185)	23.72*** (0.270)
$ln(VoterReviewCnt_{kt})$	0.308*** (0.035)	0.162*** (0.003)
$ReviewRatingDiff_{jnt}$	-0.228** (0.105)	-0.252*** (0.011)
$ReviewUpvoteRatio_{jt}$	0.670 (1.195)	0.575*** (0.100)
$VoterIsFollower_{kmt}$	1.187*** (0.144)	1.035*** (0.016)
$VoterIsFollowee_{kmt}$	0.771*** (0.136)	0.701*** (0.015)
Control variables	Yes	Yes
Constant	-66.15*** (4.381)	-24.76*** (0.296)
Observations	33,078	416,114

Note: We included a random intercept at the voter level. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

An Alternative Approach to Modeling Voter Decisions

Since the majority of votes in our sample were *helpful* votes, an alternative approach to model the voters' decision would be to classify the vote behavior into three categories: (0) the reader does not vote after reading the review (used as the base category), (1) the reader votes the review *unhelpful*, and (2) the reader votes the review *helpful*. Accordingly, we conducted a multinomial logistic regression analysis. We incorporated the same set of independent variables used in the main analysis. As shown in Table E2, the results are qualitatively consistent with those from our primary analysis, supporting the robustness of our findings.

Table E2. Result of Multinomial Logistic Regression

	(1)
Variables	***VoteType3Cat***
$VoterUpvoteRatio_{kt}$	13.45*** (0.164)
$ln(VoterReviewCnt_{kt})$	0.564*** (0.003)
$ReviewRatingDiff_{jnt}$	-0.130*** (0.010)
$ReviewUpvoteRatio_{jt}$	0.267** (0.114)
$VoterIsFollower_{kmt}$	2.372*** (0.016)
$VoterIsFollowee_{kmt}$	2.491*** (0.020)
Control variables	Yes
Constant	-13.14*** (0.211)
Observations	470,087

Note: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix F: Descriptions for Variables in Review-Level Analysis

Table F1. Descriptions for Variables in Review-Level Analysis

Variable	Description
Panel A: Cross-sectional analysis (review level)	
$Helpful\%_{j(30d)}$	The ratio of the number of *helpful* votes to total votes obtained by the focal review within 30 days after it is published
$VoterUpvoteRatio_{k(30d)}$	The average ratio of the number of *helpful* votes to total votes cast by voters prior to their vote for the focal review
$VoterReviewCnt_{k(30d)}$	The average number of reviews generated by voters prior to their vote for the focal review
$ReviewRatingDiff_{jn(30d)}$	The average absolute difference between the rating of the focal review and the average ratings of the focal restaurant prior to each vote
$Follower_{m(30d)}$	The average number of users who are following the focal reviewer prior to each vote
$Followee_{m(30d)}$	The average number of users who are followed by the focal reviewer prior to each vote
$VoterTenure_{k(30d)}$	The average number of days elapsed since the voters joined the platform prior to their vote for the focal review
$MaleVoterRatio_{k(30d)}$	The average ratio of male voters prior to each vote
$ReviewAge_{j(30d)}$	The average number of days elapsed since the focal review was published prior to each vote
$ReviewerReviewCnt_{m(30d)}$	The average number of reviews generated by the focal reviewer prior to each vote
$ReviewerAvgRating_{m(30d)}$	The average of the mean rating of reviews generated by the focal reviewer prior to each vote
$ReviewerTenure_{m(30d)}$	The average number of days elapsed since the focal reviewer joined the platform prior to each vote
$ReviewerGender_{m}$	A dummy variable indicating if the reviewer is a male (1) or female (0)
$RestReviewCnt_{n(30d)}$	The average number of reviews received by the focal restaurant prior to each vote
$RestAvgRating_{n(30d)}$	The average of the mean rating of reviews received by the focal restaurant prior to each vote
Panel B: Panel analysis (review-by-week level)	
$Helpful\%_{jw}$	The ratio of the number of *helpful* votes to total votes obtained by the focal review up to week *w*
$VoterUpvoteRatio_{k(w-1)}$	Average ratio of the number of upvotes to total votes cast by the focal voter (up to week *w*-1) prior to each vote
$VoterReviewCnt_{k(w-1)}$	Average number of reviews generated by the focal voter (up to week *w*-1) prior to each vote
$ReviewRatingDiff_{jn(w-1)}$	Average absolute difference between the rating of the focal review and the average rating of the focal restaurant (up to week *w*-1) prior to each vote
$ReviewUpvoteRatio_{j(w-1)}$	Average ratio of the number of upvotes to total votes received by the review (up to week *w*-1) prior to each vote
$Follower_{m(w-1)}$	Average number of users who are following the focal reviewer (up to week *w*-1) prior to voting for the focal review
$Followee_{m(w-1)}$	Average number of users who are followed by the focal reviewer (up to week *w*-1) prior to voting for the focal review
$VoterTenure_{k(w-1)}$	Average number of days elapsed since the voters joined the platform (up to week *w*-1) prior to their vote for the focal review
$MaleVoterRatio_{k(w-1)}$	Average ratio of male voters (up to week *w*-1) prior to each vote
$ReviewAge_{j(w-1)}$	The number of days elapsed since the focal review was published up to week *w*-1
$ReviewerReviewCnt_{m(w-1)}$	The number of reviews generated by the focal reviewer up to week *w*-1
$ReviewerAvgRating_{m(w-1)}$	The average rating of reviews generated by the focal reviewer up to week *w*-1
$ReviewerTenure_{m(w-1)}$	Average number of days elapsed since the focal reviewer joined the platform (up to week *w*-1) prior to each vote
$RestReviewCnt_{n(w-1)}$	The number of reviews received by the focal restaurant up to week *w*-1
$RestAvgRating_{n(w-1)}$	The average rating of reviews received by the focal restaurant up to week *w*-1

About the Authors

Yinan Yu is an 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. Her work has been published in *MIS Quarterly*, *Production and Operations Management*, *Journal of Management Information Systems*, and other academic journals.

Dezhi Yin is an associate professor of information systems and Muma Fellow at the Muma College of Business, University of South Florida. He received his PhD in information technology management from the 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). He has published in *MIS Quarterly*, *Information Systems Research*, *Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *Journal of Marketing Research*, *Academy of Management Journal*, *Production and Operations Management*, and other academic journals. He is a recipient of the 2024 AIS Mid-Career Award and the 2021 INFORMS ISS Sandra A. Slaughter Early Career Award. He serves as associate editor for *MIS Quarterly*, and he is a recipient of the 2023 Outstanding Associate Editor Award from the same journal.

Warut Khern-am-nuai 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. His research interests include the economics and ethics of artificial intelligence, platforms for online marketplaces, and the management of information security. He earned his PhD in management information systems from the Krannert School of Management, Purdue University, in 2016.

Alain Pinsonneault, Fellow of the Royal Society of Canada, of the National Order of Québec, and of the Association for Information Systems, is a Distinguished James McGill Professor and the Imasco Chair of Information Systems in the Desautels Faculty of Management at McGill University. His current research interests include the organizational and individual impacts of information technology, user adaptation, business model in the digital economy, the impact of AI on work, and the business value of IT. His research has appeared in numerous journals, including *Management Science, MIS Quarterly, Information Systems Research, Journal of Management Information Systems, Decision Support Systems*, and *Organization Science*. He has served on the editorial boards of several other journals, including *MIS Quarterly, Information Systems Research, Organization Science,* and *Journal of Management Information Systems.*

Zaiyan Wei is the David and Margaret Crow Rising Star Associate Professor in the Daniels School of Business at Purdue University and a faculty affiliate with Krenicki Center and the White Real Estate Finance Program. He serves on the editorial boards of premier journals, including *Information Systems Research*, *Decision Sciences Journal*, and *Information & Management*. His research examines the economics of information systems, with a focus on artificial intelligence, fintech, the sharing economy, and social media and networks. His work has appeared in leading journals, including *Information Systems Research*, *Management Science*, *MIS Quarterly*, and *Production and Operations Management*, and has received Best Paper nominations and awards at major academic conferences. In 2022, he received the Sandra A. Slaughter Early Career Award from INFORMS and the Krannert Faculty Fellow Award. He has secured grants from Daniels School of Business and the Marketing Science Institute and has served as an ad hoc judge for the Alden G. Clayton Doctoral Dissertation Competition. He earned his PhD in economics in 2015 and previously held a research associate position at the University of Arizona.