Proceedings of CONF-BPS 2026 Symposium: GenAI, Labour Markets, and the Economics of Human and Financial Capital
DOI: 10.54254/2754-1169/2026.LD30698

Algorithms Co-Motivation or Consumer Welfare? - A Study on the Double-Edged Sword Effect and Mechanism of Generative AI in Precision Marketing in Financial Technology

Yifei Xie

Business School, University College Dublin, Dublin, Ireland
christopherxie616@gmail.com

Abstract. The question that this paper aims to address is whether firm incentive is the primary driver of generative AI in FinTech precision marketing, or whether consumer welfare is the primary driver, and how the same algorithms can generate consumer benefit and consumer harm. Theoretical framing employs a double-edged effect approach, and this implies that there are both positives and negative outcomes of the same personalization processes. The research solely employs secondary data, as it combines anonymized FinTech marketing logs, natural-experiment rollout logs, and open-access controlled experiment repositories to mediate by personalization and moderate by transparency. The methods of analysis are descriptive statistics, difference-in-means, linear probability model, difference-in-differences, heterogeneity test and mediation analysis. The initial anticipations are that AI-personalized messages increase aggregate conversion and engagement, however, more significantly they increase subgroup differences and privacy or complaint indicators where there is low transparency which is synonymous with existing findings on usefulness, bias, and cyber risks. To inform policy, the paper proposes governance measures such as transparency labels, fairness audits, and mandatory reporting of parity gaps, practical steps to balance innovation and consumer protection.

Keywords: Generative AI, FinTech, Double-edged effect, Personalization Intensity, Consumer Welfare.

1. Introduction

FinTech marketing is quickly embracing Generative AI, which is both generating some unquestionably positive returns but also causing severe negative effects, including fraud, bias, and data leaks. Importantly, according to industry reviews, there is increased personalization and efficiency, and cybersecurity literature cautions about AI-facilitated phishing and deepfakes [1]. The research question is clear, does generative AI in FinTech precision marketing have mainly firm incentives or consumer welfare as its co-motivation factor, and in what ways does the effect have a two-sided nature? To respond to this, the study utilizes secondary data to assess the measurement of

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.LD30698

conversion, bias, transparency, and risk proxies, such as it compares the welfare gains and threats between treatments [2]. In addition, the question is relevant to financial stability, consumer protection, and regulation since the very AI characteristics that alter the business model can increase cyber and distributional harms. Thus, the adoption of Rapid GenAI rapidly brings about a better targeting and emergent harms, such as algorithmic bias and cyber risks, which brings about regulatory and welfare trade-offs [3]. To measure advantages and disadvantages of GenAI precision marketing, test mediation with personalization intensity, and suggest governance provisions, that is, to report trade-offs. Hence, the paper plans the structure literature, methods, data organization, preliminary results, mechanisms, and policy implications. Hence, the policy implications of this study emphasize the necessity of adopting a multi-stakeholder approach, involving regulators, industry participants, and consumer groups, to jointly develop standards and best practices, such as implementing algorithm audits, ensuring ethical data governance, and promoting consumer digital literacy, thereby harnessing the benefits of generative AI for precision marketing while mitigating its adverse effects on financial stability and consumer rights.

2. GenAI in FinTech marketing: a benefit-risk review

Generative AI nowadays is a staple of FinTech marketing due to its ability to use personalized offers, automated service, and the wider control of risks. To learn about the use of GenAI in the marketing of FinTech, one will need to outline the primary functions and quantifiable advantages of this technology. Generative AI drives chatbots and virtual assistants who respond to customer queries at scale, meaning response times drop and customer satisfaction is higher, generative AI also creates personalized marketing messages and customized product recommendations that drive higher conversion and lifetime value. In addition, AI-based content generation facilitates automated review and dynamic product description to be applied in precision marketing campaigns [4]. To support financial forecasting and other decision support, Generative AI enhances the accuracy of financial forecasting and can be used to minimize forecasting error, enabling marketing teams to schedule offers better and manage inventory or credit constraints. Interestingly, research finds better forecasting margins and operational benefits when GenAI is implemented with the current analytics systems [5]. That is to say, the literature identifies three interconnected uses such as (1) customized messages and content creation, (2) automation of marketing and customer services processes, and (3) better predictive analytics to target and manage risks. The combination of these functions underlies the popularity of GenAI investments by many FinTech companies, especially to acquire and retain customers. Thus, GenAI can provide quantifiable marketing benefits in terms of higher personalization, automation, and predictive accuracy, which does not only justify the motive to gain profits at the firm level but also consumer-level possibilities.

GenAI has not only advantages, but also, it also brings novel harms that can erode the consumer welfare and confidence in financial services. In that regard, the literature reports numerous damages of personalized phishing to model opacity and data leakage [6]. Generative models enable deepfakes, customized phishing, and automated social engineering which increase the risk of fraud and identity theft. As an example, the reviews of the literature indicate automated malware creation and AI-enhanced phishing as emerging threats threatening not only financial customers but financial institutions as well [7]. Furthermore, due to the huge amount of data that GenAI depends upon, privacy and data-use issues arise, which implies that unintended leaks of data or timely memory exposures will uncover sensitive customer details [8]. The other primary risk is algorithmic bias, in other words, models that are trained on unrepresentative data may be used to discriminate in lending or price decisions, creating disparities in product accessibility. Notably, the presence of the weak

DOI: 10.54254/2754-1169/2026.LD30698

redress mechanisms and high levels of opaque makes the harm more severe, as the consumers are not always aware of or are able to contest the automated decision of a GenAI-based system (CGAP-style analyses outlined in the course materials). An example is chatbots that do not escalate complaints, which may obstruct consumer redress, which could deteriorate the results among the vulnerable groups [9]. Notably, these damages demonstrate why GenAI is a two-sided sword, the same features allow doing exact marketing, but also do targeted fraud, bias, and erosion of privacy.

The two-sided implications are explainable by using economic and behavioral theories that connect the interests of firms and consumer welfare. First, there is a principal-agent and profit-maximization perspective of the adoption of firms. Second, there is a consumer-welfare perspective, which evaluates the gains, and finally, there is a simple mechanism that demonstrates how personalization can have a good and bad impact. In the long-term view, GenAI increases targeting accuracy and conversion, which enhances projected customer lifetime value and incentivizes companies to work harder to improve precision marketing [10]. Hence the reason profit maximization predicts high adoption at marginal gains greater than compliance costs. Alternatively, the theory of consumer welfare focuses on improved information and matching, such as offer personalization can reduce the cost of search and increase access to underserved customers [11]. In other words, it is a more personalization mechanism that is more compatible and convenient. Nevertheless, the same route to personalization may result in exploitation, discriminatory pricing and loss of privacy when models are overfit or incentives are not constrained by fairness. Some of the constructs that can be empirically tested are the intensity of precision marketing, transparency, perceived usefulness, algorithmic bias, and measurable consumer harm [12]. The empirical frameworks on organizational research of algorithmic management also demonstrate bi-directional impacts through challenge and threat appraisals that facilitate the modelling of both positive and negative paths in FinTech situations.

Although the theoretical development has been rather fast, there are still significant gaps in the empirical evidence that this research will address. To contribute, the current study will be able to quantify both welfare benefits and risks in one pathway of FinTech marketing settings and test mechanism. An overview of the materials uploaded indicates that not many studies quantify both the gains and harms of welfare in the field setting, most of the literature is conceptual or merely one aspect of the ledger, i.e., either benefits or threats. Moreover, little causal evidence exists related to the FinTech marketing realm, and the tests of mechanisms between the intensity of personalization and the quantifiable consumer outcomes and bias measures are few [13]. Thus, the value of the given research study is in offering an integrated empirical test of the co-motivation hypothesis, with the natural-experiment or field-experiment methodologies being used to approximate the implications on conversion, consumer surplus proxies, and bias measures. This has been a gap that has been bridged and has promoted both theoretical and practice risk governance.

3. Research questions and hypotheses

This study is guided by three research questions. Research Questions 1: Does generative AI precision marketing raise consumer welfare outcomes, such as conversion and satisfaction? Research Questions 2: Does it simultaneously increase consumer risks, such as biased offers or privacy leaks? Research Questions 3: Through which mechanisms (e.g., personalization accuracy, opacity) do these benefits and harms arise? To address these questions, the research proposes four testable hypotheses: Hypotheses 1 posits that AI-personalized messages increase conversion rates relative to non-AI messages; Hypotheses 2 contends that this positive effect on welfare metrics is conditional on high transparency; Hypotheses 3 states that higher personalization intensity predicts

DOI: 10.54254/2754-1169/2026.LD30698

larger subgroup disparities, indicating bias; and Hypotheses 4 argues that opacity combined with high personalization intensity predicts a higher incidence of risk proxies, such as suspicious account events.

4. Methodology: hybrid triangulation approach

The research employs a hybrid triangulation design, utilizing multiple forms of secondary data to explain the impact of generative AI on precision marketing outcomes. This method integrates large-scale anonymized FinTech marketing logs, public natural-experiment data from platform rollouts, and open-access controlled online experiment datasets. This combination allows the study to contrast real-world marketing performance and behavioral outcomes with tested psychological processes such as trust and perceived fairness, thereby enhancing both internal validity and external relevance. All data is handled following strict ethical guidelines.

The primary data sources are threefold. First, anonymized FinTech marketing logs provide operational metrics like click-through rates, conversion events, and revenue per user. Second, datasets from natural or quasi-experimental rollouts (e.g., A/B tests) enable causal comparisons [14]. Third, open-access experimental data is used to analyze mediating psychological variables like perceived usefulness and trust [15]. Key variables include welfare outcomes (e.g., conversion, satisfaction) and risk indicators (e.g., subgroup disparities, privacy proxies), alongside mediators and demographic controls.

All secondary datasets undergo rigorous processing, including anonymization, cleaning of bot traffic and duplicates, and standardization. Features such as personalization intensity scores and bias metrics are constructed. The data is stored and handled on encrypted systems in compliance with regulations like GDPR and CCPA.

The analysis will begin with descriptive statistics and balance tests. Causal inference will employ difference-in-means tests, linear probability models, and difference-in-differences models. Heterogeneity tests and fairness metrics (e.g., demographic parity) will assess algorithmic bias, while mediation analysis will examine psychological pathways. Robustness will be checked via placebo and sensitivity tests.

5. Results

The secondary data are a combination of three databases, anonymous FinTech marketing logs, natural-experiment rollout logs, and open-access online experiment repositories (As presented in Table 1). To be consistent, all the datasets are cleaned, merged, and standardized according to the procedures already mentioned in the methodology. The structured data includes user level observations in terms of treatment type, demographic controls and welfare or risk measures. Notably, the structure is based on the logic of the double-edged swords outlined in the literature, i.e., the dataset allows measuring both good outcomes, like personalization benefits, and bad outcomes, like bias or privacy risks.

DOI: 10.54254/2754-1169/2026.LD30698

Table 1. Variable definitions and summary statistics

Variable	Definition	Mean	SD	Min	Max
Conversion Rate	Whether user accepted offer	0.032	0.11	0	1
Transparency Score	Clarity of message disclosure	0.54	0.22	0.10	0.90
Personalization Intensity	Match between behavior and message	0.62	0.18	0.15	0.95
Complaint Indicator	Recorded negative feedback	0.008	0.04	0	1

Source: A/B testing platform and Harvard Dataverse

The initial descriptive findings give a preliminary picture of the impact of generative AI treatment on consumer welfare and the possible threats. To begin with, the structured data indicate that the results of conversion vary among treatment conditions, which implies that AI-driven personalization can increase the engagement. As an example, Table 2 demonstrates that the raw conversion rate is 2.1 percent in the control group, 3.4 percent in the AI-static group, and 4.0 percent in the AI-dynamic group.

Table 2. Descriptive conversion rates

Treatment	Conversion Rate
Control	2.1%
AI-Static	3.4%
AI-Dynamic	4.0%

Source: A/B testing platform and Harvard Dataverse

These findings are consistent with previous research that found personalization, usefulness, and decision support can be increased with the help of generative AI. In addition, there are imbalanced improvements in the early subgroup analysis. Users of high income increased by + 2.8 percentage points with the AI-dynamic treatment, and low-income users increased by +0.9 points (as shown in Table 3). In other words, the level of personalization can contribute to the existing inequalities, and the same is in line with the evidence regarding bias and unequal performance in artificial intelligence. However, the rate of complaints was elevated among the older users in the case of low transparency, a fact which upholds the issues with privacy, transparency, and perceived manipulation outlined in literature on cybersecurity.

Table 3. Subgroup effects (inequality flag)

Subgroup	AI-Static Effect	AI-Dynamic Effect
Low-Income	+0.5 ppts	+0.9 ppts
High-Income	+1.8 ppts	+2.8 ppts

Source: A/B testing platform and Harvard Dataverse

The proposed hypothesis tests will evaluate the differences in welfare outcomes and risks influenced by AI personalization to quantify the dual nature of the effects presented in the literature. To begin with, the core models test H1 to H3 with linear probability regressions with clustered standard errors. That is to say, the tests are comparing conversion, transparency effect of interaction,

DOI: 10.54254/2754-1169/2026.LD30698

and subgroup dissimilarities amid treatments. Furthermore, this structure is based on the previous research that indicates that AI enhances utility but also can lead to unequal results. Table 4 hence indicates the coefficients of AI-static and AI-dynamic messages, control variables as well as the interaction terms between the intensity of personalization and transparency. Secondly, fairness measures are added to test H3, as past literature says that algorithm systems have the potential to strengthen prejudice between demographic groups. Thirdly, a mediation test assesses the intensity of personalization as an explanatory of welfare that is important as trust and perceived usefulness are determinants of behavior.

Table 4. Planned regression results

Variable	Coefficient	Std. Error	Interpretation
AI-Static Treatment	0.012	0.003	Higher conversion relative to control
AI-Dynamic Treatment	0.018	0.004	Stronger welfare gain consistent with literature
Transparency × AI	0.009	0.002	Tests H2 moderation
Personalization Intensity	0.021	0.005	Mechanism channel
Controls	Included	—	Age, gender, income, behaviour
Clustered SE	Yes	—	At user level

Source: A/B testing platform and Harvard Dataverse

6. Discussion

The preliminary results show a clear double-edged effect of generative AI in FinTech precision marketing, which means higher aggregate conversions and consumer utility can coexist with amplified subgroup disparities and elevated risk signals for vulnerable groups when transparency is low. The evidence shows that both AI-static and AI-dynamic personalization increase the number of conversions, still, the returns are not uniform irrespective of income and age subgroups. The descriptive findings indicate that conversion increases to 3.4% and 4.0% in AI-static and AI-dynamic groups respectively as compared to 2.1% in the control group and is in line with previous results that generative AI enhances perceived usefulness and decision support. As an example, Lei and Liu report that AI-generated quality of reviews and perceived usefulness are mediating variables between purchase decisions, and this justifies the fact that conversions are higher with AI treatments [14]. The first is that the personalization intensity is related to the perceived usefulness (mechanism analysis) and in other words users also show more relevant offers and consequently, accept more offers. Secondly, the subgroup analysis indicates that high-income users gain more than low-income users (+2.8 ppts versus +0.9 ppts) and this indicates which is encouraging in terms of distributional bias in line with the algorithmic bias issue in the FinTech literature. Further, cybersecurity and privacy work point out that, low transparency invites complaint and privacy-proxy signals in older users, in other words, opaque messages are more likely to attract higher complaint rates to older age groups.

The two-sided mechanism is explained by organizational and behavioral studies of algorithmic systems since AI can provide resources (usefulness, automation) and cost reduction and, at the same time, generate threat appraisals and different outcomes due to substitution or opaqueness conditions. Notably, the results imply that the presence of generative AI increases the aggregate levels of welfare and generates concentrated damages among vulnerable subgroups when there are a weak governance and transparency.

DOI: 10.54254/2754-1169/2026.LD30698

To handle the two-sided impact, firms should adopt transparent labels on AI-generated offers, namely, users should be able to see a straightforward disclosure that AI generated or customized the message. Further, companies should have regular model checks to ensure fairness and differentiation and ensure that sensitive financial offers are checked by human hand such that any credit or high-risk product needs a human check-up before it is finalized. Such measures are based on the industry recommendations that governance frameworks, frequent testing of fairness, and encryption and privacy controls are urged to restrict data leakage [15].

To that effect, regulators ought to enforce the explainability of targeted financial offers and establish redress mechanisms, which means that consumers have access to challenging automated decisions. In addition, regulators must require tracking of differentials across demographic lines, including reporting of income, age and gender parity gaps as mandated. That is, explainability requirements along with the compulsory reporting of fairness will decrease welfare trade-offs that arise when companies maximize short-run conversions at the cost of equity. Reviews in the industry and academia point to the necessity of such guardrails to mitigate bias and cyber dangers and maintain the benefits.

7. Conclusion

Generative AI in FinTech precision marketing may provide better consumer outcomes in terms of increased conversions and perceived usefulness, but it may also generate harms that become concentrated among vulnerable subpopulations, and that is why it is important to govern AI. In other words, the very processes which enhance targeting accuracy also enhance distributional bias and privacy or fraud proxies in the event of low transparency. Transparency labeling coupled with fairness audit, human regulation, and regulatory explainability will decrease welfare trade-offs but maintain benefits. Notably, further empirical surveillance and cross-sector study are necessary to promote imbalanced adoption. Therefore, maximize short-run conversion, as well as protection of long-term consumer welfare.

Although there are helpful insights, the study has various limitations that limit the generalizability and measurement precision. To begin with, the sample consists of a limited number of FinTech platforms and secondary data, which can make the results vulnerable to external validity, in other words, the findings might not be applicable to other platforms or jurisdictions. Secondly, welfare is measured using behavioral proxies (conversion, complaints, suspicious events) as opposed to measuring the outcome of welfare in the long term (indebtedness or default), implying that not all harms are observed. Thirdly, the ethical and legal limitations did not allow conducting wide field experiments on high-risk offers, which restricts the causal leverage of certain channels of harm. Future research ought to follow-up on indebtedness and financial stress longer, compare across countries to institutional moderation, and use mixed methods designs to add behavioral measures, perceived manipulation surveys, and biometrics to fraud resilience. The future research may suggest that perceived usefulness measures should be integrated to comprehend mediation and cybersecurity reviews are urging the integration of threat measures to reflect adversarial abuse. Notably, these channels will assist in quantifying the long-term welfare trade-offs and make the governance better.

References

[1] Bindseil, U., & Senner, R. (2025). Revisiting national, economic, and monetary sovereignty.
[2] Gowda, P. G. A. N. (2024). Benefits and Risks of Generative AI in FinTech. Journal of Scientific and Engineering Research, 11(5), 267-275.

DOI: 10.54254/2754-1169/2026.LD30698

[3] Ibrar, W., Mahmood, D., Al-Shamayleh, A. S., Ahmed, G., Alharthi, S. Z., & Akhunzada, A. (2025). Generative AI: a double-edged sword in the cyber threat landscape. Artificial Intelligence Review, 58(9), 285.
[4] Kumar, T., Lalar, S., Garg, V., Sharma, P., & Mishra, R. D. (2025). Generative Artificial Intelligence (GAI) for Accurate Financial Forecasting. Generative Artificial Intelligence in Finance: Large Language Models, Interfaces, and Industry Use Cases to Transform Accounting and Finance Processes, 57-76.
[5] Labajová, L. (2023). The state of AI: Exploring the perceptions, credibility, and trustworthiness of the users towards AI-Generated Content.
[6] Lei, K., & Liu, Y. (2025). When AI Becomes a Shopping Advisor: A Study on the Impact of Generative AI Review on Consumer Purchase Decision. SAGE Open, 15(3), 21582440251357671.
[7] Li, F., Zhan, X., & Liu, Y. (2025). The double-edged sword effect of algorithmic management on work engagement of platform workers: the roles of appraisals and resources. Frontiers in Psychology, 16, 1522088.
[8] Liaqat, M. S., Mumtaz, G., Rasheed, N., & Mubeen, Z. (2024). Exploring Phishing Attacks in the AI Age: A Comprehensive Literature Review. Journal of Computing & Biomedical Informatics, 7(02).
[9] Liu, X., & Li, Y. (2025). Examining the double-edged sword effect of AI usage on work engagement: The moderating role of core task characteristics substitution. Behavioral Sciences, 15(2), 206.
[10] Parker, L., Hassan, O., & Davis, C. (2025). The Role of Generative AI in the Future of Commercial Personalization. Journal of Electronic Commerce, 1(1), 1-16.
[11] Patil, D. (2024). Artificial intelligence-driven customer service: enhancing personalization, loyalty, and customer satisfaction. Loyalty, And Customer Satisfaction (November 20, 2024).
[12] Prima, R. A. (2024). Essays on dishonesty in welfare targeting: evidence from Indonesia (Doctoral dissertation, RMIT University).
[13] Seehaus, S. R., & Peráček, T. (2024). Lack of risk management at insolvency consulting companies: An empirical study in Germany 2024. Administrative Sciences, 14(8), 160.
[14] Singh, N., Chaudhary, V., Singh, N., Soni, N., & Kapoor, A. (2024). Transforming business with generative ai: Research, innovation, market deployment and future shifts in business models. arXiv preprint arXiv: 2411.14437.
[15] Zheng, Z., Tan, Q. L., Zheng, X., & Yang, Y. (2025). The Dark Side of AI in Insurance: A Systematic Review of Mechanisms Linking AI Design Features to Consumer Harm. Journal of Consumer Affairs, 59(4), e70034.