Evaluating the Impact of Incentive/Non-incentive Reviews on Customer Decision-making

1st Kate Kargozari
Department of Information Science
University of North Texas
Denton, United States
katekargozari@my.unt.edu

2nd Junhua Ding
Department of Information Science
University of North Texas
Denton, United States
junhua.ding@unt.edu

3rd Haihua Chen
Department of Information Science
University of North Texas
Denton, United States
haihua.chen@unt.edu

***Abstract*—In recent years, growing social media and e-commerce has underscored the necessity of evaluating seller-customer relations. The online review system is one way to assess this relationship. However, the trustworthiness and quality of these reviews, especially incentivized ones, remains a concern. To address online review quality, this study evaluates the credibility and consistency based on the volume, length, and content of online reviews. The study tends to distinguish between incentive and organic reviews, discover the impact of incentives on customer review behavior, and consequently on improving review quality as a pivotal determinant in purchase decision-making. To confront these problems, experiments were conducted using software product reviews collected from software review websites, including Capterra, Software Advice, and GetApp. EDA results highlight the significance of review features such as cost, customer support, ease of use, and product features. The indirect impact of companies' size based on the number of employees, the direct impact of users' experience, and the different impacts of changing situations during years on the volume of incentive reviews are major findings of sentiment analysis. The A/B testing results indicate the range of having no impact to the less impact of incentive reviews on purchase decision-making regarding different review scores. We also find lower credibility, less consistency regarding volume and length, and more consistency regarding content within incentive reviews. This study suggests the necessity of potential contributions among companies to improve the quality of online review systems.**

***Index Terms*—incentive, organic, online reviews, review quality, decision-making, A/B testing**

I. Introduction

Recently social media revolutionized e-commerce by transforming how to assess seller-customer relationships. Among the critical factors in online purchasing decisions, including electronic word-of-mouth (eWOM), the price, and the website/business's reputation, online reviews, a form of eWOM, are specially crucial for shoppers decision-making [1]. Reviews categorize based on criteria. One categorization refers to ranking products or services by customer sentiment including positive, negative, and neutral. Another categorization is based on reviews' writing criteria such as experience and monetary rewards denotes reviews as organic (no incentive or non-incentivized), incentive, or fake. Despite organic/non-incentivized reviews [2]–[4] that are based on real experiences and free from external motivation or incentives, some individuals may be tempted with rewards to write either incentivized reviews reflecting their actual product purchases [2], [3], [5]–[9] or fake reviews [7], [10], [11] lacking experiential foundation.

Delivering accurate information through the online review system is vital for informed purchases and reducing bias in existing seller-provided descriptions [5]. Online review systems face the major challenge of obtaining truthful and high-quality responses from agents [12]. Factors such as social presence can mediate the relationship between online review language style and consumers' purchase intention [13]. Although businesses can save money while receiving organic reviews [1], many customers ignore posting the reviews. A direct relationship between the number of reviews and sale [6] encourages sellers to offer monetary rewards for honest reviews to boost both review count and product rating [14], [15] while reducing bias. Incentivized reviews affect customer satisfaction [2]. Incentives can impact consumers' expressions and increase positive emotions in reviews [6], which influence purchase intention, trust and satisfaction [1].

However, the sensitivity of offering incentives may have a positive or negative effect [3] with possible results of negative reviews being seen as more credible reviews [16]. High volume of online reviews for a product can cause confusion, misinformation, and misleading [7] purchase decision, which harming trust and truthfulness of the reviews. Sellers' guidance for high-quality reviews has gaps. Not all positive/negative reviews are accurate, and customer satisfaction does not always align with review sentiment. Thus, customer behavior [17]–[19] toward posting product purchase reviews influences review quality. Improving review quality [3] enhances trust, assisting purchase decision-making and facilitating valuable contributions to the new review process.

This study aims to assess the impact of incentives on customers' posting review behavior and review quality by examining the difference between incentive and organic reviews. Furthermore, the study will utilize both existing evidence on review quality along with information gained from this research to propose novel approaches for enhancing review quality reflecting on the credibility [20]–[23] and consistency [24]–[30] of the reviews while considering the impact of customers' purchase review behavior. Credible online reviews positively impact the hedonic brand image [21]. Despite the

979-8-3503-3629-0/23/$31.00 ©2023 IEEE
DOI 10.1109/AITest58265.2023.00033

alarming message by negative reviews, they are not inherently more credible than positive reviews [23]. Source credibility moderates the relationship between review comprehensiveness and review usefulness [20], [22]. Consistency can impact credibility of the review as consistent review can be either high or low quality review [28]. Consistency in content negatively impacts informational influence [24] and review helpfulness [25], [26], while positively affects online reviews credibility [29]. Depends on the study, review consistency may positively impact review usefulness [27] and brands attitudes [30]. In this study, we focus on assessing online reviews credibility and consistency based on their volume, length, and content. To achieve this goal, we design two research questions:

1) What are the significant differences between incentive and organic reviews?
2) How do incentives impact on customers' behavior on posting purchase's review, and as a result, on purchase's review quality, with impact on purchase decision-making?

We performed Exploratory Data Analysis (EDA) on various data set features pertaining to the "incentivized" status. Additionally, we conducted EDA analysis on sentiment analysis of review text to distinguish between incentivized and organic reviews. Furthermore, We applied A/B testing on review rating scores, and as the ultimate goal, examined the impact of incentives on customers' purchase decision-making.

To the best of our knowledge, previous studies have not discovered the effect of company size and years of user experience as contributing factors. Moreover, our analysis of software reviews yields valuable insights for enhancing purchase reviews and more specific software reviews that generates more targeted guidelines to enhance overall review quality.

The article's structure is as follows: Section II presents related work, follows by Methodology in Section III. We present our results in Section IV, and further discussion in Section V. The article concludes in Section VI.

II. Related Work

Previous investigations have shown that trust in online reviews is equivalent to trust in friends' recommendations [31]. Therefore, this study reviews the variance between organic and incentivized reviews, their impact on customer behavior, review quality, and customer decision-making.

A. Incentive vs. Organic

Online reviews impact purchase intention; therefore, the relationship between online review stimulus and purchase intention response is important to explore [1]. In the short term, reviewers' contribution and readability levels rise; however, over time, review quality improves, leading to stabilization of their numerical rating behaviors [8]. Study demonstrates that incentives can enhance review quality [9]. Furthermore, in accordance with the social exchange theory (SET), incentives may motivate social behavior by considering the satisfaction of individual needs, such as encouraging customers to write online reviews [9]. Disclosing incentives maintains trust, reduces bias, boosts helpfulness and increases sales [10]. The disclosure of intrinsic communication motives for writing product reviews is more authentic and less betraying [32]. However, the impact of disclosing statements on product quality judgment depends on either customers' disclosure is integral or incidental [33]. For companies, incentives increase attracting customers' attention [3] and products' rate, reduces products' return, and contributes to companies' success [9].

B. Incentive and Purchase Decision-making

Purchase decision-making considers complex situation using utility-driven systems to ease purchase decision-making by providing more details [34]. Incentivized reviews boost the effectiveness of efficient review signals to new customers [35]. In addition, incentives make users more active [36] and increase review writers by combining with social norm [4], which makes review writing more enjoyable [6], and increase review numbers [9]. Accordingly, incentives increase the volume and length of online reviews [4], [5]; and consequently increase the volume of provided information to new customers for better purchase decisions [4]. Moreover, according to loss aversion theory, review valence is more influential than review usefulness in the decision-making process [37]; therefore, incentive reviews impact purchase decision-making as they increase valence [9] by increasing emotional words in customers' WOMs [5], [6]. Disclosing incentives is crucial to prevent accuracy decrease and new consumer decision misguidance [10].

On the other hand, existing studies have investigated the importance of avoiding incentives. Offering and accepting incentives can decrease trust as follows market norms, not social norms that is a sign of human behavioral issues, raises moral concerns, increases review fraud that undermines review credibility, and establishes the interest between businesses and review posters [11]. In addition, incentives increase biased positive reviews [4]. Despite differing views [4], [5], incentives may reduce user effort to write lengthy informative reviews [38]. Moreover, customers who are uncomfortable receiving incentives for their opinion may deliver negative reviews [4], which are valuable [39].

C. Approaches of Identifying and Analyzing Incentive Reviews

Multiple studies explore the effect of monetary incentives on online reviews' quality and value using unique approaches.

Incentive reviews were identified using data mining techniques considering the overall rating, helpfulness rate, review length, and other factors. VADER algorithm was used to improve model by incorporating review sentiment scores [5]. A difference-in-differences analysis reevaluated reviewers' behavior [8]. Counterfactual thinking approach investigated incentives' effect on online review publication likelihood and valence in two experimental studies using a scenario approach, replicating and extending results by considering customers' satisfaction levels [9]. Machine-learning-based and dictionary-based approaches assessed the impact of sending efficient signals by review [35]. To analyze incentives' effect on product attention, data manipulation and statistical tests such as p-values and t-tests are used [3]. The stimulus-organism-response (S-O-R) framework developed, and a component-based structural equation modeling method (Smart PLS) used to assess online reviews and purchase intention relationships [1]. Extensive quantitative methods alongside the mixed-method experimental studies were used to evaluate review valence's impact on decision-making [37]. Regression models commonly used to assess reviews' impact on decision-making [34], that often incorporating methods like sentiment analysis [10]. A multi-methodological research design of two randomized experiments was utilized to test incentives' impact on reviews volume and length, and the potential bias in purchase decision-making [4].

III. Methodology

We introduce approaches and techniques for data collection and analysis in this section.

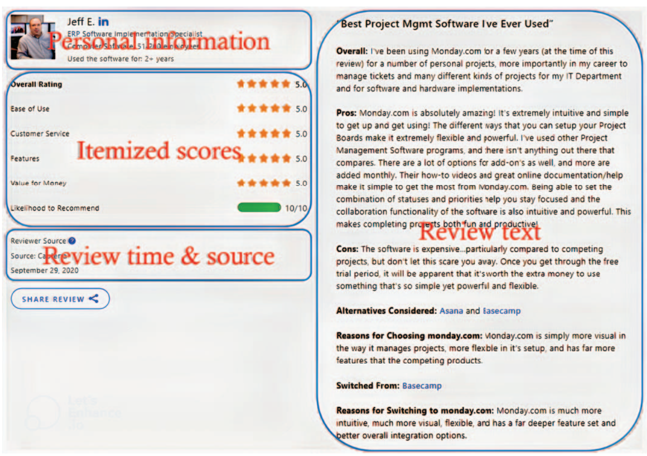

Fig. 1. A glance of the data from CACOO reviews 2022

A. Data Collection

Data was collected from software review websites, including Capterra [1], Software Advice [2], and GetApp [3], that contain user-revealed experiences. We gathered information from review sections, including "Personal Information", "Itemized Scores", "Review time & source", and "Review text", Fig. 1. "Personal information" may include the customer's name, the name abbreviation or nickname, software use duration, and more. The main focus of this study centers around "Itemized scores", which include Overall rating, ease of Use, features, value for money, and likelihood to recommend. "Review time & source" may include date and source; and the "Review text" mostly focuses on the Pros and Cons. We scraped 1189 software product reviews from review websites using Python code, selenium, and Beautiful Soup. The collected review information includes title, description, pros and cons, ratings, and review details such as name, date, company, and prior product used. Overall 62,423 non-repetitive reviews were gathered and stored in a CSV file containing 43 attributes.

B. Data Pre-processing

To pre-processed data for further analysis, we removed the "None" values from "incentivized" feature, leaving 49,998 instances in the dataset. We kept null values in other attributes to retain critical information. To binarize the "incentivized" feature, we categorized "NominalGift", "VendorReferredIncentivized", "NoIncentive", "NonNominalGift", and "VendorReferred", into two groups. Respectively, the first two were classified as "Incentive" and the last three as "NoIncentive".

Following data pre-processing applied for sentiment analysis. Expanding contractions were used to replace the short versions of the words with their complete forms to ensure that each word is treated as separate tokens that can further be analyzed individually. Non-alphabetic or non-numeric characters, such as punctuation marks, were removed in the other pre-processing step. Lemmatization used to increase sentiment accuracy while decreasing text dimension by reducing words to their base form, known as dictionary form. Tokenization is used to break the text into words for accurate sentiment prediction. Stop words such as "a" and "an" removed to reduce computational resources, lower text dimensionality, and improve sentiment analysis accuracy.

The data pre-processing followed by EDA analysis, sentiment analysis, and A/B testing. Beside "incentivized" feature, this study considers other attributes such as "overAllRating", "value_for_ money", "ease_of_use", "features", "customer_support", "likelihood_to_recommend", "year", "company_size", "time_used", "preprocessed_pros", "ReviewDescription_Sentiment", "source", "pros_Sentiment", "preprocessed_cons", "preprocessed_ReviewDescription", and "cons_Sentiment".

C. Data Analysis

1) EDA Analysis: We conducted EDA analysis to extract information, understand dataset characteristics, and identify variables' relationships.

2) Sentiment Analysis: Sentiment analysis was used to extract people's opinions [40] and compare emotional tones of incentive and organic reviews.

We used the HuggingFaceTransformers [4] for sentiment analysis. The model limitation of 200-char prevented us from combining all review texts (review description, pros, and cons). Therefore, to determine overall sentiment, We individually analyzed the sentiment of review texts and stored the results in the dataset as "ReviewDescription_Sentiment", "pros_Sentiment", and "cons_Sentiment".

As reviews rating scores reflect customers' satisfaction, Spearman's correlation coefficient is used to measure the

[1]https://www.capterra.com/project-management-software
[2]https://www.softwareadvice.com/project-management
[3]https://www.getapp.com/customer-management-software/crm
[4]https://github.com/huggingface/transformers

correlation between incentive and organic review rating scores based on sentiment, considering the review description, "incentivized", and sentiment status. To ensure accuracy, we analyzed sample of 4,000 reviews per review category and determined the 95% confidence interval (CI) using the z-test.

3) A/B Testing: A/B testing, a popular controlled experiment, known as split testing was conducted considering two alternatives, "Incentive" as (A) and "NoIncentive" as (B). Customer reviews were analyzed using "Incentive" and "NoIncentive" values to test the null hypothesis for significant difference between the two groups. The mean difference was measured between control and experimental groups using 10,000 repetitions. We ran six A/B tests comparing incentive and organic reviews across six rating attributes including "overAllRating", "value_for_ money", "ease_of_use", "features", "customer_support", and "likelihood_to_recommend".

IV. Results and Analysis

A. EAD Analysis results

After removing null values from "incentivized" feature, the EDA analysis revealed that among 49,998 remaining reviews, there are 44,255 Capterra, 3,485 Software Advice, and 2,258 GetApp reviews. The reviews were categorized into five groups including 29,466 "NominalGift", 3,272 "VendorReferredIncentivized", 16812 "NoIncentive", 90 "NonNominalGift", and 358 "VendorReferred". The first two groups with 32738 reviews labeled "Incentive", and the last three with 17,260 reviews labeled "NoIncentive".

B. Sentiment Analysis Results

Assessing the "incentivized" status of rating scores revealed more incentivized than organic reviews for scores 2 and higher. Higher volume of zero scores for incentive reviews led to decreasing product recommendation based on cost and customer support, Fig. 2.

Fig. 3 displays the tremendous escalation in software review volume, specific positive incentive reviews in 2018 and a sharp decline in 2020, revealing that changing circumstances have a greater impact on incentive than organic reviews. Rate increase may be due to increased social media usage, greater rewards for reviews, and genuine feedback posting. Declining review volume may result from COVID-19, preventing incentivized reviews, and decreasing customer trust caused by growing awareness.

Experienced users, more than two years of experience, using products tend to post positive and fewer negative incentive reviews, likely due to product familiarity and preference for benefits. Customers who use free trial post fewer reviews due to a lack of experience and confidence, they still contribute to post more incentive than organic reviews, Fig. 4.

Our study highlights that small companies with 11-50 employees, specific positive incentive reviews, have more than 7000 reviews. However, companies with 5,001-10,000 employees have less than 510 reviews. The significant gap between the number of incentive and organic reviews for smaller companies compare to larger ones is due to easier

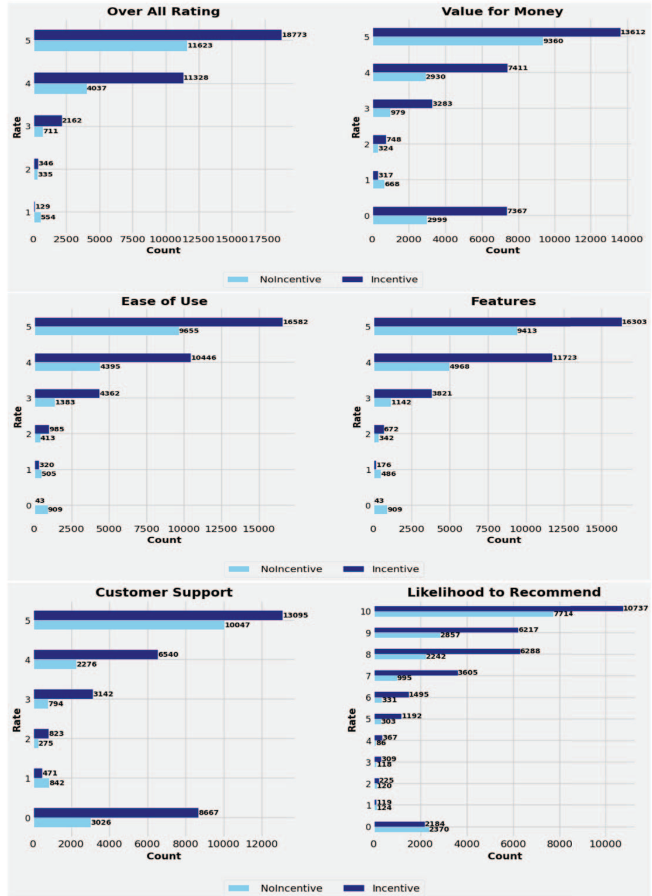

Fig. 2. Number of reviews for rating scores based on incentivized status

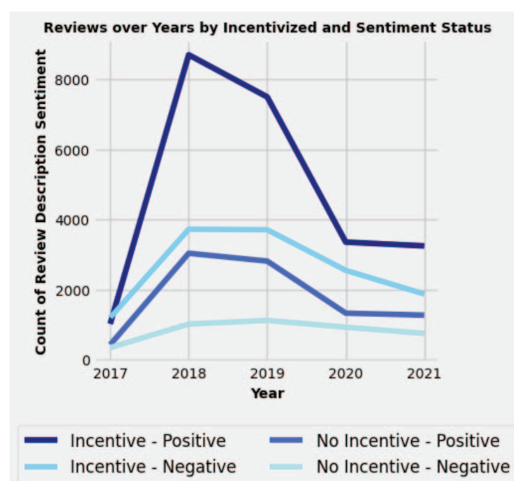

Fig. 3. Reviews over years by incentivized and sentiment status

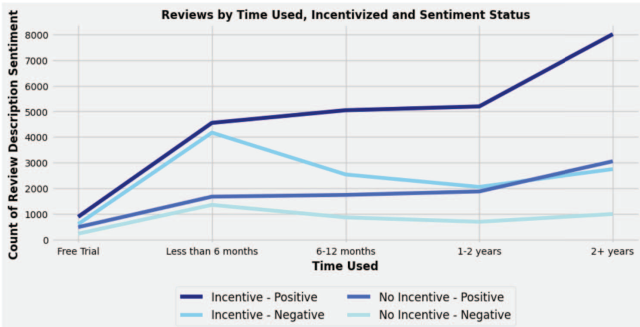

Fig. 4. Distribution of incentivized reviews by time used and sentiment

establishment and a higher likelihood of posting reviews, Fig. 5.

Analyzing review text sentiment considering "incentivized" status showed a higher incidence of positive sentiment in reviews' descriptions and pros, and negative sentiment in cons. Regardless of the review sentiment, volume of incentivized reviews is larger than organic reviews, Fig. 6.

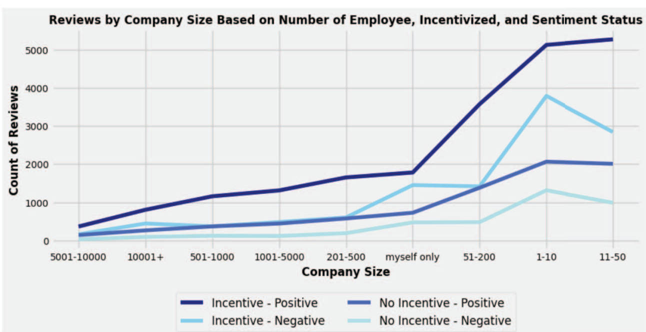

Fig. 5. Distribution of incentivized reviews by company size and sentiment

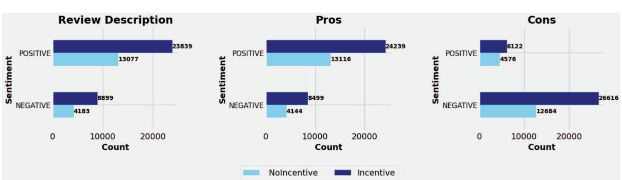

Fig. 6. Number of review description, pros, and cons based on incentivized and sentiment status

The word cloud is used to extract the top 20 words from each review text focusing on "incentivized". Considering sentiment status, the words such as "great" and "good" were frequently used in positive incentive and organic reviews for different review text, in addition to negative incentive and organic reviews for review description and pros. However, the top 20 words for negative incentive and organic do not include any negative words as expected. This could be due to removing negative words such as "not" as stop words, possibly causing the omission of negative phrases like "not good". Our results support prior research as incentives increase positive review length. Simultaneously, the volume of top 20 words is larger for positive incentives than organic and smaller for negative incentives than organic. Fig. 7 represents these results for review description.

Furthermore, measuring the average length of incentive and organic review descriptions based on the number of characters and their sentiment status reveals longer negative organic reviews, 153.91, than negative incentive reviews, 125.17. However, positive incentive reviews are longer, 104.13, than positive organic reviews, 96.45. These results are consistent with previous studies, [4], [5]. Overall these results for both negative incentive and organic reviews are higher than positive.

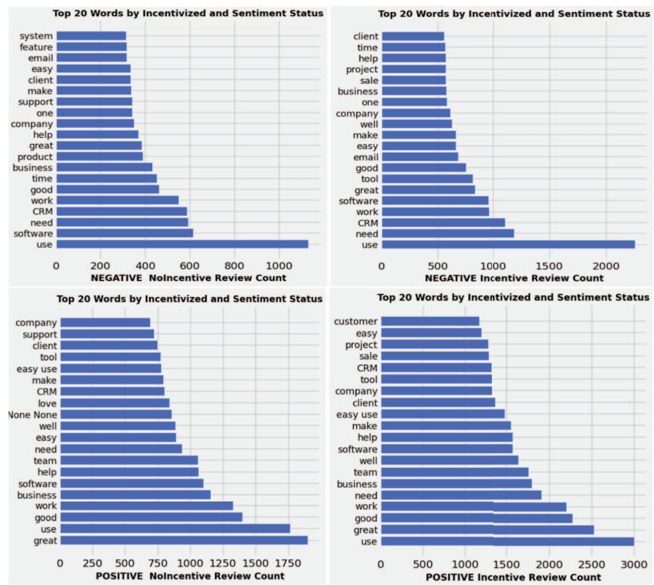

Fig. 7. Review description top 20 words based on incentivized and sentiment status

Our results of testing "Spearman's rank correlation coefficient" on review rating scores, considering 95% confidence interval, indicate stronger correlation among organic reviews, and more specifically, negative organic reviews, Fig. 8.

The highest correlation of 0.80 between "likelihood_to_recommend" and "overAllRating" appears to be influenced by high correlations between "overAllRating" with both "features" as 0.78 and "ease_of_use" as 0.76, in addition to the high correlation between "likelihood_to_recommend" with both "features" as 0.73 and "ease_of_use" as 0.72. Similar pattern with weaker positive correlations is seen in negative incentive reviews. Moreover, correlation between "features" and "ease_of_use", considering various statuses, supports the need for user-friendly software that provides easier access to features. In addition, the significant correlation between "value_for_money" and "customer_support" for negative reviews denotes weaker correlation for negative incentive , 0.60, compare to negative organic reviews, 0.66. All z-test results show 95% confidence in significant

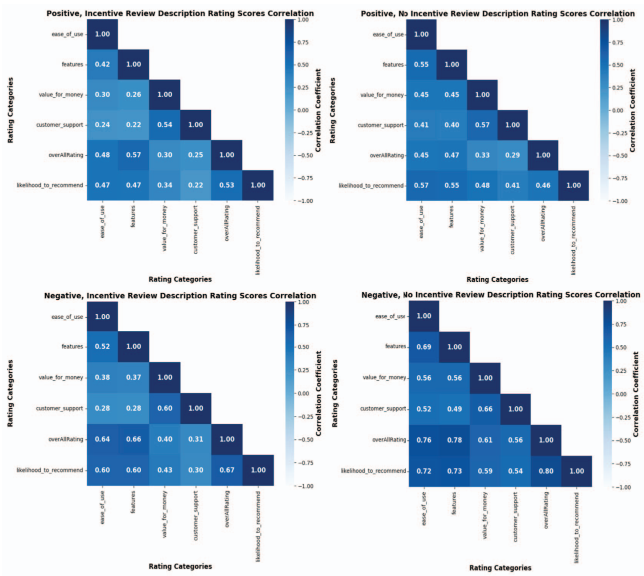

Fig. 8. Correlation among review rating scores by incentivized and review description sentiment status

correlations between review ratings due to p-values of zero.

C. A/B Testing Results

A/B test compared incentive and organic reviews for different rating scores,Table I. Organic review have a higher standard deviation (std) error for all rating scores despite higher total rating scores for incentive than organic reviews. For incentive reviews, lower std of "overAllRating" implies the consistency among ratings, and lower std error proposes a more precise estimate of the true mean. On the other hand, Organic reviews outperform significantly based on a p-value of 0.0014, with an overall impact on customer decision-making, indicated by the observed value of -0.0227. Cost-related ratings for organic reviews have a higher mean rating, and lower std that represents less variation in the rating. A P-value of 0.0000 indicates a significant difference between the two groups, and regarding the cost-related rating of the software reviews, incentives may not affect customers' decisions. In terms of ease of use, incentive reviews are rated higher. However, there is no statistically significant difference between incentive and organic reviews due to a p-value of 1.0000. Therefore, the observed value may not reflect true values, indicating insignificant impact of incentives on customer decision-making. For software review features, incentive reviews have a higher mean and consistent rating. Although the observed value of 0.1744 points to the difference between the two groups, the p-value of 1.0000 indicates no statistically significant impact of incentive reviews on customer decision-making. As discussed, customer support is essential for any product and its rating score shows a negative observed difference, meaning organic reviews have a higher average rating than expected. The significant difference between incentive and organic is because of the negative impact of incentive reviews on customer decision-making. The results of the customer's willingness to recommend products indicate more variability in organic review ratings. However, regarding willingness to recommend, the p-value of 1.0000 reveals no significant difference between groups, so it may not impact decision-making.

V. Discussion

In this section, we compare incentive and organic reviews, addressing the first research question, followed by answering the second research question by discussing the impact of incentives on customers' behavior toward posting reviews, review quality, and purchase decision-making.

A. Incentive vs. Organic

The analysis results prove that incentive reviews have more positive descriptions and pros, more negative cons, higher ratings, and minority with lower scores. Despite organic reviews, the volume of incentive reviews has changed dramatically over the years, revealing the dependency on various factors, including environmental situations (e.g., pandemics and economic problems). In addition, companies may shift their encouragement plan from offering incentives to focus on improved advertising, information sharing, consumer awareness, and distrust of review authenticity. The incentive volume can grow by growing social platforms, improving customer experience, and expanding smaller companies. Based on the result of A/B testing, incentive reviews have the higher sum of the "total rating" and lower std error for all rating scores. Overall rating is more consistent for incentive reviews.

B. Incentive Review and Customer Behavior

To answer "How do incentives impact customer behavior on posting purchase's review?", we rely on our findings from the first research question.

The analysis proves that incentives boost reviews, as the volume of incentive reviews is almost double to compare with organic reviews. Reviewers tend to rate the reviews positively, despite providing negative feedback. Higher sum of rating scores for incentive reviews compare to organic reviews indicate customers are more likely rewarded for posting positive reviews.

The dramatic alteration in the distribution of incentives over the years proves rewards as review posters' motivation. Over time, factors like commerce, economy, social networks, environmental issues, and technology can reduce, restrict, or eliminate incentives from the business platform, causing users to post fewer reviews. Despite massive changes in volume over the years, incentive reviews consistently outnumber organic reviews indicating the impact of incentives on customers' review behavior toward posting reviews. Small businesses may incentivize individuals to write incentive or fake reviews to compete in the business world and increase profits. The better product understanding enhances incentive reviews quality and quantity. Furthermore, customer support quality impacts product cost satisfaction for many customers.

TABLE I
THE STATISTICAL VALUES AND RESULTS OF A/B TESTING

Attribute	incentivized	total	sum_total_rating	mean_value	std	std_error	observed_difference	empirical_P
overAllRating	NoIncentive	17260	77620	4.497	0.913	0.007	-0.0227	0.0014
	Incentive	32738	146484	4.474	0.702	0.004		
value_for_money	NoIncentive	17260	62773	3.637	1.916	0.015	-0.2963	0.0000
	Incentive	32738	109366	3.341	1.965	0.011		
ease_of_use	NoIncentive	17260	71335	4.133	1.350	0.010	0.1455	1.0000
	Incentive	32738	140070	4.279	0.890	0.005		
features	NoIncentive	17260	71533	4.144	1.329	0.010	0.1744	1.0000
	Incentive	32738	141390	4.319	0.815	0.005		
customer_support	NoIncentive	17260	63113	3.657	1.954	0.015	-0.5050	0.0000
	Incentive	32738	103178	3.152	2.060	0.011		
likelihood_to_recommned	NoIncentive	17260	132317	7.666	3.431	0.026	0.1766	1.0000
	Incentive	32738	256756	7.843	2.685	0.015		

However, some users are discouraged from writing incentive reviews when they become aware of the potential for biased or suspicious content.

C. Incentive Review and Review Quality

Based on the analysis, higher volume of incentive reviews displays lower credibility and higher bias as they may contain non-experience-based information aimed at boosting review quantity and rating. A greater volume of incentive reviews for smaller companies may indicate bias and fake reviews, reducing credibility and consistency of the review quality. Although reviews from experienced users are more credible, those who receive incentives for their reviews tend to be less consistent in their rating compared to organic reviews, due to a significant increase in positive incentive and decrease in negative incentive reviews. In terms of cost and customer support, the significant number of zero rating scores for incentivized than organic reviews proves that incentives do not always increase positivity. Incentive and organic reviews show similar zero-rate volumes for recommendation likelihood, indicating greater consistency in organic reviews. Higher negative-to-positive cons ratio than positive-to-negative pros ratio, even for incentive reviews, suggests customer sensitivity to writing negative feedback, increasing the credibility of negative reviews. Negative software review ratings correlate more strongly than positive ratings, which may support that negative reviews are often more credible [16]. Based on the evidence, incentive reviews show inconsistent volume.

Furthermore, high volume of top 20 words in positive incentive reviews suggests a possible bias and reduced credibility compare to organic reviews. In comparison, the higher volume of these words in negative organic reviews indicates more detailed reviews, increasing credibility of organic reviews. Longer negative organic review descriptions that provide more information and detail reveals higher credibility and lower bias, outweighing incentive reviews.

To this point, our discussion on review credibility and consistency mainly focuses on reviews volume and length. Higher sum of rating scores for incentive reviews, considering different rating scores statistical values, may indicate motivated posting reviews for rewards, raising credibility concern. Offering rewards for incentive reviews reduces diversity and increases consistency in review content, despite organic reviews.

D. Incentive Review and Purchase Decision-making

Higher incentive review volume points to focusing more on the overall rating and quantity over review content. Therefore, lack of accurate and comprehensive view of the products/services, cause less consistency in incentive reviews, resulting in lower credibility and uninformed purchase decisions. Willingness to post positive incentivized reviews, based on the higher sum of all rating scores, may indicate excessive positivity, impacting customers' purchase decision. Referring to observed differences and p-values from the A/B testing results, incentive reviews have less impact on customer purchase decision based on their overall rating, may not impact customer purchase decision regarding software cost and software features, and have no significant impact on decision-making based on easiness of use. In addition, incentive reviews negatively affect customer decision-making, as shown by the customer support score evaluation.

E. Implications of the Study

Despite existing studies, our approach distinguishes our work by evaluating incentive reviews quality differently. We used EDA, sentiment analysis, and A/B testing to compare incentive and organic reviews quality and determine the impact of incentives on customers' behavior for posting reviews. Although incentive software reviews outnumber organic reviews by almost two-fold, this could be changed in either increasing or decreasing direction due to factors such as time, business platform and size, and user awareness and experience. Factors such as cost, software features, ease of use, and customer support impact software product ratings of either incentive or organic reviews. This is due to the high correlation between these features or with software's overall rating and recommendation. Furthermore, our A/B testing shows that high volume and rating of incentive reviews may not significantly affect customer purchase decisions.

Our findings could guide enhancing software review quality to improve software products and recommendation systems that aid purchase decision-making.

VI. Conclusion and Future Work

Studies may have opposing results for the same situation due to differences in population under study, research methods, and approaches used. For instance, Woolley & Sharif (2021) [6] highlights more enjoyable writing reviews with incentives, while Garnefeld et al. (2020) [9] emphasize incentives role in increasing the review rate. However, Burtch et al. (2018) [4] remark the delivering negative review by incentivized customers. While our findings support some of the existing research, the emergence of these opposing findings underscores the need for further investigation across diverse populations of various sizes and cultures. It is imperative to examine different methods and approaches to reach broadly applicable outcomes to improve the quality of online review systems. Potential collaborations among companies can achieve such outcomes.

A. Strengths and Limitations

This study has broad applicability across various research domains and is not restricted to the fast-growing world of reviews. Increasing business growth and product diversity intensifies producer competition for sales to companies, which in turn sell products to consumers. This highlight the necessity of accessing high-quality reviews. This highlights the need to access high-quality reviews. Our study has the potential to enhance the performance of product review systems, boosting customer satisfaction and efficiency by saving time and money.

Moreover, our unique contribution to the review quality study that achieves by focusing on software review quality, specifically incentive reviews, distinguishes our work from existing research in the field.

While this work has several strengths, it also exhibits some weaknesses. Current methods cannot accurately determine review sentiment due to the subjective nature of reviews, which involves human emotion and expressions. A large dataset prevents using human power to annotate reviews sentiment. Even human annotation does not ensure result accuracy due to potential human error. On the other hand, due to inability to recognize emotions and expressions, automated annotation falls short in achieving higher accuracy.

Additionally, the model's constraint limited our sentiment analysis to analyzing only up to 200 characters. Therefore, our model may have missed important aspects of some reviews as their length exceeded this limitation.

B. Future Work

We analyzed purchase review differences and assessed review credibility and consistency to evaluate the influence of incentive reviews on customer decision-making. We assessed review quality before studying its influence on purchase decisions. We discussed research problems and our findings through EDA on sentiment analysis and A/B testing. However, for future work, we plan to survey software reviews to gather and analyze information from new users, considering the subjectivity of online review and purchase decisions. Additionally, We plan to explore review quality dimensions that affect purchasing decisions, specifically objectivity [41], [42], depth [43], [44], authenticity [45], [46] and ultimately, helpfulness [47], [48], taking into consideration the incentivized status of the reviews. This will allow us to compare users' perspectives on incentive reviews' quality and their impact on purchase decisions.

References

[1] L. Zhu, H. Li, F. K. Wang, W. He, and Z. Tian, "How online reviews affect purchase intention: a new model based on the stimulus-organism-response (S-O-R) framework," *Aslib Journal of Information Management*, vol. 72, no. 4, pp. 463–488, nov 2020.

[2] M. Petrescu, K. O'Leary, D. Goldring, and S. Ben Mrad, "Incentivized reviews: Promising the moon for a few stars," *Journal of Retailing and Consumer Services*, vol. 41, no. December 2016, pp. 288–295, 2018.

[3] J. Ai, D. Gursoy, Y. Liu, and X. Lv, "Effects of offering incentives for reviews on trust: Role of review quality and incentive source," *International Journal of Hospitality Management*, vol. 100, no. February 2021, p. 103101, 2022. [Online]. Available: https://doi.org/10.1016/j.ijhm.2021.103101

[4] G. Burtch, Y. Hong, R. Bapna, and V. Griskevicius, "Stimulating Online Reviews by Combining Financial Incentives and Social Norms," *Management Science*, vol. 64, no. 5, pp. 2065–2082, may 2018. [Online]. Available: http://pubsonline.informs.org/doi/10.1287/mnsc.2016.2715

[5] A. Costa, J. Guerreiro, S. Moro, and R. Henriques, "Unfolding the characteristics of incentivized online reviews," *Journal of Retailing and Consumer Services*, vol. 47, pp. 272–281, mar 2019.

[6] K. Woolley and M. A. Sharif, "Incentives Increase Relative Positivity of Review Content and Enjoyment of Review Writing," *Journal of Marketing Research*, vol. 58, no. 3, pp. 539–558, jun 2021.

[7] M. Niaz Imtiaz, T. Ahmed, and A. Paul, "Incentivized Comment Detection with Sentiment Analysis on Online Hotel Reviews," 2020. [Online]. Available: https://doi.org/10.22541/au.159559938.84764895

[8] M. Zhang, X. Wei, and D. D. Zeng, "A matter of reevaluation: Incentivizing users to contribute reviews in online platforms," *Decision Support Systems*, vol. 128, jan 2020.

[9] I. Garnefeld, S. Helm, and A. K. Grötschel, "May we buy your love? psychological effects of incentives on writing likelihood and valence of online product reviews," *Electronic Markets*, vol. 30, no. 4, pp. 805–820, 2020.

[10] G. Cui, Y. Chung, L. Peng, and W. Zheng, "The importance of being earnest: Mandatory vs. voluntary disclosure of incentives for online product reviews," *Journal of Business Research*, vol. 141, pp. 633–645, mar 2022.

[11] M. Luca and G. Zervas, "Fake it till you make it: Reputation, competition, and yelp review fraud," in *Management Science*, vol. 62, no. 12, 2016, pp. 3412–3427. [Online]. Available: http://pubsonline.informs.org/journal/mnsc

[12] V. Kamble, N. Shah, D. Marn, A. Parekh, and K. Ramchandran, "The Square-Root Agreement Rule for Incentivizing Objective Feedback in Online Platforms," *SSRN Electronic Journal*, 2019. [Online]. Available: https://ssrn.com/abstract=3488831

[13] Z. Liu, S. hui Lei, Y. lang Guo, and Z. ang Zhou, "The interaction effect of online review language style and product type on consumers' purchase intentions," *Palgrave Communications*, vol. 6, no. 1, pp. 1–8, 2020. [Online]. Available: http://dx.doi.org/10.1057/s41599-020-0387-6

[14] L. Tuan, P. Thi, M. Ly, N. Thanh, L. Trong, and T. Tran, "Online reviews as a pacifying decision-making assistant," *Journal of Retailing and Consumer Services*, vol. 64, no. September 2021, p. 102805, 2022. [Online]. Available: https://doi.org/10.1016/j.jretconser.2021.102805

[15] H. Zhang, A. Yang, A. Peng, L. F. Pieptea, J. Yang, and J. Ding, "A Quantitative Study of Software Reviews Using Content Analysis Methods," *IEEE Access*, vol. 10, no. December, pp. 124 663–124 672, 2022.

[16] S. Kusumasondjaja, T. Shanka, and C. Marchegiani, "Credibility of online reviews and initial trust: The roles of reviewer's identity and review valence," *Journal of Vacation Marketing*, vol. 18, no. 3, pp. 185–195, 2012.

[17] S. Jamshidi, R. Rejaie, and J. Li, "Characterizing the dynamics and evolution of incentivized online reviews on Amazon," *Social Network Analysis and Mining*, vol. 9, no. 1, 2019. [Online]. Available: https://www.amazon.com/gp/bestsellers/

[18] U. Gneezy, S. Meier, and P. Rey-Biel, “When and why incentives (don’t) work to modify behavior,” *Journal of Economic Perspectives*, vol. 25, no. 4, pp. 191–210, 2011.
[19] T. Chen, P. Samaranayake, X. Cen, M. Qi, and Y.-C. Lan, “The Impact of Online Reviews on Consumers’ Purchasing Decisions: Evidence from an Eye-Tracking Study,” *Frontiers in Psychology*, vol. 13, no. June, 2022.
[20] N. Aghakhani, O. Oh, D. Gregg, and H. Jain, “How Review Quality and Source Credibility Interacts to Affect Review Usefulness: An Expansion of the Elaboration Likelihood Model,” *Information Systems Frontiers*, no. 0123456789, 2022. [Online]. Available: https://doi.org/10.1007/s10796-022-10299-w
[21] U. Chakraborty and S. Bhat, “Credibility of online reviews and its impact on brand image,” *Management Research Review*, vol. 41, no. 1, pp. 148–164, 2018.
[22] R. Filieri, C. F. Hofacker, and S. Alguezaui, “What makes information in online consumer reviews diagnostic over time? The role of review relevancy, factuality, currency, source credibility and ranking score,” *Computers in Human Behavior*, vol. 80, pp. 122–131, 2018. [Online]. Available: https://doi.org/10.1016/j.chb.2017.10.039
[23] J. Mackiewicz, D. Yeats, and T. Thornton, “The Impact of Review Environment on Review Credibility,” *IEEE Transactions on Professional Communication*, vol. 59, no. 2, pp. 71–88, 2016.
[24] S. W. Hung, C. W. Chang, and S. Y. Chen, “Beyond a bunch of reviews: The quality and quantity of electronic word-of-mouth,” *Information and Management*, vol. 60, no. 3, p. 103777, 2023. [Online]. Available: https://doi.org/10.1016/j.im.2023.103777
[25] N. Aghakhani, O. Oh, and D. Gregg, “Beyond the Review Sentiment: The Effect of Review Accuracy and Review Consistency on Review Usefulness,” *ICIS 2017: Transforming Society with Digital Innovation*, no. April, 2018.
[26] K. L. Xie, C. Chen, and S. Wu, “Online Consumer Review Factors Affecting Offline Hotel Popularity: Evidence from Tripadvisor,” *Journal of Travel and Tourism Marketing*, vol. 33, no. 2, pp. 211–223, 2016. [Online]. Available: http://dx.doi.org/10.1080/10548408.2015.1050538
[27] N. Aghakhani, O. Oh, D. G. Gregg, and J. Karimi, “Online Review Consistency Matters: An Elaboration Likelihood Model Perspective,” *Information Systems Frontiers*, vol. 23, no. 5, pp. 1287–1301, 2021.
[28] K. Zhao, A. C. Stylianou, and Y. Zheng, “Sources and impacts of social influence from online anonymous user reviews,” *Information and Management*, vol. 55, no. 1, pp. 16–30, 2018. [Online]. Available: https://doi.org/10.1016/j.im.2017.03.006
[29] V. D. Tran, M. D. Nguyen, and L. A. Lng, “The effects of online credible review on brand trust dimensions and willingness to buy: Evidence from Vietnam consumers,” *Cogent Business and Management*, vol. 9, no. 1, 2022. [Online]. Available: https://doi.org/10.1080/23311975.2022.2038840
[30] H. H. Wu, P. Tipgomut, H. F. Chung, and W. K. Chu, “The mechanism of positive emotions linking consumer review consistency to brand attitudes: A moderated mediation analysis,” *Asia Pacific Journal of Marketing and Logistics*, vol. 32, no. 2, pp. 575–588, 2020.
[31] E. Bigne, K. Chatzipanagiotou, and C. Ruiz, “Pictorial content, sequence of conflicting online reviews and consumer decision-making: The stimulus-organism-response model revisited,” *Journal of Business Research*, vol. 115, pp. 403–416, jul 2020.
[32] M. H. Gerrath and B. Usrey, “The impact of influencer motives and commonness perceptions on follower reactions toward incentivized reviews,” *International Journal of Research in Marketing*, vol. 38, no. 3, pp. 531–548, sep 2021.
[33] C. Du Plessis, A. T. Stephen, Y. Bart, and D. Goncalves, “When in Doubt, Elaborate? How Elaboration on Uncertainty Influences the Persuasiveness of Consumer-Generated Product Reviews When Reviewers Are Incentivized,” *SSRN Electronic Journal*, 2017. [Online]. Available: http://ssrn.com/abstract=2821641
[34] H. Yin, S. Zheng, W. Yeoh, and J. Ren, “How online review richness impacts sales: An attribute substitution perspective,” *Journal of the Association for Information Science and Technology*, vol. 72, no. 7, pp. 901–917, 2021.
[35] M. Siering, J. Muntermann, and B. Rajagopalan, “Explaining and predicting online review helpfulness: The role of content and reviewer-related signals,” *Decision Support Systems*, vol. 108, pp. 1–12, 2018. [Online]. Available: https://doi.org/10.1016/j.dss.2018.01.004
[36] S. Jamshidi, R. Rejaie, and J. Li, “Trojan horses in amazon’s castle: Understanding the incentivized online reviews,” in *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018*, 2018, pp. 335–342. [Online]. Available: https://www.amazon.com/gp/bestsellers/
[37] Y. Jia and I. L. Liu, “Do consumers always follow “useful” reviews? The interaction effect of review valence and review usefulness on consumers’ purchase decisions,” *Journal of the Association for Information Science and Technology*, vol. 69, no. 11, pp. 1304–1317, 2018.
[38] M. Tang, Z. Xu, Y. Qin, C. Su, Y. Zhu, F. Tao, and J. Ding, “A Quantitative Study of Impact of Incentive to Quality of Software Reviews,” *Proceedings - 2022 9th International Conference on Dependable Systems and Their Applications, DSA 2022*, pp. 54–63, 2022.
[39] X. Li, C. Wu, and F. Mai, “The effect of online reviews on product sales: A joint sentiment-topic analysis,” *Information and Management*, vol. 56, no. 2, pp. 172–184, 2019.
[40] M. E. Basiri, N. Ghasem-Aghaee, and A. R. Naghsh-Nilchi, “Exploiting reviewers’ comment histories for sentiment analysis,” *Journal of Information Science*, vol. 40, no. 3, pp. 313–328, 2014.
[41] M. Hair and T. Ozcan, “How reviewers’ use of profanity affects perceived usefulness of online reviews,” *Marketing Letters*, vol. 29, no. 2, pp. 151–163, 2018.
[42] C. Luo, X. Luo, Y. Xu, M. Warkentin, and C. L. Sia, “Examining the moderating role of sense of membership in online review evaluations,” *Information and Management*, vol. 52, no. 3, pp. 305–316, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.im.2014.12.008
[43] S. Bi, Z. Liu, and K. Usman, “The influence of online information on investing decisions of reward-based crowdfunding,” *Journal of Business Research*, vol. 71, pp. 10–18, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.jbusres.2016.10.001
[44] C. Janze and M. Siering, “” Status effect” in user-generated content: Evidence from online service reviews,” *2015 International Conference on Information Systems: Exploring the Information Frontier, ICIS 2015*, pp. 1–15, 2015.
[45] S. Chatterjee, R. Chaudhuri, A. Kumar, C. Lu Wang, and S. Gupta, “Impacts of consumer cognitive process to ascertain online fake review: A cognitive dissonance theory approach,” *Journal of Business Research*, vol. 154, no. January 2022, p. 113370, 2023. [Online]. Available: https://doi.org/10.1016/j.jbusres.2022.113370
[46] C. L. Campagna, N. Donthu, and B. Yoo, “Brand authenticity: literature review, comprehensive definition, and an amalgamated scale,” *Journal of Marketing Theory and Practice*, vol. 31, no. 2, pp. 129–145, 2021.
[47] C. Xu, X. Zheng, and F. Yang, “Examining the effects of negative emotions on review helpfulness: The moderating role of product price,” *Computers in Human Behavior*, vol. 139, no. September 2022, p. 107501, 2023. [Online]. Available: https://doi.org/10.1016/j.chb.2022.107501
[48] L. Luo, J. Liu, H. Shen, and Y. Lai, “Vote or not? How language mimicry affect peer recognition in an online social QA community,” *Neurocomputing*, vol. 530, pp. 139–149, 2023. [Online]. Available: https://doi.org/10.1016/j.neucom.2023.01.086