Does AI-Generated Review Summarization Affect Consumer Purchasing Behavior?— An Empirical Study Based on the Amazon Platform

Huike Wang
Central University of Finance and Economics
2023212424@email.cufe.edu.cn

Tianmei Wang*
Central University of Finance and Economics
wangtianmei@cufe.edu.cn

Abstract

The product reviews on e-commerce platforms have become an important reference for consumers' purchase decisions. However, the sheer volume of reviews can lead to information overload, which may negatively affect consumer purchasing behavior. The AI-generated review summarization brings an opportunity to solve this dilemma. We collected 48,019 panel data of 3,781 products from the Amazon platform between February 2023 and February 2024, and employed a difference-in-differences model to investigate the impact mechanism of AI-generated review summarization on consumer purchasing behavior. The research found that AI-generated review summarization promotes consumer purchasing behavior. The number of product reviews has a U-shaped moderating effect, and the average product rating and rating dispersion both have positive moderating effects. Additionally, compared to experience products, search products amplify the impact of AI-generated review summarization on consumer purchasing behavior. The findings offer theoretical support for enhancing the design of AI-human interactions on e-commerce platforms.

Keywords: AI-generated review summarization, User-generated content, Purchasing behavior, Difference-in-Differences model, Online shopping

1. Introduction

With the rapid development of internet technology and smart devices, online shopping has become the norm for most consumers. Product reviews on e-commerce platforms play a crucial role in consumers' purchase decisions, which are valuable in reducing information asymmetry and facilitating better purchasing choices (Filieri, 2015; Mudambi & Schuff, 2010). As the number of online reviews continues to grow, the information overload caused by massive online reviews (Hu & Krishen, 2019) makes it difficult for consumers to quickly capture the key product features or usage experiences they need, and search costs skyrocket (Jabr & Rahman, 2022), which reduces purchase intentions or delays purchase decisions, or even decision avoidance (Kaushik et al., 2018). The advancement of AI technology provides a solution to this problem (Stolz et al., 2024). In August 2023, Amazon introduced the AI-generated review summarization feature, which utilizes artificial intelligence to create concise summaries of consumer reviews and places them at the top of the review section after filtering out identified fake or useless reviews. In addition to providing a overview summary text, the summaries will highlight key product attributes such as performance, ease of use, and stability. Each attribute will be assigned a sentiment label, and the AI will also generate separate summary for each label (Schermerhorn, 2023). It is evident that AI-generated review summarization aims to help consumers understand products more quickly and effectively, thereby enhancing their purchase intentions.

Interestingly, scholars hold two opposing viewpoints on the effectiveness of AI-generated summarization in real-world applications. On the one hand, AI-generated summarization can rapidly extract key information from large volumes of data, and enable consumers to efficiently obtain the important information, which significantly reduces search costs and improves decision-making efficiency (Liu et al., 2024; Mohammad Rajiur Rahman et al., 2023). AI-generated summarization can distill complex information into easily comprehensible texts. It is especially beneficial for time-pressed consumers, as it significantly optimizes their product browsing and decision-making processes. On the other hand, AI-generated summarization, as an emerging technology, may miss some important details during information processing, which can result in summaries that are not sufficiently comprehensive or accurate, eventually leading to information processing biases. The summarization currently lacks credibility, and its accuracy needs improvement (Goodman et al., 2024). Therefore, despite the obvious technological advantages of AI-generated summarization, there is no consensus on its impact on consumer purchasing behavior on e-commerce platforms.

Cognitive Load Theory (CLT) can be applied to explain how consumers evaluate and process information for decision-making. In fact, consumers

URI: https://hdl.handle.net/10125/109399
978-0-9981331-8-8

have limited cognitive resources (Sweller, 1988). The huge amount of online reviews accumulated on e-commerce platforms, which contain substantial noisy data, present significant challenges related to information redundancy and low value. Based on the CLT, this study argues that AI-generated review summarization, which provides a high-level summary and refinement of online reviews, alters the way consumers process information and make decisions. It helps reduce uncertainty and cognitive costs, then actively promoting consumer purchasing behavior. Therefore, this study attempts to use CLT to reveal whether AI-generated summarization affects consumer purchasing behavior and what differences exist in its impact on consumer purchasing behavior. This study provides a theoretical foundation for enhancing the design of AI-human interactions on e-commerce platforms. Next, this paper will have the following sections. We will introduce the related literature in Section 2 to review the research on the impact of user-generated content on consumer purchase behavior and its intrinsic mechanism. Section 3 will present the research model and hypotheses. Section 4 will describe the data and methods used in this study, followed by Section 5, which will give and explain the empirical results. Finally, Section 6 will summarize the research findings, provide managerial implications, and offer suggestions for future research directions.

2. Related work

2.1. User-generated content: online reviews

In online shopping, potential consumers mainly rely on two sources of signals to obtain product information, merchants and peer consumers (Cheung et al., 2014). Merchants usually convey objective product information to consumers, while peer consumers primarily provide subjective feedback on products through online reviews (Wells et al., 2011). Online reviews can effectively reduce information asymmetry and have a greater impact on the purchasing behavior of potential consumers (Cheung et al., 2014). Focusing on online reviews as a key type of user-generated content, scholars have mainly explored its impact on consumer purchasing behavior from three perspectives: textual information, social interaction information, and multimedia information.

Textual information is the most visual source of information presented in online reviews. Some studies have explored the impact of review characteristics on consumer purchase decisions. Characteristics such as review length and topic diversity positively impact product sales (Zhai et al., 2024). It has also been demonstrated that sentiment disposition significantly affects consumers' purchase intentions (Gao et al., 2024), and in particular, negative reviews (Kaushik et al., 2018) or extreme reviews (Román et al., 2023) can adversely affect product sales. Social interaction information refers to the interactive behaviors that consumers use to express their personal views in e-commerce platforms, such as rating purchased products, liking or commenting on other users' reviews. Social interactive information can further break down information barriers, reduce consumers' perceived uncertainty, and then influence purchase decisions, which has become a research hotspot. Ratings can intuitively reflect consumers' overall attitudes towards products. Ratings from proprietary platforms (Filieri, 2015), third-party platforms (Luo et al., 2021), and the raters themselves (Luo et al., 2021) all influence consumers' purchase intentions. Additionally, likes (Liang et al., 2023) and the Q&A feature (Warut Khern-am-nuai et al., 2023) in the review section, as a more agile interaction method, can strengthen social connections between consumers, enhance their sense of belonging and loyalty to the e-commerce platforms, and thereby increase their purchase intentions. Multimedia information, such as product images or videos posted by consumers in online reviews, which serves as tangible feedback on product experiences, has also attracted significant scholarly attention. Research indicates that videos in the review section positively affect consumers' perception of review usefulness, thus increasing their purchase intentions (Park et al., 2023). Images in the review section can mitigate the cognitive costs associated with textual information overload, but they can also create expectation gaps, negatively impacting post-purchase satisfaction (Guan et al., 2023).

It is evident that online reviews, as one of the most important user-generated content, significantly influence consumer purchasing behavior. AI-generated review summarization is essentially an intelligent text derived from product reviews. Compared to traditional online reviews, its advantages include quickly extracting key product features and reducing consumers' cognitive load. However, its drawbacks may include the potential to filter out some highlights or interesting comments that consumers would like to see. Therefore, the impact of AI-generated review summarization on consumer purchasing behavior deserves to be explored in depth.

2.2. User-generated content and consumer purchasing behavior

Reviewing the existing research, the intrinsic mechanisms by which user-generated content influences consumer purchasing behavior can be divided into the content quality and information processing. On the one hand, enhancing review quality is crucial to help consumers perceive useful information, and ultimately drives purchasing behavior (Beck et al., 2023). Comprehensive and specific reviews provide more product details and usage feedback, reflecting a higher quality of reviews, which can effectively enhance consumers' perception of the usefulness of reviews, and in turn, significantly increase consumers' purchase intentions (Filieri, 2015; Luo et al., 2021; Mudambi & Schuff, 2010). Additionally, Yin et al. (2022) found that review content consistency also affects potential consumers' purchase intensions, and reviews with higher consistency are more likely to diminish consumers' perceived uncertainty, which in turn facilitates purchases.

On the other hand, information processing also plays a crucial role in influencing consumer decision-making behavior. Since online reviews disclose an excessive amount of detailed information, consumers are likely to suffer from information overload when facing a large volume of reviews. Due to limited processing capacity, it is difficult for consumers to make rational decisions, ultimately reducing their purchase intentions (Ma et al., 2022; Zhai et al., 2024). Therefore, many scholars are exploring how to balance review overload and consumers' cognitive costs through information processing mechanism. Some studies have utilized information processing methods to enhance the presentation of UGC, thereby promoting consumer purchasing behavior. Presenting selected and concise reviews (Jabr & Rahman, 2022) can significantly reduce consumers' information burden, improving the efficiency and satisfaction of their purchase decisions. Pinning high-quality, helpful and popular reviews positively affects product sales, whereas pinned reviews with negative sentiment can inhibit consumers' purchase intentions (Kaushik et al., 2018). Additionally, some studies have also focused on how e-commerce platforms use information processing techniques such as text mining to demonstrate key features of UGC. Specific keywords in online reviews (Lu et al., 2023) can attract consumers' attention. Similarly, the text and color of tags in the review section (Xu & Zhang, 2018) can also influence consumer purchasing decisions. However, excessive keywords or tags can even lead to information overload, thus reducing purchase intentions.

It is evident that UGC influences consumer purchasing behavior by enhancing content quality and information processing mechanisms. On the one hand, AI-generated review summarization, as an emerging form of AI-pinned review, affects consumers' perceived usefulness of the reviews, which in turn influences purchase intentions. On the other hand, AI-generated review summarization extracts the key information from reviews by means of information processing, which affects consumers' information processing load. Additionally, existing studies have confirmed that AI-generated summarization does influence user behavior. It has been shown to enhance consumers' perceived usefulness of video content in online learning (Mohammad Rajiur Rahman et al., 2023) and optimize clinical decision-making in online healthcare (Liu et al., 2024). However, consumer purchasing behavior in online shopping differs from that in online learning and healthcare scenarios, making the research conclusion less universally applicable. Therefore, this paper focuses on the AI-generated review summarization feature introduced on the Amazon.com to explore how this technology impacts consumer purchasing behavior in online shopping.

3. Research model and hypotheses

This study utilizes Cognitive Load Theory (CLT) as the theoretical foundation to understand how AI-generated review summarization influences consumer purchasing behavior. CLT explains how individuals allocate and utilize cognitive resources when solving problems or making decisions, reflecting the costs and burdens imposed on working memory by the focal task (Sweller, 1988). In online shopping situations, cognitive processing plays a crucial role in the purchase process (Racat & Plotkina, 2023). According to the CLT, individuals have limited cognitive resources. These resources are consumed during internet activities such as online product information retrieval. If cognitive resources are exhausted or the cognitive demands of the task increase, it may result in cognitive deterioration, thereby affecting decision-making (Li et al., 2024). E-commerce platforms accumulate a vast amount of product reviews, which leads to issues of information redundancy and low-value content, making the cognitive demands of the purchasing task excessively high (Jabr & Rahman, 2022). Consumers inevitably consume a large amount of cognitive resources when filtering and processing information. When the cognitive cost required to process reviews exceeds their limited cognitive resources, high cognitive load can reduce their purchase intentions (Shobhit Kakaria et al., 2023).

Amazon has officially stated that the primary intention behind launching the AI-generated review summarization feature is to help consumers quickly capture key information from product reviews, reduce the cognitive cost of retrieving product information. (Schermerhorn, 2023). In online shopping, although consumers can obtain more useful information by browsing product reviews, AI-generated review summarization can quickly extract important features from massive reviews. It presents consumers with concise, processed review texts, thus reducing the cognitive resources required to browse product information. This achieves an effective balance between cognitive resources and information processing, optimizing the purchase decision-making process. Collectively, we propose hypothesis one:

H1: *AI-generated review summarization has a positive effect on consumer purchasing behavior.*

The number of online reviews reflects a product's popularity and the information load within the reviews (Hu & Krishen, 2019), which plays an important role in consumer purchasing decisions. According to CLT, consumers have limited cognitive resources (Sweller, 1988). In scenarios with relatively few reviews, the information processing cost for consumers to browse the reviews does not exceed their cognitive resource limits, making them more willing to read all or most of the reviews. Additionally, when dealing with a smaller dataset of reviews, the content accuracy of AI-driven summarization also tends to be lower (Peal et al., 2022). As the number of reviews increases, AI-generated review summaries may struggle to maintain strong consistency with the content of the multiple reviews viewed. This inconsistency can lead to a reduction in consumers' purchase intentions, potentially hindering their final purchase decisions (Yin et al., 2022).

However, when the number of reviews accumulates to a certain threshold, the information processing cost exceeds consumers' available cognitive resources, leading to information overload and a reduced willingness to read reviews one by one. Moreover, with a larger dataset of reviews, the AI's ability to generate accurate summaries is also significantly enhanced (Jiang et al., 2021). In such cases, as the number of reviews increases, AI-generated review summarization can help consumers obtain useful information from massive reviews quickly and alleviate decision fatigue caused by cognitive load, thereby enabling them to make purchasing decisions efficiently (Schermerhorn, 2023). Collectively, we propose hypothesis two:

H2: *The number of online reviews has a U-shaped moderating effect on the impact of AI-generated review summarization on consumer purchasing behavior.*

Ratings based on verified purchasers are informative of objective product values, reflecting consumers' average preferences for a product (Chen et al., 2021). Compared to products with lower average ratings, those with higher ratings typically indicate greater consumer satisfaction and more favorable features (Chen & Chang, 2018). AI-generated review summarization can effectively highlight the key strengths of highly-rated products and the positive opinions of previous consumers, which helps consumers reduce the burden of comprehension while enhancing their perceptions of the product's features (Xu & Zhang, 2018), thereby supporting more informed purchase decisions. Collectively, we propose hypothesis three:

H3: *Average rating positively moderates the effect of AI-generated review summarization on consumer purchasing behavior.*

Rating dispersion refers to the variability in rating feedback, which can reflect the uncertainty of product quality and the heterogeneity of product characteristics (Li, 2018). Compared to products with concentrated ratings, greater rating dispersion reflects that consumers have differing attitudes toward the product, then inconsistency among product reviews is enhanced, making it more challenging to extract the essence of the reviews (Wang & Cui, 2023). In such scenarios, AI-generated review summarization can efficiently extract the product's key features, clearly and objectively presenting its strengths and weaknesses to consumers. This enhances the perceived usefulness of the reviews, reduces their uncertainty (Lee et al., 2021), and facilitates purchase decisions. Collectively, we propose hypothesis four:

H4: *Rating dispersion positively moderates the effect of AI-generated review summarization on consumer purchasing behavior.*

The research model of this study is illustrated in Figure 1.

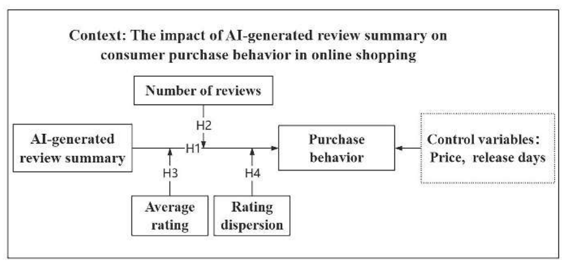

Figure 1. Research model.

4. Data and methods

This study focuses on Amazon (www.amazon.com), one of the largest e-commerce platforms in the world. We tracked and collected product data from February 2023 to February 2024 by web scraping. This dataset includes information on product prices, sales rank, and the presence of AI-generated review summaries, among other variables. We selected products for this study for the following criteria. First, considering the principle of comprehensiveness, this study includes a range of product categories such as beauty, household, electronics, and food, all of which have broad consumer appeal. Second, to achieve diversity, this study incorporates both experience and search products, as well as products of varying prices, to enhance the reliability of the empirical results (Guan et al., 2023; Mudambi & Schuff, 2010). Third, to maintain the robustness of the sales data, seasonal clothing and perishable vegetables, which are highly sensitive to seasonal fluctuations, were excluded. Additionally, following the principle of randomness, a random sample of products was selected from each category (Jabr & Rahman, 2022). After data cleaning, we constructed a panel dataset comprising total 48,019 records for 3,781 products. The product review page is shown in Figure 2. During the 13-month time window, the implementation of the AI-generated review summarization feature on product pages serves as a naturally occurring external policy intervention. This intervention is completely exogenous in the experimental setup. Additionally, the adoption of the AI-generated review summarization feature for products is entirely determined by the platform, leaving merchants no choice. Therefore, the issue of self-selection can be excluded. Based on the above discussion, this study designated August 2023 as the policy intervention point. Products were divided into treatment and control groups based on whether they were affected by the policy intervention (i.e., whether the AI-generated review summarization feature was adopted). A difference-in-differences model was then employed for the subsequent research.

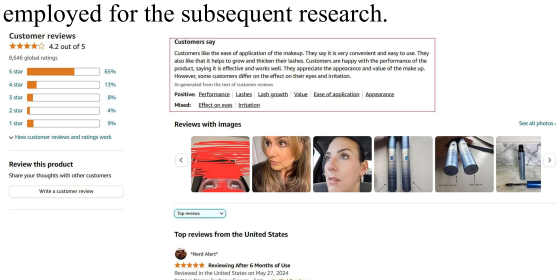

Figure 2. A screenshot of product review page.

In research related to e-commerce platforms, product sales are generally used to measure consumer purchasing behavior (Gao et al., 2024; Ma et al., 2022). However, Amazon does not disclose specific information about the sales of listed products. Given the challenges in acquiring sales data, many scholars have employed proxy and estimation methods to approximate product sales. Some studies have shown that there is a near-linear relationship between product sales rank and actual sales, making it reasonable to use rank as a proxy or to estimate sales (Brynjolfsson et al., 2003). Consequently, this study roughly reflects product sales by determining the correlation between each product's sales rank and its sales to approximate overall sales performance (Kaushik et al., 2018; Warut Khern-am-nuai et al., 2023). The treatment group dummy variable ($Treat_i$) is used to distinguish whether a product has adopted the AI-generated review summarization feature. The adoption time dummy variable ($Post_{it}$) is used to distinguish two periods, before and after the introduction of the AI-generated review summarization feature. Considering that certain product characteristics may also influence consumer purchase, this study includes product price ($price_{it}$) and the number of days since product release ($Release\_Days_{it}$) (Gao et al., 2024) as control variables. Additionally, to address the large differences in values between variables, and then to eliminate the impact of heteroscedasticity, this study takes the logarithm of continuous variables after adding 1 to their values. The definitions of the variables are provided in Table 1, and the descriptive statistics are presented in Table 2.

Table 1. Variable definition

Variables	Description
$Sales_{it}$	The sales of product *i* in month *t*.
$Treat_i$	A dummy variable; =1 if product *i* is in the treatment group, =0 if product *i* is in the control group.
$Post_{it}$	A dummy variable; =1 if product *i* in month *t* has implemented the AI-generated review summarization feature, otherwise, =0.
$Num\_Reviews_{it}$	Growth in the number of reviews for product *i* in month *t*.
Avg_Rating_{it}	Average rating value of product *i* in month *t*.
$Rating\_disp_{it}$	Standard deviation of ratings of product *i* in month *t*.
$Type_i$	Dummy variable; =1 for experience products, 0 for search products.
$Price_{it}$	Average price of product *i* in month *t*.
$Release\_Days_{it}$	Duration of product *i* from release to *t*.

Table 2. Descriptive statistics

Variables	Mean	Std. Dev.	Min	Max
$Sales$	10839.71	17299.88	0	203373
$Treat$	0.399	0.490	0	1
$Post$	0.534	0.499	0	1
$Num\_Reviews$	928.23	6387.005	0	304111
Avg_Rating	4.398	0.346	1	5
$Rating\_Disp$	0.239	0.436	0	3.536
$Type$	0.608	0.488	0	1
$Price$	20.602	41.749	0	5048.9
$Release\_Days$	2028.248	1647.481	0	9163

Next, following the approach of Guan et al. (2023), we constructed a difference-in-differences model to examine the direct effect of AI-generated review summarization on consumer purchasing behavior. The formula is specified as follows:

$$Sales_{it} = \beta_0 + \beta_1 Treat_i \times Post_{it} + \sum_{j=2}^{3} \beta_j Controls_{i,t} + D_i + D_t + \varepsilon_{it}$$

Where the subscripts *i* and *t* respectively represent the product and time. $Treat_i \times Post_{it}$ is the interaction term between the treatment group dummy variable and the adoption time dummy variable. β_1 is the coefficient of focus in this study, measuring the direct effect of AI-generated review summarization on consumer purchasing behavior. $Controls_{i,t}$ are control variables, D_i represents individual fixed effects, D_t represents time fixed effects, and ε_{it} is the random error term. Subsequently, to test the moderating effect of the number of product reviews on the relationship between AI-generated review summarization and consumer purchasing behavior, we followed the approach of Mudambi & Schuff (2010) by generating the interaction terms *Treat*Post*Num_reviews* and *Treat*Post*Num_reviews*$^{\wedge 2}$. Therefore, we constructed a difference-in-differences model incorporating moderating variable. The formula is specified as follows:

$$Sales_{it} = \beta_0 + \beta_1 Treat_i \times Post_{it} + \beta_2 Num\_Reviews_{it} + \beta_3 Treat_i \times Post_{it} \times Num\_Reviews_{it} + \beta_4 Treat_i \times Post_{it} \times Num\_Reviews_{it}^{\wedge 2} + \sum_{j=5}^{6} \beta_j Controls_{i,t} + D_i + D_t + \varepsilon_{it}$$

Where β_3 and β_4 are the coefficients of focus in this study, used to measure the moderating effect of the number of product reviews and to further examine the existence of a U-shaped relationship. In addition, the moderating effects of average rating and rating dispersion were also tested using the addition of the interaction terms *Treat*Post*Avg_Rating* and *Treat*Post*Rating_Disp* respectively to the DID model.

5. Results

5.1. Main result

Table 3 presents the main regression results of the DID model. Column (1) shows the direct regression of the interaction term *Treat*Post* with the dependent variable *Sales*, while column (2) includes control variables in the regression model based on column (1). The regression coefficients of *Treat*Post* are 0.137 and 0.145, respectively, and both are significant. It indicates that AI-generated review summarization has a significant positive effect on consumer purchasing behavior, confirming Hypothesis 1.

Table 3. Difference-in-Differences Model: The effect of AI-generated review summarization

	(1)	(2)
Variables	***Sales***	***Sales***
*Treat*Post*	0.1370***	0.1450***
	(4.1100)	(4.4180)
Price		-0.0924***
		(-3.9280)
Release_Days		0.6860***
		(13.1500)
Observations	48,019	48,019
ID FE	YES	YES
Month FE	YES	YES
Adjusted R^2	0.0097	0.0234

Note: T-statistics values in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

5.2. Moderating effect

Table 4 presents the regression results for the moderating effect of the number of product reviews. Column (1) includes the interaction terms *Treat*Post* Num_reviews* and *Treat*Post*Num_reviews*^{2} in the regression model. As shown in Table 4, the coefficient for the interaction term *Treat*Post*Num_reviews* is -0.0842, and for the *Treat*Post*Num_reviews*$^{\wedge 2}$ is 0.0071. Both coefficients are significant, indicating that the number of product reviews may exhibit a U-shaped moderating effect. Next, we conducted a U-test to verify the existence of the U-shaped. The results indicate that the extreme point is 5.97, which lies within the interval [0, 11.741]. Additionally, the p-value is 0.0069 < 0.01, indicating that the U-shaped moderating effect is significant. In summary, the number of product reviews exhibits a U-shaped moderating effect on the impact of AI-generated review summarization on consumer purchasing behavior, indicating that Hypothesis 2 is supported.

Table 4. The moderating effect of the number of product reviews

	(1)
Variables	***Sales***
*Treat*Post*Num_reviews*	-0.0842***
	(-4.8320)
*Treat*Post*Num_reviews*$^{\wedge 2}$	0.0071***
	(3.4180)
*Treat*Post*	0.3190***
	(7.2510)
Controls	YES
ID FE	YES
Month FE	YES
Adjusted R^2	0.0266

Note: T-statistics values in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

Next, the moderating effects of average product rating and rating dispersion were examined, with the results shown in Table 5. From column (1), the coefficient of the interaction term *Treat*Post*Avg_Rating* is 0.102, while in column (2), the coefficient of the interaction term *Treat*Post*Rating_Disp* is 0.188. Both coefficients are statistically significant, indicating that both average rating and rating dispersion have a positive moderating effect, thereby providing support for Hypothesis 3 and 4.

Table 5. The moderating effect of average rating and rating dispersion

	(1)	(2)
Variables	***Sales***	***Sales***
*Treat*Post*	0.1520***	0.1230***
	(4.5760)	(3.7360)
*Treat*Post*Avg_Rating*	0.1020*	
	(1.9390)	
*Treat*Post*Rating_Disp*		0.1880***
		(3.5430)
Controls	YES	YES
ID FE	YES	YES
Month FE	YES	YES
Adjusted R^2	0.0205	0.0160

Note: T-statistics values in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

5.3. Heterogeneity analysis

To further examine the impact of AI-generated review summarization on different types of products, we categorized products into search products and experience products (Nelson, 1970). The results of the subgroup regression analysis are presented in Table 6. From columns (1) and (2), it shows that the coefficients of the interaction term *Treat×Post* for search products and experience products are 0.1670 and 0.0954, respectively (0.1670 > 0.0954). This suggests that the positive effect of AI-generated review summarization on consumer purchasing behavior may be stronger for search products. To further validate the reliability of the conclusions, we conducted a suest test for group differences. The results indicate that the chi-squared test p-value is 0.0582 < 0.1, confirming a significant difference between search products and experience products.

For search goods, users are more concerned with standardized product features. AI-generated review summarization can provide consumers with semi-structured product attribute information, reducing the complexity of information processing and enhancing decision-making efficiency. In contrast, users purchasing experience goods are more concerned with personal experiences and differences. AI-generated review summarization can only provide general reference information, having a relatively weaker impact on consumer purchasing behavior (Guha Majumder et al., 2022; Román et al., 2023).

Table 6. The heterogeneity effect of products

	(1) Search	(2) Experience
Variables	***Sales***	***Sales***
*Treat*Post*	0.1670***	0.0954***
	(5.1190)	(4.4110)
Controls	YES	YES
ID FE	YES	YES
Month FE	YES	YES
Observations	18,815	29,204
Adjusted R^2	0.0083	0.0166

Note: T-statistics values in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

5.4. Robustness tests

Parallel trend test. Before the shock of the AI-generated review summarization, it is an important premise of the difference-in-differences method that the treatment and the control groups meet the parallel trend assumption. We designated August 2023 as the point when the products were shocked, then conducted a parallel trend test with product sales as the dependent variable. The results are shown in Figure 3. It can be observed that before the AI-generated review summarization shock, there was no significant difference in product sales between the treatment group and the control group, satisfying the premise of the parallel trend assumption. After the shock, as products gradually adopted the AI-generated review summarization feature, a significant positive difference emerged between two groups.

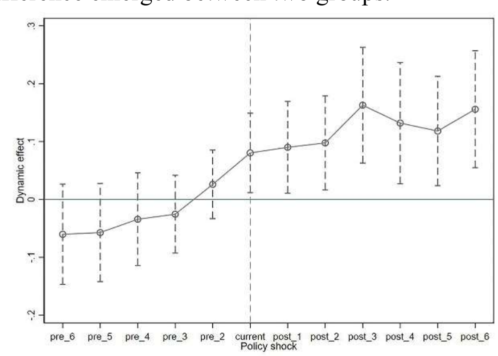

Figure 3. Parallel trend test.

Placebo test. To address the possibility that the estimation results may be influenced by random disturbances, we plotted a kernel density graph for an individual placebo test. Repeatedly perform 500 random assignments of the treatment effect and test the distribution of the estimated coefficients of the policy variable. The results are shown in Figure 4. It can be observed that the estimated coefficients of the randomly assigned policy variable (*sales*) are centered around 0, and most of the p-values of these estimated coefficients are above 0.1, and the estimator values are far from the benchmark regression estimated coefficient value of 0.145, indicating that the DID estimation results are relatively robust.

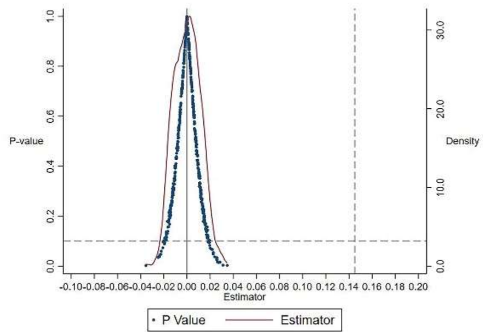

Figure 4. Placebo test.

Additional robustness checks. To avoid sample selection bias within groups, we used one-to-one and one-to-three nearest neighbor matching methods within a caliper to match the treatment and control groups. Columns (1) and (2) of Table 7 show the results for the matched samples. It can be observed that the direction and significance of all coefficients are highly consistent with those in Table 3. It is worth noting that consumers will only make a purchase after entering the page. Since the action of consumers clicking on the product page occurs before they see the AI-generated review summaries, it can be said that the clicking behavior is entirely unaffected by the AI-generated review summarization. Therefore, we replaced the explanatory variable with product clicks and conducted the DID regression again. The results are shown in column (3). It can be seen that the coefficient of the interaction term *Treat×Post* is no longer significant, indicating that no other policy effects were found during the time window period. Furthermore, to avoid the biasing effects of extreme values, we performed two-way winsorization on continuous variables at the 1st and 99th percentiles, the regression results are shown in column (4). The research conclusions are consistent with the previous findings, further indicating that our research results are robust.

Table 7. Other robustness test results

Variables	(1) *Sales*	(2) *Sales*	(3) *Clicks*	(4) *Sales*
*Treat*Post*	0.125***	0.120***	-0.0106	0.144***
	(2.492)	(3.221)	(-0.372)	(4.387)
Controls	YES	YES	YES	YES
ID FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Observations	16,957	29,121	48,019	48,019
Adjusted R^2	0.0257	0.0218	0.1110	0.0218

Note: T-statistics values in parentheses;
*** p<0.01, ** p<0.05, * p<0.1

6. Conclusion and implications

Amazon introduced AI-generated review summarization, which aims to alleviate consumers' information overload when purchasing products, thereby shaping an innovative AI-human interaction ecosystem in the online shopping scenarios. Based on Cognitive Load Theory, this paper proposes the pathways of AI-generated review summarization influencing consumer purchasing behavior. It highlights that the impact is moderated by the number of online reviews, average rating, and rating dispersion. Then, an empirical test using the difference-in-differences method was conducted. The research found that AI-generated review summarization technology has a positive effect on product sales, with a stronger positive effect on search products compared to experience products. Additionally, the number of online product reviews has a U-shaped moderating effect, while average rating and rating dispersion both have positive moderating effects.

This study has several managerial implications. First, this study reveals the positive impact of AI-generated review summarization technology on consumer purchasing behavior, providing valuable insights for other e-commerce platforms. E-commerce platforms should actively explore and implement this technology to enhance consumer engagement, improve the purchasing experience, and increase conversion rates, thereby boosting market competitiveness. Second, this study analyzes the moderating effects of the number of product reviews, average rating and rating dispersion, as well as the heterogeneous impact on different type of products. This analysis contributes to understanding the optimization issues associated with implementing AI-generated review summarization technology. When introducing AI-generated review summarization technology, e-commerce platforms should adopt differentiated management measures based on the types and characteristics of products. For example, for search products with a large number of reviews, it is essential to continuously enhance the accuracy of AI-generated summaries.

This study has certain limitations, providing possible works for further exploration. First, the products selected in this study were random. whether the research conclusions are applicable to other products or e-commerce platforms remains to be explored. Due to the limitations of the available data, we were unable to obtain the specific values for each individual rating, and it is currently only possible to roughly calculate the month-to-month rating dispersion based on the average rating for the month.

Future research could expand the dataset or conduct cross-platform studies to test the generalizability of these conclusions. Second, this study has not yet empirically tested the inrernal path of the effects. Future research could use surveys or scenario experiments to measure consumers' internal motivations. Finally, our work has not considered individual differences. The reactions and preferences of different consumer groups (such as age, gender, purchasing habits, etc.) toward AI technology still need to be verified.

7. Acknowledgements

This work was jointly supported by National Key Research and Development Program of China (2021YFF0900800) and the National Natural Science Foundation of China (72072194).

8. References

Balan U, M., & Mathew, S. K. (2020). Personalize, Summarize or Let them Read? A Study on Online Word of Mouth Strategies and Consumer Decision Process. *Information Systems Frontiers*, *23*, 627–647. https://doi.org/10.1007/s10796-020-09980-9

Beck, B. B., Wuyts, S., & Jap, S. D. (2023). Guardians of Trust: How Review Platforms Can Fight Fakery and Build Consumer Trust. *Journal of Marketing Research*. https://doi.org/10.1177/00222437231195576

Brynjolfsson, E., Hu, Y. (Jeffrey), & Smith, M. D. (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science, 49*(11), 1580–1596. https://doi.org/10.1287/mnsc.49.11.1580.20580

Chen, C.-C., & Chang, Y.-C. (2018). What drives purchase intention on Airbnb? Perspectives of consumer reviews, information quality, and media richness. *Telematics and Informatics, 35*(5), 1512–1523. https://doi.org/10.1016/j.tele.2018.03.019

Chen, P., Hitt, L. M., Hong, Y., & Wu, S. (2021). Measuring Product Type and Purchase Uncertainty with Online Product Ratings: A Theoretical Model and Empirical Application. *Information Systems Research, 32*(4), 1470–1489. https://doi.org/10.1287/isre.2021.1041

Cheung, C. M. K., Xiao, B. S., & Liu, I. L. B. (2014). Do actions speak louder than voices? The signaling role of social information cues in influencing consumer purchase decisions. *Decision Support Systems*, *65*(SI), 50–58. https://doi.org/10.1016/j.dss.2014.05.002

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261–1270. https://doi.org/10.1016/j.jbusres.2014.11.006

Gao, M., Wang, J., & Liu, O. (2024). Is UGC sentiment helpful for recommendation? An application of sentiment-based recommendation model. *Industrial Management and Data Systems*. https://doi.org/10.1108/imds-05-2023-0335

Goodman, K. E., Yi, P. H., & Morgan, D. J. (2024). AI-Generated Clinical Summaries Require More Than Accuracy. *JAMA*, *331*(8), 637–638. https://doi.org/10.1001/jama.2024.0555

Guan, Y., Tan, Y., Wei, Q., & Chen, G. (2023). When Images Backfire: The Effect of Customer-Generated Images on Product Rating Dynamics. *Information Systems Research*, *34*(4), 1641–1663. https://doi.org/10.1287/isre.2023.1201

Guha Majumder, M., Dutta Gupta, S., & Paul, J. (2022). Perceived usefulness of online customer reviews: A review mining approach using machine learning & exploratory data analysis. *Journal of Business Research*, *150*, 147–164. https://doi.org/10.1016/j.jbusres.2022.06.012

Hu, H., & Krishen, A. S. (2019). When is enough, enough? Investigating product reviews and information overload from a consumer empowerment perspective. *Journal of Business Research*, *100*, 27–37. https://doi.org/10.1016/j.jbusres.2019.03.011

Jabr, W., & Rahman, M. (2022). Online Reviews and Information Overload: The Role of Selective, Parsimonious, and Concordant Top Reviews. *MIS Quarterly*, *45*(3), 1517–1550. https://doi.org/10.25300/misq/2022/16169

Jiang, W., Chen, J., Ding, X., Wu, J., He, J., & Wang, G. (2021). Review Summary Generation in Online Systems: Frameworks for Supervised and Unsupervised Scenarios. *ACM Transactions on the Web*, *15*(3), 1–33. https://doi.org/10.1145/3448015

Kaushik, K., Mishra, R., Rana, N. P., & Dwivedi, Y. K. (2018). Exploring reviews and review sequences on e-commerce platform: A study of helpful reviews on Amazon.in. *Journal of Retailing and Consumer Services*, *45*(0969-6989), 21–32. https://doi.org/10.1016/j.jretconser.2018.08.002

Lee, S., Lee, S., & Baek, H. (2021). Does the dispersion of online review ratings affect review helpfulness? *Computers in Human Behavior, 117,* 106670. https://doi.org/10.1016/j.chb.2020.106670

Li, X. (2018). Impact of Average Rating on Social Media Endorsement: The Moderating Role of Rating Dispersion and Discount Threshold. *Information Systems Research,* *29*(3), 739–754. https://doi.org/10.1287/isre.2017.0728

Liang, C., Wu, J., & Li, X. (2023). Do “Likes” in a Brand Community Always Make You Buy More?. *Information Systems Research*. https://doi.org/10.1287/isre.2022.0008

Liu, S., McCoy, A. B., Wright, A. P., Nelson, S. D., Huang, S. S., Ahmad, H. B., Carro, S. E., Franklin, J., Brogan, J., & Wright, A. (2024). Why do users override alerts? Utilizing large language model to summarize comments and optimize clinical decision support. *Journal of the American Medical Informatics Association*, *31*(6), 1388–1396. https://doi.org/10.1093/jamia/ocae041

Lu, B., Ma, B., Cheng, D., & Yang, J. (2023). An Investigation on Impact of Online Review Keywords on Consumers' Product Consideration of Clothing. *Journal of Theoretical and Applied Electronic Commerce Research*, *18*(1), 187–205. https://doi.org/10.3390/jtaer18010011

Luo, L., Duan, S., Shang, S., & Pan, Y. (2021). What makes a helpful online review? Empirical evidence on the effects of review and reviewer characteristics. *Online Information Review*, *45*(3), 614–632. https://doi.org/10.1108/oir-05-2020-0186

Ma, G., Ma, J., Li, H., Wang, Y., Wang, Z., & Zhang, B. (2022). Customer behavior in purchasing energy-saving products: Big data analytics from online reviews of e-commerce. *Energy Policy*, *165*, 112960. https://doi.org/10.1016/j.enpol.2022.112960

Mohammad Rajiur Rahman, Raga Shalini Koka, Shah, S. K., Solorio, T., & Jaspal Subhlok. (2023). Enhancing lecture video navigation with AI generated summaries. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-11866-7

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. https://doi.org/10.2307/20721420

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*, *78*(2), 311–329. https://doi.org/10.1086/259630

Park, K., Lee, S., Shahryar Doosti, & Tan, Y. (2023). Provision of helpful review videos: Effects of video characteristics on perceived helpfulness. *Production and Operations Management*, *32*(7), 2031–2048. https://doi.org/10.1111/poms.13969

Peal, M., Hossain, M. S., & Chen, J. (2022). Summarizing consumer reviews. *Journal of Intelligent Information Systems*, *59*, 193–212. https://doi.org/10.1007/s10844-022-00694-9

Peter, S., & Riemer, K. (2024). Creative Assistants with Style: Making Sense of Generative AI as "Style Engines." *Proceedings of the 57th Hawaii International Conference on System Sciences*, 3980–3989. https://hdl.handle.net/10125/106866

Racat, M., & Plotkina, D. (2023). Sensory-enabling Technology in M-commerce: The Effect of Haptic Stimulation on Consumer Purchasing Behavior. *International Journal of Electronic Commerce*, *27*(3), 354–384. https://doi.org/10.1080/10864415.2023.2226900

Román, S., Riquelme, I. P., & Iacobucci, D. (2023). Fake or credible? Antecedents and consequences of perceived credibility in exaggerated online reviews. *Journal of Business Research*, *156*, 113466. https://doi.org/10.1016/j.jbusres.2022.113466

Schermerhorn, V. (2023, August 14). *How Amazon continues to improve the customer reviews experience with generative AI*. US about Amazon. https://www.aboutamazon.com/news/amazon-ai/amazon-improves-customer-reviews-with-generative-ai

Shobhit Kakaria, Farzad Saffari, Ramsøy, T. Z., & Enriqué Bigné. (2023). Cognitive load during planned and unplanned virtual shopping: Evidence from a neurophysiological perspective. *International Journal of Information Management*, *72*, 102667. https://doi.org/10.1016/j.ijinfomgt.2023.102667

Stolz, K., Hammerschmidt, T., & Posegga, O. (2024). How Conversational Agents Influence Purchase Decisions of Online Fashion Shoppers toward Sustainable Consumption: Exploring Nudges for Green Decision-Making. *Proceedings of the 57th Hawaii International Conference on System Sciences,* 4889–4898. https://hdl.handle.net/10125/106970

Sweller, J. (1988). Cognitive Load during Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Wang, M., & Cui, X. (2023). The role of rating feedback and its implications for solver submission behavior in crowdsourcing contests. *Information & Management, 60*(5), 103790–103790. https://doi.org/10.1016/j.im.2023.103790

Warut Khern-am-nuai, Hossein Ghasemkhani, Qiao, D., & Kannan, K. (2023). The Impact of Online Q&As on Product Sales: The Case of Amazon Answer. *Information Systems Research*. https://doi.org/10.1287/isre.2023.1233

Wells, J. D., Valacich, J. S., & Hess, T. J. (2011). What Signal Are You Sending? How Website Quality Influences Perceptions of Product Quality and Purchase Intentions. *MIS Quarterly*, *35*(2), 373. https://doi.org/10.2307/23044048

Xu, C., & Zhang, Q. (2018). The dominant factor of social tags for users' decision behavior on e-commerce websites: Color or text. *Journal of the Association for Information Science and Technology*, *70*(9), 942–953. https://doi.org/10.1002/asi.24118

Yin, D., de Vreede, T., M. Steele, L., & de Vreede, G.-J. (2022). Decide Now or Later: Making Sense of Incoherence Across Online Reviews. *Information Systems Research*, *34*(3), 1211–1227. https://doi.org/10.1287/isre.2022.1150

Zhai, M., Wang, X., & Zhao, X. (2024). The importance of online customer reviews characteristics on remanufactured product sales: Evidence from the mobile phone market on Amazon.com. *Journal of Retailing and Consumer Services*, *77*, 103677. https://doi.org/10.1016/j.jretconser.2023.103677