Does AI Disclosure in Discriminatory Pricing Backfire? The Moderating Role of Price Sensitivity and Explanation for Price Differences

Xiao Peng
Zhejiang University, City University of Hong Kong
pengxiao@zju.edu.cn

Xixian Peng
Zhejiang University
pengxx@zju.edu.cn

David (Jingjun) Xu
City University of Hong Kong
davidxu@cityu.edu.hk

Abstract

This research aimed to examine the moderating role of price sensitivity and explanation for price differences in the relationship between AI disclosure and consumers' revenge behavior, as well as to explore the potential mediating effect of inferred motives. A scenario-based lab experiment was conducted, involving 121 participants who engaged in an online airline ticket booking context. The findings of this study revealed that the positive impact of AI disclosure on revenge behavior was amplified among consumers with high price sensitivity, and this relationship was mediated by inferred motives. Additionally, the provision of explanations alongside AI disclosure was found to increase revenge behavior. These findings contribute to the understanding of consumers' psychological processes and revenge behavior within the context of discriminatory pricing empowered by AI. Moreover, the study offers practical implications for managers aiming to mitigate the negative consequences of discriminatory pricing.

Keywords: Discriminatory pricing, AI disclosure, Price sensitivity, Explanation, Inferred motives

1. Introduction

Discriminatory pricing, also referred to as price discrimination, involves charging different prices for the same product or service to different customer segments or individuals (Hau et al., 2021). With the advancements in machine learning algorithms and AI technology, businesses have increasingly adopted discriminatory pricing strategies. Automated algorithms that implement discriminatory pricing may disproportionately affect minority groups due to biases in training data, proxy variables linked to socioeconomic status, feedback loops that reinforce historical discrimination, algorithmic opacity, and the perpetuation of systemic inequalities (Seele et al., 2021). Instances of this pricing approach have yielded unfavorable outcomes; for example, customers have complained about elevated prices based on factors such as low battery levels or premium phone models (Chowdhry, 2016). The price differences in discriminatory pricing, especially higher prices than others arouse consumers' perception of unfairness in pricing, leading to diminished trust in companies and hesitancy in making purchases among consumers (Borgesius & Poort, 2017). Furthermore, such pricing practices have been found to evoke negative inferred motives of companies among consumers, resulting in revenge behavior, including actions like complaints, negative word-of-mouth, reducing purchases in companies, and switching to alternative competitors (Chung & Petrick, 2013). Given the significant negative impact of consumers' revenge behavior, businesses have recognized the need to focus on effectively managing discriminatory pricing (Allender et al., 2021).

The widespread adoption of AI in discriminatory pricing practices by companies (Garvey et al., 2023) has prompted a growing need for regulations that mandate the disclosure of AI usage. This arises from consumers' desire to be informed about the use of algorithms or AI systems monitoring their activities (MacCarthy, 2020). Although prior research has highlighted consumers' negative reactions and aversion toward AI or algorithms (Dietvorst et al., 2018), the specific impact of AI disclosure in the context of discriminatory pricing is still not well understood. Therefore, this study aimed to contribute to the existing literature by investigating the effects of AI disclosure on revenge behavior. Moreover, inferred motives, refer to the consumer's perception of the underlying intentions or motives of a company or brand based on the information disclosed (Joireman et al., 2013) (e.g., positive motivation refers to the company helping customers, while negative motivation refers to the company maximizing its own interests and utilizing customers), which is highly related to revenge behavior. Therefore, we also explored the mediation effect of inferred motives on the relationship between AI disclosure and revenge behavior.

URI: https://hdl.handle.net/10125/107208
978-0-9981331-7-1

While previous research has primarily focused on situational factors influencing the impact of AI disclosure such as utilitarian vs. hedonic contexts (Longoni & Cian, 2022), it is important to acknowledge that consumer characteristics themselves can also play a role in determining the effectiveness of AI disclosure. In this regard, we propose that price sensitivity, which refers to the extent to which consumers are responsive to price fluctuations and value considerations when making purchasing decisions, is a crucial factor to consider (Goldsmith & Newell, 1997). Previous study indicated that consumers with high price sensitivity are more likely to exhibit resistance toward discriminatory pricing (Weisstein et al., 2013). However, the role of price sensitivity in amplifying the effects of AI disclosure on consumers' inferred motives of company and revenge behavior has not been investigated. Therefore, our study aimed to address this research gap by examining the moderating effect of price sensitivity on the relationship between AI disclosure and revenge behavior in the context of discriminatory pricing.

Furthermore, businesses utilize the tactic of offering explanations for price differences as a means to alleviate the adverse repercussions associated with discriminatory pricing practices (Li & Jain, 2016). Extant literature has consistently underscored the efficacy of providing explanations across diverse settings (Mao & Benbasat, 2000). Nevertheless, we are interested in whether consumers perceive the integration of AI disclosure and explanations for price differences as a facade or pretext for enacting discriminatory pricing by a company, leading to inferred negative motives of the company and consumers' revenge behavior. Consequently, this study endeavored to bridge this research gap by examining the interaction effect of AI disclosure and explanation on revenge behavior.

The primary objective of this study was to investigate the moderating role of price sensitivity and explanation in the relationship between AI disclosure and consumers' revenge behavior, as well as to explore the potential mediating effect of inferred motives. To test our hypotheses, we conducted a scenario-based lab experiment involving 121 participants in the context of online airline ticket booking, a prevalent domain where discriminatory pricing is frequently employed (Shukla et al., 2019). This research contributes to the understanding of the psychological processes and revenge behavior exhibited by consumers in the context of discriminatory pricing empowered by AI and provides practical implications for managers seeking to address the negative consequences of discriminatory pricing.

The structure of this paper is as follows: Firstly, we present a review of relevant literature pertaining to consumer responses to discriminatory pricing and the impact of AI disclosure. Secondly, we develop our theoretical framework and formulate hypotheses. Thirdly, we describe our research methodology, encompassing details of the experimental design, measures, and procedures employed. Fourthly, we present our findings and conduct a comprehensive analysis to evaluate the hypotheses. Lastly, we discuss the implications of our findings, identify limitations, and suggest potential avenues for future research.

2. Literature Review

2.1. Discriminatory Pricing

Discriminatory pricing is seen as a mechanism through which businesses exert power over consumers without their consent. This ability is reflected in the company's access to vast amounts of data, the ability to track consumers, and the computational power that enables real-time price adjustments (Joireman et al., 2013). Consumers' perception of an ethical failure, particularly in relation to discriminatory pricing, can significantly impact their relationship with the firm and lead to negative and even punitive reactions (Wang & Krishna, 2012). A perception of discrimination pricing as unfair, inequitable, and non-transparent can damage the trust relationship (Garbarino & Maxwell, 2010). Consumers' responses can range from a simple change in attitude, increased comparative search behavior, negative electronic word-of-mouth (Shea, 2010), complaints (Chung & Petrick, 2013), no-repurchase, non-purchase to resistance, and revenge (Brunk, 2010).

In the age of big data, advanced AI technologies, and the abundance of extensive data facilitated the practice of discriminatory pricing. Simultaneously, social networks may amplify the manifestation of companies' hegemonic influence in discriminatory pricing and serve as a fertile ground for negative word-of-mouth and more aggressive and threatening resistance behaviors towards retailers. Among these reactions, revenge is the most detrimental consumer response for companies. Bechwati and Morrin (2003) defined revenge desire as the consumer's retaliatory feeling towards a company, such as the desire to inflict harm upon the company, typically following an extremely negative purchasing experience. Revenge behavior includes complaining, spreading negative word-of-mouth, reducing purchases, and switching to other competitors (Chung & Petrick, 2013), which can have a significant adverse impact on companies. However, little research has focused on the factors influencing consumer revenge behavior in discriminatory pricing. Further investigation is needed to determine what factors influence consumer revenge behavior and the underlying mechanisms.

2.2. AI Disclosure

Although companies have embraced the adoption of AI in their business practices, leveraging its advanced capabilities and efficiency in collecting and processing information (Kumar et al., 2016), previous research has highlighted consumers' negative reactions towards AI or algorithm aversion (Dietvorst et al., 2018). These negative reactions stem from perceived factors such as a lack of knowledge and empathy (Luo et al., 2019), a perceived absence of benevolent intention (Garvey et al., 2023), a failure to consider uniqueness (Yalcin et al., 2022), or a perceived deficiency in mental and emotional attributes (Longoni et al., 2019).

Various studies have demonstrated the negative effects of AI disclosure in different contexts. For instance, in the context of structured outbound sales calls, AI disclosure was found to significantly reduce customers' purchase rates, leading to a substantial decrease of 79.7% (Luo et al., 2019). Similarly, in situations where the offered price for a second-hand performance ticket or ride service was lower than expected, AI disclosure was found to have a detrimental effect on consumers’ offer responses (Garvey et al., 2023). Furthermore, recent research has revealed that consumers tend to exhibit less positive reaction when they become aware that a favorable decision, such as the acceptance of an application, was made by an AI (Yalcin et al., 2022).

However, the influence of AI disclosure on consumers' revenge behavior in the context of discriminatory pricing remains largely unexplored. There is a need for empirical investigation to validate and further explore these relationships, contributing to a deeper understanding of consumer responses to AI disclosure and the underlying psychological mechanisms at play.

Furthermore, there are several potential moderators that may influence the impact of AI disclosure. While previous studies have explored the moderating role of factors such as utilitarian vs. hedonic contexts (Longoni & Cian, 2022), outcome favorability (Yalcin et al., 2022), anthropomorphic degree (Crolic et al., 2022), construct level (Kim & Duhachek, 2020), there remains a gap in understanding the moderating effects of price sensitivity (Hufnagel et al., 2022) and explanation (Li & Jain, 2016) in the context of discriminatory pricing. Investigating how these critical factors interact with AI disclosure to shape consumer behavior in the context of discriminatory pricing would be highly valuable.

Although existing research has made significant contributions, our study stands out in several important ways. Firstly, it adds to the consumer behavior literature by focusing specifically on consumers' revenge behavior in the context of discriminatory pricing. Secondly, it advances our understanding of consumer reactions to AI by examining the effects of AI disclosure and the underlying mechanisms involved in discriminatory pricing. Lastly, it contributes to the literature on discriminatory pricing by exploring the interactions between price sensitivity and AI disclosure, as well as between explanation and AI disclosure.

3. Hypotheses Development

Discriminatory pricing raises ethical concerns, as it can exploit vulnerable consumers or perpetuate social inequalities (Seele et al., 2021). AI disclosure reveals and emphasizes the use of discriminatory pricing, consumers may perceive the business as engaging in unethical behavior. In addition, AI systems are often viewed as black boxes, where their decision-making processes are not fully understood by the general public (Rai, 2020). When AI is responsible for discriminatory pricing, individuals, particularly those who face higher prices than others, may perceive a lack of accountability for the decision. This perceived lack of accountability can fuel feelings of helplessness as individuals seek to exert control or seek justice through their actions. Even worse, consumers may have increased desire for revenge, defined as retaliatory actions taken by consumers in response to perceived mistreatment or dissatisfaction (Ayadi et al., 2017).

Moreover, Consumers form judgments and make inferences about a company's motives based on the price outcome and other information including AI disclosure. AI disclosure reveals and emphasizes that the business is engaging in unethical behavior of discriminatory pricing. This perception further amplifies the negative inference about the business's motives, as consumers believe that the company is intentionally deceiving them for financial gain, taking advantage of consumer, or pursuing profit maximization. These inferred negative motives intensify revenge behavior as consumers seek retribution for what they perceive as deliberate mistreatment. Therefore, we put forward the following hypothesis:

H1: *Inferred negative motives mediate the positive impact of AI disclosure on revenge behavior.*

Previous research has highlighted the importance of price sensitivity in managing consumer behavior of discriminatory pricing(Goldsmith et al., 2010). Building upon this literature, we propose that the impact of AI disclosure on revenge behavior, may be moderated by individual differences in price sensitivity. Price sensitivity refers to the degree to which consumers are responsive to price changes and value considerations when making purchasing decisions.

Highly price-sensitive consumers are likely to be more vigilant in evaluating product or service attributes, including AI disclosure information. They are more attuned to the potential benefits or drawbacks associated with AI utilization and are more prone to react strongly when they perceive any unfair or misleading practices. Therefore, we put forward the following hypothesis:

H2: *Price sensitivity strengthens the positive impact of AI disclosure on revenge behavior.*

In addition to the moderating effect of price sensitivity, we assumed inferred motives as a potential mediator between AI disclosure and revenge behavior. Consumers with high price sensitivity are more likely to scrutinize AI disclosure information and attribute motives based on their price-conscious decision-making tendencies. Due to their heightened scrutiny, consumers with high price sensitivity pay closer attention to AI disclosure and are more likely to engage in detailed information processing. As a result, they may be more sensitive to any perceived discrepancies or inadequacies in the disclosure, leading to stronger inferences about the business's negative motives and increased revenge behavior. Therefore, we put forward the following hypothesis:

H3: *Price sensitivity increases revenge behavior by strengthening the positive impact of AI disclosure on inferred negative motives.*

Effective communication is crucial in managing consumer perceptions and expectations regarding discriminatory pricing. Providing explanations to consumers about why prices vary over time or why prices differ from those of others may result in favorable consumer perceptions. However, it is possible that providing explanations for price differences alongside AI disclosure may inadvertently contribute to negative perceptions and subsequent revenge behavior. The underlying reason lies in that transparency may highlight biases. AI systems, while capable of complex decision-making, can still reflect and perpetuate societal biases and prejudices present in the data they were trained on. When AI is used for discriminatory pricing, explanations for discriminatory pricing may inadvertently reveal the biases within the system and perceived as confusing or misleading. The transparency provided by these explanations could anger individuals who feel personally affected by the biases. Consequently, consumers may be more inclined to engage in revenge behavior. Therefore, we put forward the following hypothesis:

H4: *Explanation for price differences strengthens the positive impact of AI disclosure on revenge behavior.*

In addition to the moderating effect of explanation for price differences, H5 introduces the concept of inferred negative motives as a mediator between AI disclosure, explanations, and revenge behavior. When consumers receive explanations for discriminatory pricing with AI disclosure, consumer may perceive this as companies' defense of discriminatory behavior, they may develop negative attributions about the company's motives. Explanations with AI disclosure may fuel suspicion, raise doubts about the company's intentions, and contribute to perceptions of mistreatment or unfair practices. Consumers may interpret explanation for price difference with AI disclosure as an attempt to hide unethical practices, prioritize profit over customer well-being, or engage in deceptive behaviors. Consequently, these inferred negative motives can amplify revenge behavior as consumers seek retribution for perceived mistreatment. Therefore, we put forward the following hypothesis:

H5: *Explanation for price differences increases revenge behavior by strengthening the positive impact of AI disclosure on inferred negative motives.*

Figure 1 shows our research model.

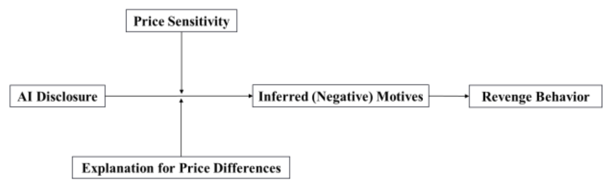


Figure 1. Research framework.

4. Methodology

To examine the proposed hypotheses, a scenario-based experiment was conducted utilizing a 2 (AI disclosure: with vs. without) × 2 (Explanation: with vs. without) between-subjects design. Consumer price sensitivity was a measured variable which was used to classify participants into high versus low price sensitivity group. Participants were presented with a scenario and instructed to imagine themselves within the described context. Subsequently, questionnaire data were collected to assess participants' perceptions and reactions toward discriminatory pricing.

4.1. Participants and Procedures

The study involved students from a Chinese public university, recruited through an online forum. Those who passed a check to ensure their attention were included in the analysis. Each participant received 10 RMB after completing the study. A total of 121 participants (31% male) met these criteria and completed the survey. While most participants were students around 23 years old, 75% had bought air tickets before, and 97% spent over 1000 RMB per month. This made them suitable for our study about buying air tickets. Participants were randomly placed into different groups, and there were no significant demographic differences among these groups.

For our research, we designed stimuli materials based on booking airline tickets. These materials included real airline information and reasonable prices. In all groups, we presented a scenario where ticket prices increased after a week. This reflected real situations where airfares often go up over time. To manipulate AI disclosure, we added a note "The following price is set by AI". To manipulate the explanation, only participants in the condition of with explanation will see a note “Due to factors such as time, passenger traffic, and costs, air ticket prices may vary. Different users may receive varying levels of discounts, resulting in price differences.” in each booking interface. Figure 2 shows these materials for the different conditions, along with translation notes of the disclosure and explanation in Chinese.

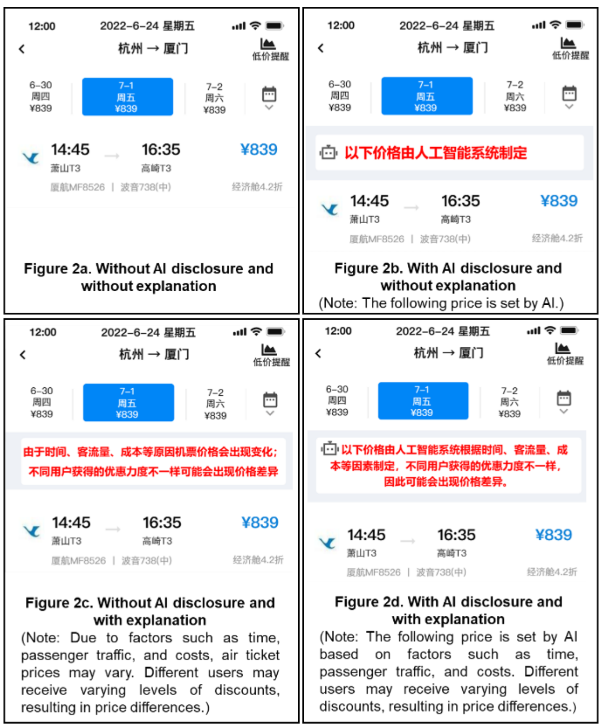


Figure 2a. Without AI disclosure and without explanation

Figure 2b. With AI disclosure and without explanation
(Note: The following price is set by AI.)

Figure 2c. Without AI disclosure and with explanation
(Note: Due to factors such as time, passenger traffic, and costs, air ticket prices may vary. Different users may receive varying levels of discounts, resulting in price differences.)

Figure 2d. With AI disclosure and with explanation
(Note: The following price is set by AI based on factors such as time, passenger traffic, and costs. Different users may receive varying levels of discounts, resulting in price differences.)

Figure 2. Stimuli materials.

At the start, we assigned participants randomly to one of four conditions. We asked them to imagine wanting to buy a 639 RMB air ticket for a trip. A week later, they found out their ticket price had increased by 200 RMB to 839 RMB. Meanwhile, their friend's ticket cost 739 RMB, which was 100 RMB less than theirs. Before they answered the survey questions, we asked about the price changes to ensure they read the materials attentively. The survey's first part aimed to understand their reactions in this situation, while the second part collected personal information like age, gender, job, monthly spending, experience buying air tickets, and price sensitivity.

4.2. Measures

The measurement items for each variable were adapted from previous research and were somewhat modified to best fit the study context. Inferred motives on a two-item scale was adapted from Campbell (2007), including the motive of this airline for the price difference is good/bad for you (1 = “good,” and 9 = “bad”) and their agreement with the statement, “The intention of this airline’s ticket price increase was to take advantage of you (the customer)”. Three items were adapted from Chung & Petrick (2013) to measure revenge behavior. For example, “I will switch to other competitors because of the price changes on the most recent trip with the airline.” (1 = “strongly disagree,” and 9 = “strongly agree”). Four items were adapted from Goldsmith & Newell (1997) to measure revenge price sensitivity. For example, “In general, the price or cost of buying an airline ticket is important to me.” (1 = “strongly disagree,” and 9 = “strongly agree”). To ensure the content validity of our measurements, the original scale was translated and revised by professors and Ph.D. students specializing in English and management.

5. Results

5.1. Assessment of Measurement Model

The initial step involved the validation of the measurement model through the application of factor analysis. The outcome of this procedure revealed that the collective variance elucidated by all underlying factors accounted for 63.00%. The Kaiser-Meyer-Olkin (KMO) coefficient, gauging the interrelation of variables and the appropriateness of factor analysis, exhibited a value of 0.69. This outcome indicated a substantial interdependence among variables and the suitability of pursuing factor analysis. Moreover, Bartlett's test of sphericity yielded a statistically significant outcome ($p < 0.001$), affirming the existence of correlations among variables and endorsing the pertinence of employing factor analysis for the dataset.

To assess the presence of common method bias, we employed the Harman's single-factor method. This approach involves subjecting all variables to an exploratory factor analysis, examining the unrotated factor analysis results to determine the minimum number of factors required to explain the variance in the variables. If only one factor is extracted or if a particular factor exhibits notably high explanatory power, it suggests the presence of significant common method bias. Utilizing this method, we derived that the variance explained by the first factor amounts to 34.4%, which is below the critical threshold of 50% (Podsakoff et al., 2003). This indicates that there is no substantial common method bias present in the study.

As posited by Hair et al. (2019), two fundamental criteria were subjected to scrutiny: construct reliability and construct validity,. To assess construct reliability, the metric of composite reliability (CR) was invoked. The computed CR values for each individual construct surpassed the designated threshold of 0.7, thus signifying an acceptable degree of reliability for the measurement items. Convergent validity necessitates robust correlations among items appraising the same construct. The factor loadings for all items eclipsed the threshold of 0.6, substantiating a robust correlation with their respective constructs. Furthermore, the computed average variance extracted (AVE) for each construct exceeded the benchmark of 0.5, thereby attesting that more than half of the variance within the indicators could be attributed to their corresponding constructs. These outcomes effectively corroborate the presence of convergent validity within the measurement model. Discriminant validity, acting as a determinant of the distinctiveness of discrete constructs, was scrutinized. Results demonstrated that the square root of AVE for each construct exceeded the correlation coefficients between said construct and other coexisting constructs. This substantiates that the measurement model prominently demonstrates discriminant validity.

5.2. Hypothesis Testing

5.2.1 Manipulation checks. Initially, a manipulation check was administered to validate the integrity of our experimental design. The findings revealed that a greater proportion of participants in the AI disclosure group acknowledged their awareness of the pricing being determined by AI, compared to those in the absence of AI disclosure condition (t = 3.71, $p < .01$). This outcome underscores the efficacy of our manipulation involving AI disclosure. Furthermore, the results demonstrated that a higher percentage of participants in the explanation-provided group acknowledged their awareness of the price being accompanied by an explanation, in contrast to participants in the absence of explanation condition (t = 6.40, $p < .01$). This underscores the successful implementation of our manipulation pertaining to the provision of explanations.

5.2.2 Mediation effect of inferred motives. To test our hypothesis H1, used PROCESS macro (v. 3.3) Model 4 (Hayes, 2017) to conduct a mediation analysis with revenge behavior as the dependent variable, AI disclosure (0 = without, 0 = with) as the independent variable and inferred motives as the mediator. Results showed that the total effect of AI disclosure on revenge behavior was significantly positive (Effect = 0.33, SE = 0.18, 90% CI = [0.03, 0.63]). And the indirect effect AI disclosure → inferred motives → revenge was significant (Effect = 0.07, SE = 0.06, 90% CI = [0.001, 0.178]). These results indicated the mediation effect of inferred motives in the positive relationship between AI disclosure and revenge behavior, supporting H1.

5.2.3 Moderating effect of price sensitivity. To test our hypothesis H2, we conducted a 2 (AI disclosure: with vs. without) × 2 (price sensitivity: high vs. low) ANOVA. Participants with price sensitivity lower than the mean of 7.82 were divided into the low price sensitivity group (N = 54), and others were divided into the high price sensitivity group (N = 67). Results revealed a significant interaction effect between AI disclosure and price sensitivity on revenge behavior (F (1, 117) = 3.01, $p < 0.1$, see Figure 3).

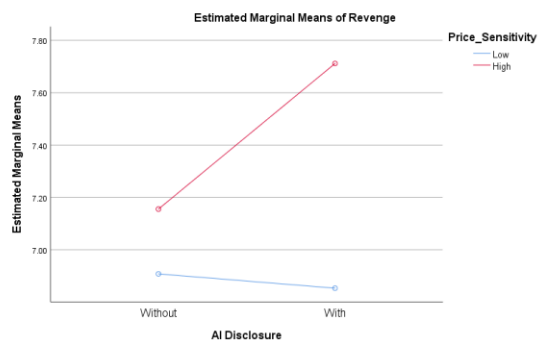


Figure 3. Revenge behavior as a function of AI disclosure and price sensitivity.

Specifically, AI disclosure increased revenge behavior to a greater extent for consumer with high price sensitivity (M_{High_WithAI} = 7.71 vs. $M_{High\_WithoutAI}$ = 7.16; t = 2.13, $p < 0.05$) than consumer with low price sensitivity (M_{Low_WithAI} = 6.85 vs. $M_{Low\_WithoutAI}$ = 6.91; t = 0.25, $p > 0.1$). Thus, hypothesis H2 was supported.

5.2.4 Mediated moderation analysis of price sensitivity. To test our hypothesis H3, we conducted a 2 (AI disclosure: with vs. without) × 2 (price sensitivity: high vs. low) ANOVA. Results revealed a significant interaction effect between AI disclosure and price sensitivity on inferred motives (F (1, 117) = 3.92, $p < 0.05$, see Figure 4). Specifically, AI disclosure increased the negative degree of inferred motives to a greater extent for consumer with high price sensitivity (M_{High_WithAI} = 7.04 vs. $M_{High\_WithoutAI}$ = 6.23; t = 2.32, $p < 0.05$) than consumer with low price sensitivity

(M_{Low_WithAI} = 6.12 vs. $M_{Low\_WithoutAI}$ = 6.62; t = 0.44, p > 0.1).

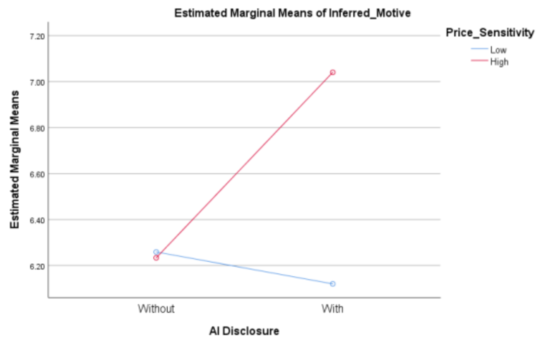
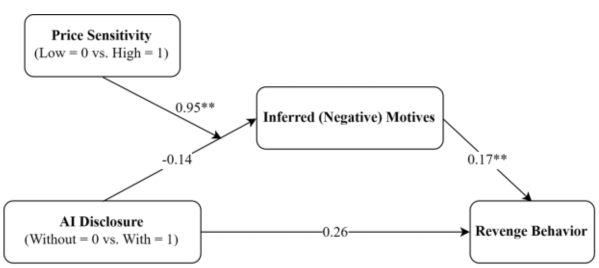


Figure 4. Inferred motives as a function of AI disclosure and price sensitivity.

Next, we used PROCESS macro (v. 3.3) Model 7 (Hayes, 2017) to conduct a mediated moderation analysis with revenge behavior as the dependent variable, AI disclosure (0 = without, 1 = with) as the independent variable, price sensitivity (0 = low, 1 = high) as the moderator and inferred motives as the mediator. In this model, the moderating effect of price sensitivity takes place before the mediator. Results revealed that the index of mediated moderation for inferred motives did not include 0 (Effect = 0.16, SE = 1.11, 90% CI = [0.01, 0.37], see Figure 5).



Notes: **p < .05.

Figure 5. Results of mediated moderation analysis.

Specifically, the indirect effect AI disclosure → inferred motives → revenge was significant on the moderator in the condition of high price sensitivity (Effect = 0.13, SE = 0.09, 90% CI = [0.02, 0.31]). For the condition of low price sensitivity, however, the corresponding indirect effect was not significant (Effect = -0.02, SE = 0.06, 90% CI = [−0.12, 0.07]). This indicated that price sensitivity increases revenge behavior by strengthening the positive impact of AI disclosure on inferred negative motives. Thus, hypothesis H3 was supported.

5.2.5 Moderating effect of explanation. To test our hypothesis H4, we conducted a 2 (AI disclosure: with vs. without) × 2 (explanation: with vs. without) ANOVA. Results revealed a significant interaction effect between AI disclosure and explanation on revenge behavior (F (1, 117) = 5.46, p < 0.05, see Figure 6). Specifically, AI disclosure increased revenge behavior to a greater extent when with explanations ($M_{WithExp\_WithAI}$ = 7.55 vs. $M_{WithExp\_WithoutAI}$ = 6.79; t = 2.90, p < 0.01) than without explanations ($M_{WithoutExp\_WithAI}$ = 7.18 vs. $M_{WithoutExp\_WithoutAI}$ = 7.27; t = 0.34, p > 0.1). Thus, hypothesis H4 was supported.

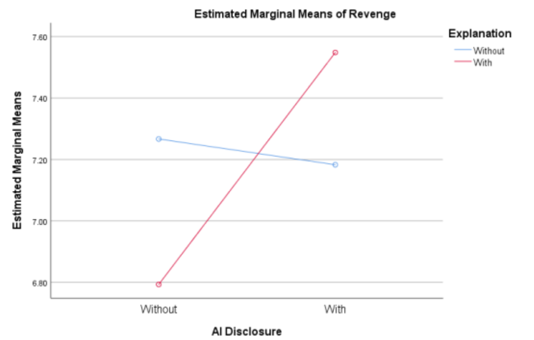


Figure 6. Revenge behavior as a function of AI disclosure and explanation.

5.2.6 Mediated moderation analysis of explanation. To test our hypothesis H5, we conducted a 2 (AI disclosure: with vs. without) × 2 (explanation: with vs. without) ANOVA. Results revealed an insignificant interaction effect between AI disclosure and explanation on inferred motives (F (1, 117) = 0.13, p > 0.1, see Figure 7). Results of mediated moderation analysis revealed that the index of mediated moderation for inferred motives included 0 (Effect = 0.03, SE = 0.09, 90% CI = [-0.09, 0.21]). Specifically, the indirect effect AI disclosure → inferred motives → revenge was not significantly different whether explanations were present or not. Thus, hypothesis H5 was not supported.

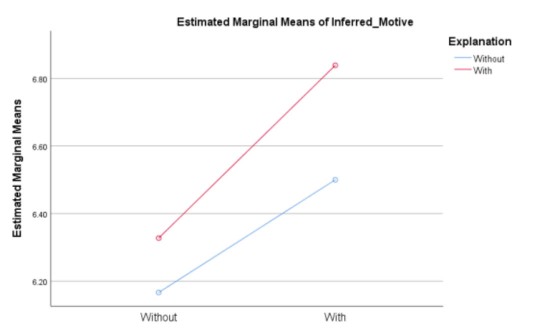


Figure 7. Inferred motives as a function of AI disclosure and explanation.

6. Discussion

6.1. Key Findings

This study examined how price sensitivity or explanation for price differences moderates the relationship between AI disclosure and consumers' revenge behavior, and whether this process is mediated by inferred motives. Results revealed that AI disclosure has a positive impact on revenge behavior through the mediation of inferred negative motives and there was a significant interaction effect between AI disclosure and price sensitivity on revenge behavior. AI disclosure increased revenge behavior to a greater extent for consumers with high price sensitivity compared to those with low price sensitivity. Inferred motives played a mediator role between AI disclosure and revenge behavior, and this mediation effect was more pronounced for consumers with high price sensitivity.

In addition, there was a significant interaction effect between AI disclosure and explanation on revenge behavior. AI disclosure increased revenge behavior to a greater extent when accompanied by explanations compared to when no explanations were provided. However, the presence or absence of explanations does not significantly affect inferred motives underlying revenge behavior.

These findings suggest that AI disclosure has a greater impact on revenge behavior especially for consumers with high price sensitivity or when explanations for discriminatory pricing are provided. This study underlies several theoretical and managerial implications.

6.2. Contribution

This research makes significant contributions to the existing literature in several ways. Firstly, it contributes to the understanding of consumer reactions towards AI by investigating the effect of AI disclosure in discriminatory pricing and focusing on consumers' revenge behavior. While previous studies have explored negative consumer reactions and reduced engagement when AI is involved in decision-making (Longoni et al., 2019), this study extends the examination of AI disclosure to the context of discriminatory pricing. Moreover, this study specifically focuses on revenge behavior, which is a crucial potential outcome resulting from AI disclosure but has not been previously examined.

Secondly, this research contributes to the literature on discriminatory pricing by examining the interaction between price sensitivity and AI disclosure in influencing revenge behavior. Previous studies have identified various factors, such as individualistic cultural characteristics, peer influence, and pricing strategies, that affect consumer behavior in the context of discriminatory pricing (Lastner et al., 2019; Li et al., 2018). However, no prior research has investigated the moderating role of price sensitivity in the impact of AI disclosure. Additionally, this study sheds light on the underlying mechanisms through which AI disclosure and price sensitivity influence consumers' revenge behavior. The findings demonstrate that AI disclosure increases revenge behavior to a greater extent for consumers with high price sensitivity, and this process is mediated by inferred motives.

Finally, this study contributes to the literature on revenge behavior by examining the interaction effect of explanation and AI disclosure on revenge behavior in the context of discriminatory pricing. In contrast to previous studies indicating the effectiveness of explanations for mitigating the negative effects of discriminatory pricing (Li & Jain, 2016), this study reveals that providing explanations is not always effective. When explanations are provided alongside AI disclosure, consumers may perceive them as attempts to conceal unethical practices or defend discriminatory pricing, leading to increased revenge behavior. Overall, this study provides novel insights into the impact of AI disclosure on consumers' revenge behavior and highlights the role of explanations in this process.

6.3. Managerial Implications

This study has important managerial implications for organizations engaging in discriminatory pricing practices and implementing AI systems. The following implications can be drawn.

Companies should carefully consider how they disclose the use of AI in decision-making processes, particularly in the context of discriminatory pricing. Given that AI disclosure can significantly increase revenge behavior, it is crucial to exercise caution and prudence in communicating the role of AI systems to consumers, avoiding negative reactions and potential retaliation from consumers. If companies can clearly explain the factors and algorithms involved in pricing decisions, it would build trust and reduce perceived unfairness.

This study indicated that price sensitivity plays a significant role in the impact of AI disclosure on revenge behavior. Managers should recognize that consumers with high price sensitivity are more likely to exhibit revenge behavior when AI is involved in discriminatory pricing. Therefore, companies could consider tailoring pricing strategies for different consumer segments based on their price sensitivity levels. This customization can help minimize negative reactions and mitigate the potential for revenge behavior.

Providing explanations alongside AI disclosure does not always mitigate revenge behavior. Companies should be cautious when offering explanations for pricing decisions, as they may be perceived as attempts to justify unfair practices or deceive consumers. It is crucial to ensure that explanations are genuine, transparent, and align with ethical standards. When providing explanations, companies should focus on addressing consumer concerns and clarifying the decision-making process rather than using explanations as a defensive mechanism.

6.4. Limitations and Future Research

Despite its contributions, this study has certain limitations that open avenues for future research.

This study identified inferred motives as a mediator between AI disclosure and revenge behavior. However, other potential mediating mechanisms may exist. Future research should explore additional psychological and cognitive factors that may mediate the relationship between AI disclosure and revenge behavior. For example, factors like trust, perceived fairness, perceived control, or emotional reactions could be examined as potential mediators.

This study focused on immediate reactions to AI disclosure and revenge behavior. Future research could explore the long-term effects of AI disclosure on consumer attitudes, intentions, and behaviors. Examining whether the initial revenge behavior subsides over time or leads to more enduring negative outcomes would provide a more comprehensive understanding of the impact of AI disclosure.

The study's findings are based on a specific sample of participants, which may limit the generalizability of the results. Future research should consider diverse samples, including participants from different demographic backgrounds, cultures, and regions, to assess the robustness of the findings across different populations.

Future research could delve into the development and evaluation of managerial interventions aimed at mitigating revenge behavior resulting from AI disclosure. Investigating strategies such as improved communication, trust-building initiatives, or fairness-enhancing mechanisms could help organizations manage and minimize the negative consequences of AI disclosure.

7. Acknowledgments

This work was partially supported by the Strategic Research Grant at the City University of Hong Kong (Grant No. 7005473) and the Hong Kong Research Grant Council (Nos. CityU 11500421 and CityU 11500322).

8. References

Allender, W. J., Liaukonyte, J., Nasser, S., & Richards, T. J. (2021). Price Fairness and Strategic Obfuscation. *Marketing Science*, *40*(1), 122–146. https://doi.org/10.1287/mksc.2020.1244

Ayadi, N., Paraschiv, C., & Rousset, X. (2017). Online dynamic pricing and consumer-perceived ethicality: Synthesis and future research. *Recherche et Applications En Marketing (English Edition)*, *32*(3), 49–70. https://doi.org/10.1177/2051570717702592

Bechwati, N. N., & Morrin, M. (2003). Outraged consumers: Getting even at the expense of getting a good deal. *Journal of Consumer Psychology*, *13*(4), 440–453. https://doi.org/10.1207/S15327663JCP1304_11

Borgesius, F. Z., & Poort, J. (2017). Online Price Discrimination and EU Data Privacy Law. *Journal of Consumer Policy*, *40*(3), 347–366. https://doi.org/10.1007/s10603-017-9354-z

Campbell, M. C. (2007). “Says who?!” How the source of price information and affect influence perceived price (un) fairness. *Journal of Marketing Research*, *44*(2), 261–271. https://doi.org/10.1509/jmkr.44.2.261

Chowdhry, A. (2016, May 25). Uber: Users Are More Likely To Pay Surge Pricing If Their Phone Battery Is Low. *Forbes*. https://www.forbes.com/sites/amitchowdhry/2016/05/25/uber-low-battery/

Chung, J. Y., & Petrick, J. F. (2013). Price Fairness of Airline Ancillary Fees: An Attributional Approach. *Journal of Travel Research*, *52*(2), 168–181. https://doi.org/10.1177/0047287512457261

Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the Bot: Anthropomorphism and Anger in Customer–Chatbot Interactions. *Journal of Marketing*, *86*(1), 132–148. https://doi.org/10.1177/00222429211045687

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Garbarino, E., & Maxwell, S. (2010). Consumer response to norm-breaking pricing events in e-commerce. *Journal of Business Research*, *63*(9), 1066–1072. https://doi.org/10.1016/j.jbusres.2008.12.010

Garvey, A. M., Kim, T., & Duhachek, A. (2023). Bad News? Send an AI. Good News? Send a Human. *Journal of Marketing*, *87*(1), 10–25. https://doi.org/10.1177/00222429211066972

Goldsmith, R. E., Flynn, L. R., & Kim, D. (2010). Status Consumption and Price Sensitivity. *Journal of Marketing Theory and Practice*, *18*(4), 323–338. https://doi.org/10.2753/MTP1069-6679180402

Goldsmith, R. E., & Newell, S. J. (1997). Innovativeness and price sensitivity: Managerial, theoretical and methodological issues. *Journal of Product & Brand Management*, *6*(3), 163–174. https://doi.org/10.1108/10610429710175682

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hau, H., Hoffmann, P., Langfield, S., & Timmer, Y. (2021). Discriminatory Pricing of Over-the-Counter Derivatives. *Management Science*, *67*(11), 6660–6677. https://doi.org/10.1287/mnsc.2020.3787

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hufnagel, G., Schwaiger, M., & Weritz, L. (2022). Seeking the perfect price: Consumer responses to personalized price discrimination in e-commerce. *Journal of Business Research*, *143*, 346–365. https://doi.org/10.1016/j.jbusres.2021.10.002

Joireman, J., Grégoire, Y., Devezer, B., & Tripp, T. M. (2013). When do customers offer firms a “second chance” following a double deviation? The impact of inferred firm motives on customer revenge and reconciliation. *Journal of Retailing*, *89*(3), 315–337. https://doi.org/10.1016/j.jretai.2013.03.002

Kim, T. W., & Duhachek, A. (2020). Artificial Intelligence and Persuasion: A Construal-Level Account. *Psychological Science*, *31*(4), 363–380. https://doi.org/10.1177/0956797620904985

Kumar, V., Dixit, A., Javalgi, R., & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, *44*(1), 24–45. https://doi.org/10.1007/s11747-015-0426-9

Lastner, M. M., Fennell, P., Folse, J. A., Rice, D. H., & Porter III, M. (2019). I guess that is fair: How the efforts of other customers influence buyer price fairness perceptions. *Psychology & Marketing*, *36*(7), 700–715. https://doi.org/10.1002/mar.21206

Li, K. J., & Jain, S. (2016). Behavior-Based Pricing: An Analysis of the Impact of Peer-Induced Fairness. *Management Science*, *62*(9), 2705–2721. https://doi.org/10.1287/mnsc.2015.2265

Li, W., Hardesty, D. M., & Craig, A. W. (2018). The impact of dynamic bundling on price fairness perceptions. *Journal of Retailing and Consumer Services*, *40*, 204–212. https://doi.org/10.1016/j.jretconser.2017.10.011

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Longoni, C., & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The “word-of-machine” effect. *Journal of Marketing*, *86*(1), 91–108. https://doi.org/10.1177/0022242920957347

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, *38*(6), 937–947. https://doi.org/10.1287/mksc.2019.1192

MacCarthy, M. (2020, March 9). AI needs more regulation, not less. *Brookings*. https://www.brookings.edu/research/ai-needs-more-regulation-not-less/

Mao, J. Y., & Benbasat, I. (2000). The use of explanations in knowledge-based systems: Cognitive perspectives and a process-tracing analysis. *Journal of Management Information Systems*, *17*(2), 153–179. https://doi.org/10.1080/07421222.2000.11045646

Newcomer. (2017). Uber Starts Charging What It Thinks You’re Willing to Pay. *Bloomberg*. https://www.bloomberg.com/news/articles/2017-05-19/uber-s-future-may-rely-on-predicting-how-much-you-re-willing-to-pay

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879. https://doi.org/10.1037/0021-9010.88.5.879

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Seele, P., Dierksmeier, C., Hofstetter, R., & Schultz, M. D. (2021). Mapping the ethicality of algorithmic pricing: A review of dynamic and personalized pricing. *Journal of Business Ethics*, *170*(4), 697–719. https://doi.org/10.1007/s10551-019-04371-w

Shea, L. J. (2010). Using consumer perceived ethicality as a guideline for corporate social responsibility strategy: A commentary essay. *Journal of Business Research*, *63*(3), 263–264. https://doi.org/10.1016/j.jbusres.2009.04.021

Shukla, N., Kolbeinsson, A., Otwell, K., Marla, L., & Yellepeddi, K. (2019). Dynamic Pricing for Airline Ancillaries with Customer Context. *Kdd’19: Proceedings of the 25th Acm Sigkdd International Conferencce on Knowledge Discovery and Data Mining*, 2174–2182. https://doi.org/10.1145/3292500.3330746

Wang, Y., & Krishna, A. (2012). Enticing for me but unfair to her: Can targeted pricing evoke socially conscious behavior? *Journal of Consumer Psychology*, *22*(3), 433–442. https://doi.org/10.1016/j.jcps.2011.11.004

Weisstein, F. L., Monroe, K. B., & Kukar-Kinney, M. (2013). Effects of price framing on consumers’ perceptions of online dynamic pricing practices. *Journal of the Academy of Marketing Science*, *41*(5), 501–514. https://doi.org/10.1007/s11747-013-0330-0

Yalcin, G., Lim, S., Puntoni, S., & van Osselaer, S. M. J. (2022). Thumbs Up or Down: Consumer Reactions to Decisions by Algorithms Versus Humans. *Journal of Marketing Research*, 002224372110700. https://doi.org/10.1177/00222437211070016