Generative AI, Platform Stances, and Content Creator Behavior

Hongxian Huang Runshan Fu Anindya Ghose

Abstract

Generative Artificial Intelligence (AI) technologies have emerged as a transformative force in the content creator economy. This paper examines how the AI stances signaled by specific platform policy decisions shape creator behavior on visual arts platforms. We leverage two independent natural experiments on leading Chinese platforms: Lofter's launch of an AI image generator, which signaled a pro-AI stance, and Graffiti Kingdom's prohibition of AI-generated artwork, which signaled an anti-AI stance. Our analysis shows that creators decreased their activity on Lofter following the AI generator launch, while activity increased on Graffiti Kingdom after the AI prohibition. Multiple lines of evidence show these effects stem from policy-communicated stance signaling: creators respond to what the focal AI policy communicates about AI's role, the platform's commitment to human creators, and the future competitive environment for human-created work, rather than merely to specific tool features or enforcement actions. Heterogeneity analysis reveals that higher-popularity, multi-homing, and AI-averse creators show larger activity reductions on Lofter. Through analysis of creator posts, we identify three primary concerns driving resistance: replacement risk, perceived low quality, and copyright infringement. To our knowledge, this is the first paper to causally identify how the AI stance signaled by a platform policy decision affects creator behavior. Our findings have implications for how platforms communicate AI-related policy decisions, and for policymakers on fostering a constructive relationship between AI and human creators.

Keywords: Generative AI, Content Creator, Platform Stance

1 Introduction

The rapid advancement of Generative Artificial Intelligence (AI) has unleashed a new wave of technological innovation across various domains. These AI systems can now generate high-quality content ranging from images and music to text and code, exhibiting capabilities that increasingly parallel human creativity (Epstein et al. 2023). For instance, in August 2022, an AI-generated artwork won first prize at the Colorado State Fair Fine Arts Exhibition, demonstrating that Generative AI can even compete with professional artists' works.[1]

Generative AI offers a wealth of opportunities for content creators. The AI tools can automate routine aspects of the creative process, allowing creators to focus more on conceptualization and artistic vision. They act as powerful assistants that enhance creative possibilities through novel suggestions and variations. Moreover, by lowering technical barriers, Generative AI has the potential to democratize content creation, which could enable broader participation in creative industries. It is therefore no surprise that this disruptive technology has garnered considerable attention and sparked various reactions across the industry. Several digital platforms dedicated to AI-generated works have emerged, including Midjourney, Artbreeder, and Craiyon. Online entertainment platforms like TikTok have also integrated Generative AI filters and AI-powered virtual assistants.

However, the rise of Generative AI also creates substantial challenges for the creator economy from two broad directions. First, intellectual property (IP) concerns remain acute: (i) the authorship and ownership status of AI-assisted or AI-generated outputs remains unsettled, with recent U.S. Copyright Office guidance and judicial rulings reiterating a human-authorship baseline (Blaszczyk 2025, Crusey 2025, Lemley 2023); (ii) model training often relies on datasets assembled without clear consent or compensation for rights holders (Lucchi 2025, Sag 2023); and (iii) Generative AI systems enable "style cloning", which can appropriate the economic value of distinctive artistic signatures even without literal copying (Sobel 2024, Shan et al. 2023). Second, Generative AI threatens content creators with displacement (Eloundou et al. 2023, Green 2024). Because models can generate near-instant images and text at effectively zero marginal cost, they flood the marketplace with similar outputs, expanding supply and compressing prices for creators' works. The threat is greatest where demand centers on standardized outputs (e.g., poster-style visuals, character renders, short captions) and the demand side is relatively indifferent to authorship.

Platforms have responded to these tensions by adopting different AI-related policies. We define a platform's *stance* toward Generative AI as its expressed orientation toward the role that AI creation and AI-generated content should play on the platform. A stance can be expressed through many platform actions; our focus is on the stance communicated by a single focal policy decision, which we infer from the substantive direction of that policy (for example, whether it enables, permits, labels, segregates, restricts, or prohibits AI creation). In this paper, we use "stance" as operational shorthand for what a focal policy itself signals, which we view as one important and observable component of the platform's broader ideology toward AI rather than a comprehensive characterization of it. Such a focal policy is also a publicly visible governance choice about how creative production on the platform will be structured going forward, and the stance it conveys carries implications for that future production environment. Some platforms have actively developed AI

[1] See The Verge: An AI-generated artwork won first place at a state fair fine arts competition, and artists are pissed. Last accessed July 22, 2026.

tools, which signal an embrace of AI creation. Others have prohibited AI-generated content, which signals a stronger commitment to human creators and a protected environment for human-created work. Still others have adopted intermediate positions such as requiring labels or creating separate sections for AI content. These responses create a spectrum from pro-AI to anti-AI stances.

Work on psychological contracts suggests that participants form beliefs about reciprocal obligations from the organizations with which they interact (Rousseau 1989), and recent evidence extends this to digital platforms (She et al. 2020, Frazier 2024). In this sense, when a focal AI policy communicates a visible stance toward Generative AI, creators may read the policy as a signal about the future value of human-created work, the platform's willingness to protect creators like them, and where their creative effort is best invested going forward. Such belief updating can also concern the expected competitive environment on the platform: for example, an anti-AI prohibition may signal that human creators will face a more protected, less AI-dominated production environment, which can lead incumbent creators to allocate more effort to the platform or reactivate. Understanding how creators respond to such policy-communicated stance signals is therefore critical, not only for platforms seeking to maintain vibrant creator communities, but also because Generative AI systems rely on high-quality human-created content for training and improvement. As OpenAI co-founder Ilya Sutskever explained, data is the "fossil fuel" of AI.[2]

These considerations raise three critical research questions. First, how do content creators respond to the AI stances signaled by specific platform policy decisions? Second, what are the main driving forces behind these responses? Third, how do responses vary across different creator segments?

To answer these questions, we leverage two separate natural experiments on leading Chinese platforms: Lofter and Graffiti Kingdom. Lofter, founded by NetEase in 2011, is one of China's leading content creation platforms for visual arts and fan fiction, with over 80 million interest tags and 12.8 million content creators. On March 6th, 2023, Lofter launched a full-featured AI image generator (marketed as a "Profile Picture Generator" but functionally similar to Midjourney), signaling a pro-AI stance through active platform investment in AI creation tools. Graffiti Kingdom, with nearly 20 years of history, is a platform renowned for original, high-quality illustrations, with over 500,000 registered users. On February 22nd, 2023, Graffiti Kingdom announced that AI-generated artwork was prohibited and that all previously uploaded works would be reviewed. Violators of this policy would be banned from their accounts, consistent with the anti-AI stance communicated by the prohibition, while their accounts and any past works that were not produced using AI would remain. The two focal policy decisions, Lofter's launch of a platform-developed AI tool and Graffiti Kingdom's explicit AI prohibition, anchor opposite ends of the AI-policy stance spectrum, and both intervention types (platform-developed AI tools and explicit AI prohibitions) have been adopted at multiple comparable visual-art platforms (see Online Appendix Tables A.1 and A.2). By examining these two policy events that communicate opposite AI stances, we can establish boundary conditions for creator responses and provide insights into how creators might respond to intermediate AI-policy signals.

A key identification challenge is that platform-wide policy changes affect all users simultaneously. To address this, we apply the Temporal Causal Inference (TCI) framework (Turjeman and Feinberg 2023), which constructs control groups from creators who joined the platform at different times but are observed

[2] See The Verge: What Ilya Sutskever sees as the next leap in AI. Last accessed July 22, 2026.

at similar platform tenure and estimate treatment effects via Difference-in-Differences. We validate the findings with alternative estimation methods (Generalized Synthetic Control), alternative identification strategies (Regression Discontinuity in Time, Panel Difference Estimators), alternative outcome measures, and alternative comparison-group constructions and sample restrictions.

Our analysis reveals that on Lofter, the launch of the AI image generator led to a significant decrease in uploaded works. Notably, this negative effect persisted even after Lofter removed the tool. By contrast, we find a positive average treatment effect on Graffiti Kingdom, suggesting that the platform's prohibition of Generative AI increased the activity of its creators. Analysis of creator posts on Lofter reveals three primary concerns driving resistance: replacement risk, perceived low quality of AI-generated artwork, and copyright infringement. Notably, our more fine-grained analysis suggests that most concerns address the AI stance communicated by Lofter's policy decision rather than specific tool features. Heterogeneity analysis shows that higher-popularity, multi-homing, and AI-averse creators exhibit larger reductions in activity on Lofter, whereas on Graffiti Kingdom, the differences across popularity and tenure levels are not statistically significant.

The treatment effects we estimate capture both direct policy impacts and stance signaling effects, and we argue the signaling channel is the dominant driver. For Lofter, the AI tool was removed within one week of its launch, yet the negative effects persist into subsequent months, which cannot reflect continued mechanical access to the tool and most plausibly reflects the pro-AI stance signaled by Lofter's AI-tool launch. For Graffiti Kingdom, enforcement was minimal (fewer than 0.05% of users restricted), and the weekly platform-level data show no dip-and-rebound pattern that would be expected if reduced competition were driving the effect. The positive response is therefore most plausibly attributable to the signaling content of the anti-AI stance communicated by Graffiti Kingdom's prohibition.

Several pieces of additional evidence point to signaling effects as the primary driver of creator responses. First, analysis of creator posts reveals that only 7% of concerns address the AI tool directly, while roughly 50% address Lofter's policy decision or the AI stance it communicated, confirming that creators respond primarily to stance signals rather than specific tool features. Second, creators who had previously expressed negative attitudes toward AI exhibit larger activity reductions, suggesting that those most attuned to stance signals are most responsive. Together, these results demonstrate that creators on these visual arts platforms respond to AI-policy stance signals—favoring anti-AI policy signals over pro-AI policy signals.

This study makes several contributions. First, to our knowledge, this is the first paper to provide causal evidence on how the AI stances signaled by platform policy decisions shape content creator behavior. We show that creators respond primarily to stance signals rather than to specific AI tools: the negative effects on Lofter persisted even after the AI tool was removed, and analysis of creator posts reveals that the negative reactions mainly target Lofter's policy decision or the AI stance it communicated. Second, we provide the first systematic analysis of why creators on visual arts platforms resist Generative AI. Through LLM-assisted content analysis, we identify three primary concerns, which offer concrete targets for platform policy design. Third, our heterogeneity analysis reveals that the creators most likely to reduce activity following Lofter's pro-AI policy signal are precisely those platforms can least afford to lose: higher-popularity creators, and multi-homing creators with lower switching costs. Meanwhile, longer-tenure creators show smaller

reductions, and creators who previously expressed negative AI sentiment exhibit larger reductions. This suggests that platforms can use sentiment monitoring to identify at-risk segments before implementing AI policies. Finally, by examining two policy decisions that communicate opposite AI stances (see Online Appendix Table A.1), we find consistent evidence and provide a more comprehensive picture of how creators respond to different AI-policy stance signals.

The rest of the paper is organized as follows: Section 2 reviews related literature and explains our contribution; Section 3 introduces our research contexts and data; Section 4 presents our main empirical analysis using the TCI framework, including identification strategy, estimation results, and mechanism analysis; Section 5 explores heterogeneous effects; Section 6 provides business and policy suggestions and concludes.

2 Literature Review & Contribution

A growing literature examines the economic impacts of AI across business contexts, including chatbot disclosure, algorithmic pricing, and AI-assisted service (Luo et al. 2019, Zhang et al. 2021, Kim et al. 2022), along with the factors shaping AI adoption (Gray 2017, Castelo et al. 2019, Castelo and Ward 2016, Yu et al. 2018). With the emergence of Generative AI, researchers have begun examining its impact on individual productivity (Peng et al. 2023, Noy and Zhang 2023, Shin et al. 2024), creativity (Shin et al. 2023, Zhou and Lee 2024), and workforce dynamics (Eloundou et al. 2023), as well as its potential risks to the content-creator and broader economy (Spitale et al. 2023, Burtch et al. 2023).

Our research connects to two streams of literature. First, building on the literature examining the economic impacts of (Generative) AI, we engage with the literature on platform governance, including the emerging work on AI governance (Hanisch et al. 2023, Kathuria et al. 2026, Costabile 2024), which highlights that the central question is not whether Generative AI exists as a technology, but how platforms govern its presence and shape its use. Existing work in this area largely studies how specific AI-related policies or technologies affect users and creators: Wittenberg et al. (2025) evaluates AI-content labeling strategies; Lloyd et al. (2025) examines how Reddit communities respond to AI-generated content; and Goldberg and Lam (2025) models how Generative AI adoption shifts the market equilibrium of content creation. By contrast, we compare two analogous content-creation platforms that experienced focal policy decisions communicating contrasting AI stances, and trace how the corresponding policy signals shape creator behavior. Our contribution thereby connects to the broader literature on institutional signaling, which examines how visible organizational choices update stakeholder beliefs and shape subsequent behavior (Spence 1973, Connelly et al. 2011).

The second strand pertains to the creator economy, digital labor supply, and the design of content creation platforms. Prior work examines content creators' intrinsic and extrinsic motivations (Toubia and Stephen 2013, Sun and Zhu 2013, Liu and Feng 2021) and emphasizes the central role of platforms in structuring creator incentives and opportunities (Bleier et al. 2024, Peres et al. 2024, Khobzi et al. 2025). Studies of platform design levers further show how monetary incentives shift output volume and composition (Chen et al. 2008, Sun and Zhu 2013), and how gamification and reputation systems steer effort and engagement (Anderson et al. 2013). We add to this literature by showing how creators adjust effort, participation,

attachment, and activity when platform policy decisions communicate visible stances toward Generative AI under uncertainty about the future returns to human creative work.

3 Context and Data

3.1 Platforms Overview and Research Design

Our empirical analysis leverages two separate natural experiments on two Chinese platforms: Lofter and Graffiti Kingdom. Each policy event provides distinct causal evidence about creator responses to the stance signaled by a focal AI policy: one signaling a pro-AI stance and the other an anti-AI stance. We analyze each platform separately using an identical empirical framework. Since the two focal policy decisions communicate stances at opposite ends of the pro-AI–to–anti-AI spectrum (see Online Appendix Table A.1), this parallel design enhances external validity: converging findings across contrasting policy contexts strengthen the generalizability of our results to platforms adopting AI-policy signals between these extremes.

Lofter, founded by NetEase in 2011, is one of China's leading content creation platforms for visual arts and fan fiction. The platform spans interests including gaming, anime, photography, painting, design, and writing, with over 80 million interest tags and 12.8 million content creators.[3] Graffiti Kingdom, with almost 20 years of history, is a platform renowned for its original, high-quality works in various styles. It serves as a hub for illustrators, with over 500,000 registered users and approximately 15,000 daily visits.[4] Figure 1 shows screenshots of the platform interfaces, highlighting where user-generated works appear on Lofter and Graffiti Kingdom.

Both Lofter and Graffiti Kingdom are visual arts platforms that emphasize original, human-created art. Both prohibit reposted or copied materials unless properly labeled. On Lofter, users must choose one of six Creative Commons licenses at publication.[5] Graffiti Kingdom specifies that "the copyright and interpretation of original content published and displayed by Graffiti Kingdom members belong to the original author".[6] On both platforms, Generative AI models can produce similar works at near-zero marginal cost and rapid turnaround.

From the perspective of where value primarily attaches, content creation platforms span a continuum from *output-centric*, where attention and economic value attach mainly to discrete works that can be searched, displayed, and transacted as relatively standalone products, to *creator-centric*, where value attaches more strongly to the creator's persona, ongoing audience relationship, and repeated interactions with followers, so that any single piece of content is less separable from the creator. Visual arts portfolio platforms such as DeviantArt and ArtStation tend to be more output-centric, while short-form video and social platforms such as TikTok and YouTube tend to be more creator-centric. Since most creators on Lofter and Graffiti Kingdom are not widely recognized, authorship usually carries little premium relative to the work itself, and both platforms sit closer to the output-centric end of this continuum.

[3] See Apple App Store: Lofter. Last accessed July 22, 2026.
[4] See Baidu Baike: Graffiti Kingdom. Last accessed July 22, 2026.
[5] See Wikipedia: LOFTER. Last accessed July 22, 2026.
[6] See Graffiti Kingdom Terms of Service. Last accessed July 22, 2026.

Engagement-oriented measures support this placement. Although we cannot construct a complete repeat-commenter network, we observe follower counts and supporter counts, which proxy for the extent to which audience engagement organizes around stable, creator-centric communities. On Lofter, the median creator has only 9 followers, the 75th percentile 69, and the 95th percentile 1,181. Supporters, defined as users who have financially rewarded a creator, are similarly sparse: the median is 0, the 75th percentile 1, and the 95th percentile 60. On Graffiti Kingdom, follower data are incomplete, but even among the roughly 6,000 most popular creators (those whose works were ever selected into the platform's "Most Popular" group), the median is only 18, the 75th percentile 49, and the 95th percentile 590. For scale, the major creator-centric monetization programs set eligibility well above these levels: YouTube's Partner Program requires 1,000 subscribers and TikTok's Creator Rewards Program at least 10,000.[7] Only the top few percent of Lofter creators reach even the lower of these bars, and the typical creator on both platforms falls far below them. Creator-centric attachment thus exists on Lofter and Graffiti Kingdom, but it is thin and concentrated in the upper tail rather than the dominant mode of participation. Accordingly, we characterize the two platforms as comparatively more output-centric than creator-centric, rather than purely output-centric.

Both platforms rely primarily on third-party advertising and on a fraction of creator tips, so platform success is tied to creator activity. Creators are motivated by recognition, audience appreciation, and modest monetary gain through tips and off-platform opportunities. The typical user is much closer to a hobbyist or intrinsically motivated creator than to a professional pursuing income or employment: direct monetary ties are limited, activity is generally low and intermittent, and commission-related language (e.g., "Accepting Commissions" / "Open for Assignments") appears in only about 0.7% of Lofter posts.

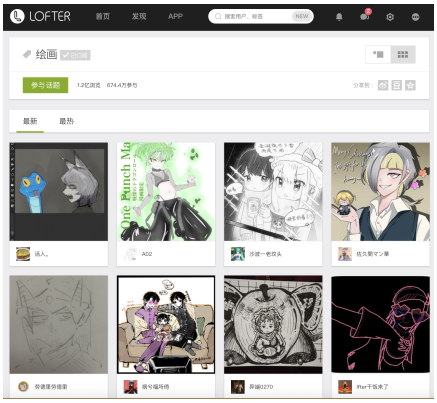
The platforms differ in two respects. First, Lofter supports multiple content formats (painting, writing, photography), while Graffiti Kingdom caters exclusively to illustrators. Second, Lofter's user base is substantially larger due to backing from NetEase, a major Chinese internet company. For our analysis, we restrict Lofter data to the *painting* community, since the Generative AI tool they introduced primarily impacts visual artists. This restriction ensures that the creator populations on both platforms face similar exposure to AI-generated competition. The next two subsections outline the natural experiments and data collection in detail.

3.2 Natural Experiments

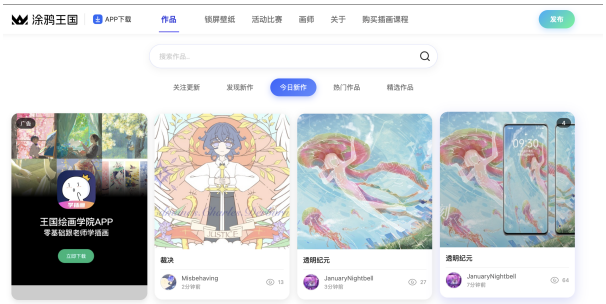
On March 6th, 2023, Lofter launched a Generative AI image generator marketed as the "Profile Picture Generator" (also called the "Lofter Drawing Machine"). Despite its name, this is not merely an avatar tool. Instead, it is a full-featured AI image generator that allows users to generate and download images from text prompts, similar to Midjourney. The tool's functionality was widely documented in contemporary news reports, and its release generated substantial attention and controversy both in the broader public sphere and among users on the platform.[8] Lofter's development of this Generative AI feature represents active platform investment in AI creation tools, signaling a pro-AI stance. The feature was removed within one week after its launch, but its introduction is the treatment we analyze. Lofter's decision to develop this tool was likely

[7] See TikTok's Creator Rewards Program and YouTube's Partner Program. Last accessed July 22, 2026.

[8] See Jiemian News, Sina Finance, Zhejiang News, and Pandaily coverage from March 2023. Last accessed July 22, 2026.



(a) Lofter



(b) Graffiti Kingdom

Figure 1: Screenshots of Lofter and Graffiti Kingdom

influenced by the growing popularity of Generative AI and the success of platforms like TikTok and Red Note in integrating similar features.

On February 22nd, 2023, Graffiti Kingdom announced that AI-generated artwork was prohibited on the platform, becoming the first content creation platform in China to officially ban AI-generated content. The policy was stringent: all previously uploaded works would be reviewed, and users who violated this policy would be banned from their accounts.[9] This prohibition signals an anti-AI stance. Graffiti Kingdom's decision may have been influenced by prior protests on platforms like ArtStation.[10]

Online Appendix Table A.3 reconstructs the publicly documented AI-related actions on both platforms from July 2022 through June 2023. The chronology shows that each platform's subsequent actions remained within its focal policy episode: Lofter's clarifications, use restrictions, withdrawal, and creator-protection announcements were reactions to the launch and the ensuing backlash, while Graffiti Kingdom simply maintained and enforced its prohibition. Our search did not identify evidence that either platform initiated another AI-related policy during this window.

These two focal policy decisions anchor opposite ends of the AI-policy stance spectrum, as shown in Online Appendix Table A.1. Other visual arts platforms adopt intermediate AI-policy positions (e.g., permissive with labeling, restrictive with constraints). We analyze each platform's natural experiment separately, and consistent patterns across both would strengthen external validity and extend our conclusions to intermediate AI-policy signals. The two treatments also occurred at broadly similar points in time, which further improves comparability by making the broader external environment outside the platforms largely similar across the two settings.

3.3 Data Collection and Variables

For Lofter, we focus on the painting community, since the Generative AI image generator primarily affects visual artists. Our final Lofter dataset includes 197,771 artists and 10,240,721 artwork posts, covering all

[9] According to an official staff member of the platform, however, banned users' accounts and any of their previously uploaded works that were not generated by AI would be preserved.

[10] See Vice: Artists are in revolt against AI art on ArtStation. Last accessed July 22, 2026.

the creators who were active between January 2020 and December 2023. We additionally collect the full text of 1,518,242 non-work posts, which are text-based posts similar to tweets where creators discuss their thoughts, feelings, and opinions; we use these for the mechanism analysis in Section 4.3. Throughout the paper we refer to artwork posts as *works* and non-work posts as *posts*.

For Graffiti Kingdom, our final dataset covers all creators active from January 2019 to December 2023, comprising 42,418 artists and 711,381 artwork posts. Graffiti Kingdom does not have a non-work post feature analogous to Lofter's, so such content is not available on this platform. According to an official Graffiti Kingdom staff member, although users who violated the AI policy could be banned from their accounts, neither the accounts themselves nor any previously uploaded non-AI-generated content were deleted. As a result, the policy did not mechanically remove users or their historical human-created works from our sample, which helps rule out the concern that the estimated effects are driven by policy-induced sample loss rather than by changes in creators' subsequent behavior. Detailed data-collection methodology for both platforms appears in Online Appendix B.

We aggregate the data at the creator-month level. Our focal outcome is *Num of Works*, the number of works uploaded by creator i in month t. We also construct *Active*, a binary indicator equal to one if creator i uploads at least one work in month t, which we use in robustness checks. We also measure creator characteristics using *like_avg* (lifetime average likes per work) and *review_avg* (lifetime average comments per work). These time-invariant metrics proxy for creator popularity and audience engagement on these two-sided platforms, as well as for potential financial returns through on-platform tips and off-platform opportunities (Li et al. 2021).

We use a monthly frequency because content creators need time to produce their work; even at this frequency, creators upload fewer than two works on average on both platforms, and a finer frequency would substantially increase zero inflation in the outcome variable (Lambert 1992).[11] To ensure consistent comparison periods, we standardize each "month" as a 28-day period (four weeks) rather than using calendar months of varying lengths.

Since the analytical sample depends on the cohort-matching scheme introduced in Section 4, descriptive statistics for that sample appear after the empirical specification. Full-sample descriptive statistics are provided in Online Appendix B.

4 Main Analysis: Temporal Causal Inference

In this section, we present our main empirical analysis of how the AI stances signaled by platform policy decisions causally affect content creators' behavior. We begin by outlining the key identification challenge and our empirical strategy. We then validate our approach through assumption tests before presenting estimation results. Finally, we analyze creator-generated content to understand the mechanisms driving their responses.

[11] Robustness checks confirm that our results are not driven by a small number of highly productive contributors; see the super-user exclusion in Online Appendix G (Table G.7).

4.1 Empirical Strategy

Both policy changes were implemented across the entire platform simultaneously, which eliminates untreated control groups required under standard causal inference frameworks. To tackle this challenge, we apply the Temporal Causal Inference (TCI) framework proposed by Turjeman and Feinberg (2023) to construct control groups.[12] The key idea of TCI is to shift the analysis from natural calendar time to platform tenure time. While all content creators experienced the policy changes on the same date, they were at different points in their platform tenure (i.e., the number of months since joining the platform) when these changes occurred. This allows us to compare content creators who joined later and were exposed to the treatment (the "treatment group") with those who joined earlier and were not exposed to the treatment when they were at similar tenures (the "control group"). The framework builds on the assumption that while users joined the platform on different dates, they should exhibit similar behavior patterns at the same tenure points. It is important to note that this strategy is implemented within each platform separately: each platform has its own treated and control cohorts, and we do not pool the two samples or rely on cross-platform comparability for identification.

We construct the control groups as follows. First, we divide creators into cohorts based on their first posting time on each platform: cohort i contains creators whose first post was i months before the treatment.[13] We retain cohorts 8–55 on Lofter and 7–42 on Graffiti Kingdom, which together cover more than 75% of creators on each platform. [14] By construction, the treated cohorts used in the TCI analysis consist of creators who had joined the focal platform before the corresponding treatments; the resulting treatment effects therefore capture behavioral responses among pre-policy creators, including possible reactivation, rather than entry by creators who joined after the policy. For each treated cohort, following Turjeman and Feinberg (2023), the control group consists of cohorts that joined 3–7 months earlier;[15] only treatment-control pairs that pass both the Kolmogorov–Smirnov (KS) and Granger causality tests are retained for the subsequent analysis (see Online Appendix D).

Figure 2 illustrates the matching structure. For a treated cohort i (creators who first posted i months before treatment), we pool five control cohorts $i+3, i+4, \ldots, i+7$. The treated cohort contributes i pre-treatment periods, indexed from $-(i-1)$ through 0, and three post-treatment periods, indexed as 1, 2, and 3. Month 0 is the last pre-treatment period; the treatment occurs at month 1. For each control cohort, we retain a tenure-matched observation window of $i+3$ periods that occurs entirely before the treatment. For example, when cohort 10 is the treatment group, the observation window spans months $-9, -8, \ldots, 0, 1, 2, 3$ (13 periods); from each of cohorts 13 through 17, we retain a 13-period window matched in tenure (from cohort 13, months -12 through 0; from cohort 14, months -13 through -1; and so on, until cohort 17, from

[12] Robustness checks using alternative identification strategies that do not rely on the TCI framework (Regression Discontinuity in Time and a Panel Difference Estimator) are included in Online Appendix G; both yield results consistent with our main findings.

[13] We use first posting time rather than joining time because we do not observe joining dates and because first posting marks the start of active participation as a creator.

[14] We exclude cohorts joining fewer than 8 months (Lofter) or 7 months (Graffiti Kingdom) before treatment to ensure sufficient pre-treatment observations. The 48 retained Lofter cohorts and 36 retained Graffiti Kingdom cohorts together capture more than 75% of creators on each platform, while the remaining 100+ earlier cohorts cover less than 25% of creators and are excluded.

[15] Robustness checks using alternative control-cohort constructions, including a nearby alternative specification with cohorts $i+6$ to $i+10$, a no-common-period specification with cohorts $i+11$ to $i+15$, and a narrower sample restricted to the 24 cohorts from the two years prior to treatment, appear in Online Appendix G (Table G.7).

which we retain months −16 through −4). Each control window therefore matches the treated cohort exactly in tenure profile while occurring entirely before the treatment, and each user appears in only one cohort. The first two columns of Online Appendix Tables C.1 and C.2 list the full set of cohort pairs.

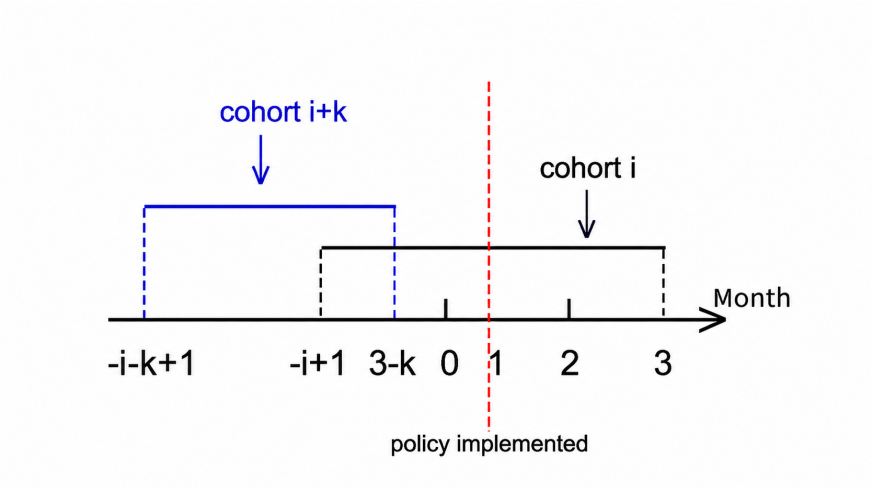


Figure 2: Cohort Matching in TCI

With the treatment-control pairs formed, we then estimate the treatment effects using standard causal inference frameworks. Following Turjeman and Feinberg (2023), we use the Difference-in-Differences (DiD) model (Angrist and Pischke 2009, Lechner et al. 2011) as our primary estimation method. Specifically, when cohort j is the treatment group, we estimate the time-varying treatment effects in the three post-treatment months using the following specification:

$$Y_{it} = \mu_i + \lambda_t + \sum_{m=1}^{3} \tau_{jm} Treat_i \times PostMonth_{mt} + \epsilon_{it}. \tag{1}$$

In this equation, the subscript i denotes content creator i, and the subscript t refers to the months since content creator i posted her first work (i.e., tenure) on the platform. $PostMonth_{mt} = 1$ if tenure t for the treatment group (i.e., cohort j) falls within post-treatment month $m \in \{1, 2, 3\}$ (or the corresponding month in the untreated comparison window for the control group), and 0 otherwise. The treatment indicator $Treat_i = 1$ if content creator i is in the treatment group, and 0 otherwise. μ_i is the content creator fixed effect, λ_t represents the tenure fixed effect, and ϵ_{it} is the error term. For the dependent variable Y_{it}, we examine the log of the number of works uploaded by content creator i to the platform during tenure month t. We add one to the number of works before taking the logarithm, as some observations are zero.[16]

Equation (1) is our baseline time-varying specification. We additionally estimate two related variants: a pooled DiD that replaces the three post-month indicators with a single $Post$ indicator, yielding the canonical

[16]In Online Appendix G, we address concerns about the $\log(Y + 1)$ "log-like" transformation (Chen and Roth 2024) using a binary outcome variable, Poisson pseudo-maximum likelihood (PPML) estimation, and alternative offset constants in the log transformation.

ATT; and an event study with full lead-lag indicators for every period, allowing us to inspect pre-treatment dynamics.

After separately estimating $\hat{\tau}_{jm}$ from each treatment-control pair j, we aggregate across cohorts following Turjeman and Feinberg (2023): $\bar{\tau}_{m,\text{treated}} = \sum_j \left(\text{num}_j / \sum_i \text{num}_i\right) \hat{\tau}_{jm}$, with standard deviation $\text{Std}\left(\bar{\tau}_{m,\text{treated}}\right) = \left\{\sum_j \left(\text{num}_j / \sum_i \text{num}_i\right)^2 \left[(\hat{\tau}_{jm} - \bar{\tau}_{m,\text{treated}})^2 + \sigma^2_{\tau_{jm}}\right]\right\}^{1/2}$, where $\sigma_{\tau_{jm}}$ is the standard error of $\hat{\tau}_{jm}$ and num_j denotes the size of cohort j. The same aggregation procedure applies to the pooled DiD and event-study variants.

We validate the TCI framework through three identification-assumption checks: Kolmogorov–Smirnov tests for distributional similarity between treated and control cohorts, Granger causality tests for SUTVA violations, and parallel-trend tests via the event-study form. All cohorts pass the KS test; the small number that fail the Granger test are dropped from the analysis with minimal impact on sample size. Detailed discussion appears in Online Appendix D. Table 1 reports descriptive statistics for the final estimation sample.

Table 1: Summary Statistics of Key Variables in the Final Estimation Sample

Variable	**Definition**	**N**	**Mean**	**Std.Dev.**	**Min**	**Max**
Panel A: Lofter						
Num of Works	Number of works uploaded during the month	5,245,026	1.572	19.448	0	14,726
Active	Equals 1 if creator uploads ≥1 work	5,245,026	0.049	0.217	0	1
review_avg	Average number of reviews received	5,245,026	0.863	7.857	0	2,682
like_avg	Average number of likes received	5,245,026	57.555	983.143	0	534,697
Panel B: Graffiti Kingdom						
Num of Works	Number of works uploaded during the month	791,737	0.939	4.328	0	511
Active	Equals 1 if creator uploads ≥1 work	791,737	0.143	0.350	0	1
review_avg	Average number of reviews received	791,737	0.026	0.415	0	167
like_avg	Average number of likes received	791,737	1.417	7.277	0	1,532

Note. The unit of analysis is the creator-month. The table reports descriptive statistics for the final Temporal Causal Inference estimation sample; *N* denotes the number of creator-month observations.

4.2 Estimation Results

With the control group constructed under the TCI framework, we use Difference-in-Differences (DiD) to estimate the treatment effects. Table 2 presents the estimated treatment effects for both Lofter and Graffiti Kingdom, reporting both the canonical ATT and the time-varying effects by post-treatment month.[17]

For Lofter, we observe a consistent negative impact on the number of works across all post-treatment periods (columns 1–2 of Table 2). The effect grows in magnitude from Month 1 to Month 2 and remains at a similar level in Month 3. In Online Appendix F, we extend the post-treatment horizon to six months and continue to find significant negative effects, indicating that the impact persists. For Graffiti Kingdom, we observe positive yet shrinking treatment effects across all post-treatment periods (columns 3-4).

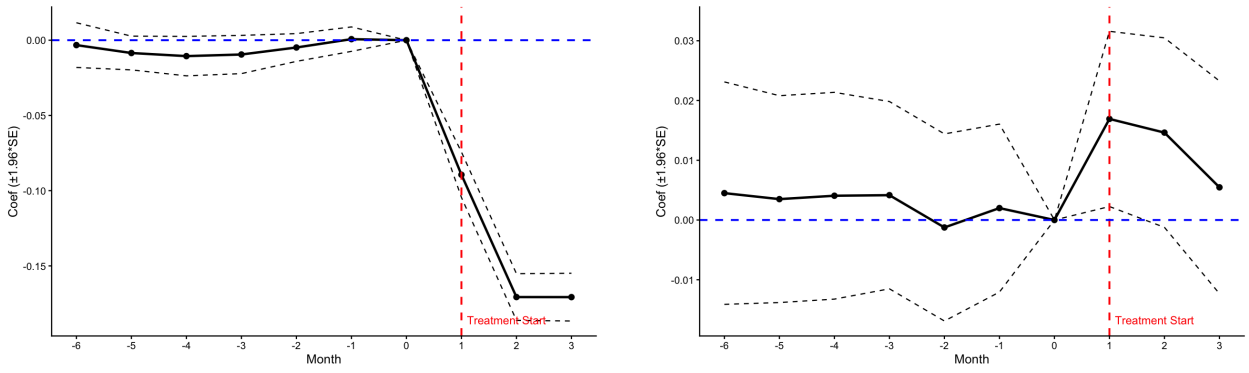
[17] The DiD estimates are corroborated by Generalized Synthetic Control as an alternative TCI estimator; see Online Appendix G (Table G.1).

Table 2: Temporal Causal Inference Results Using Difference-in-Differences

	Lofter		Graffiti Kingdom	
	(1)	(2)	(3)	(4)
Dependent Variable	log(Num of Works + 1)			
Treat × Post	-0.0936***		0.0224***	
	[0.0056]		[0.0041]	
Treat × Post Month 1		-0.0520***		0.0267***
		[0.0063]		[0.0050]
Treat × Post Month 2		-0.1154***		0.0254***
		[0.0058]		[0.0050]
Treat × Post Month 3		-0.1134***		0.0150***
		[0.0056]		[0.0046]
Estimation Method	Difference-in-Differences			
Creator Fixed Effects	Yes	Yes	Yes	Yes
Tenure Fixed Effects	Yes	Yes	Yes	Yes
Number of Creators	140,061	140,061	24,178	24,178

Note. The unit of analysis is the creator-month. Post Month 1–3 denote the first three post-treatment periods. Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Figure 3 visualizes the event-study coefficients; the corresponding coefficient table appears in Online Appendix Table D.1. The pre-treatment coefficients are small and not statistically distinguishable from zero, supporting the parallel-trends assumption underlying the DiD specification. The post-treatment estimates are consistent with the time-varying pattern shown in Table 2: negative for Lofter and positive but shrinking for Graffiti Kingdom. Because it estimates a separate coefficient per tenure month, the event study is more flexible but less precise than the main specification: Lofter's coefficients remain significant throughout, while Graffiti Kingdom's coefficients become insignificant in the later periods. We therefore interpret the effect of Graffiti Kingdom's policy as positive, at least in the short run.



(a) Lofter (b) Graffiti Kingdom

Figure 3: Event Study: Treatment Effects by Tenure Month

Note: Month 0 is the omitted reference period and the last pre-treatment period; Month 1 is the first post-treatment period.

The estimated treatment effects on each platform combine two components: a direct policy effect (the

mechanical impact of the policy itself on creators) and a signaling effect (the impact of the AI stance communicated by the policy). We identify the signaling component on each platform by combining sign restrictions on the direct effect with the timing of the policy.

On Lofter, introducing a new tool should likely have a non-negative direct effect: under free disposal, creators are weakly better off when an additional option becomes available. The observed negative overall effect therefore implies a negative signaling effect. One could argue that the direct effect could itself be negative if creators react adversely to the tool's existence (e.g., concerns that others will flood the feed with AI-generated content, or that their works will be scraped for training). However, Lofter retracted the tool within one week of launch, so any direct effect of the tool's availability should not operate beyond the first post-treatment month. The continued negative effects in Months 2 and 3 (when the tool was already gone) therefore most plausibly reflect signaling rather than the direct effect. The retraction was part of Lofter's effort to contain the backlash: within the same month, the platform also issued public apologies and announced prospective creator-protection measures (see Online Appendix Table A.3). We read these follow-up actions as remedial efforts rather than as a change in stance; none of them replaced the launch with an anti-AI governance policy, and we found no independently verifiable evidence that the announced measures were substantively implemented during the three post-treatment months we analyze.

On Graffiti Kingdom, the direct policy effect is plausibly non-positive: the prohibition imposes a negative direct effect on the small fraction of violators but does not directly alter the creative environment for the vast majority of non-violating creators. The observed positive overall effect therefore implies a positive signaling effect. In principle, violator removal could reduce competition for non-violators and therefore produce a positive direct effect; however, the data do not support this channel: only about 100 users (under 0.05% of the user base) were restricted, and the weekly platform-level upload data show no dip-and-rebound pattern that a reduced-competition channel would predict. The positive response is therefore most plausibly attributable to the signaling content of the anti-AI stance communicated by Graffiti Kingdom's prohibition.

Together, the two policy events point to a consistent stance-signaling story: creator activity responds to the signaling content of a platform's AI-related policy, declining after a pro-AI policy signal and rising after an anti-AI policy signal.

Finally, we also address two additional concerns. First, because the two treatments partially overlap in time, user multi-homing across Lofter and Graffiti Kingdom could confound our estimates. We analyze all Lofter posts to detect self-reported accounts on other platforms. Only 96 (around 0.05%) Lofter users ever disclosed a Graffiti Kingdom account, compared to 16,868 who disclosed a Weibo account. This minimal cross-posting suggests that multi-homing between our two focal platforms poses no substantive identification threat. Second, platforms might strategically time policies based on pre-existing creator activity trends, violating exogeneity. We estimate a hazard regression testing whether pre-event levels, trends, and volatility in key creator metrics predict policy adoption timing (Kooperberg et al. 1995). All estimated coefficients are statistically insignificant (Table E.1, Online Appendix E), supporting that the policy timing is plausibly exogenous.

We further confirm our findings through robustness checks targeting four distinct concerns. (i) The results do not depend on the TCI framework: alternative identification strategies that do not rely on TCI

(Regression Discontinuity in Time and a Panel Difference Estimator) yield consistent estimates. (ii) The choice of estimator within TCI does not drive the results: Generalized Synthetic Control yields estimates consistent with our main DiD specification. (iii) Concerns about the $\log(Y+1)$ transformation are addressed using a binary activity indicator, Poisson pseudo-maximum likelihood, and alternative log-offset constants. (iv) The results are not driven by extreme contributors or particular cohort choices: we obtain consistent estimates after excluding super-users (top 1%), using an alternative nearby control window ($i+6$ to $i+10$) and a no-common-period control window ($i+11$ to $i+15$), and restricting the sample to the 24 most recent cohorts. Full details appear in Online Appendix G.

4.3 Understanding the Mechanisms

A natural theoretical interpretation of the signaling channel is that the stance communicated by a platform's focal AI policy carries information for creators about the future place of human creative labor on the platform. This signaling-based reading draws on institutional signaling theory (Spence 1973, Connelly et al. 2011), platform psychological contracts (Rousseau 1989, She et al. 2020), and evidence that AI-related information shifts beliefs and that creators respond to platform governance signals (Menkhoff 2025, Duffy and Meisner 2023). Under this lens, the pro-AI stance signaled by Lofter's AI-tool launch may have reduced activity because creators read it as a signal that the future value of their effort would decline, that the platform was less committed to creators like them, or that effort would be better invested elsewhere; the anti-AI stance signaled by Graffiti Kingdom's prohibition may have increased activity because it communicated stronger support for human creators and a more protected, less AI-dominated competitive environment for human-created work, which may have encouraged pre-policy creators to reallocate effort to, or reactivate on, the platform. We interpret this belief-mediated response as one pathway through which the stance signal operates rather than as a separate direct policy effect. This is broadly consistent with recent work showing that Generative AI can increase individual creative productivity while raising concerns about substitution and reducing collective diversity (Zhou and Lee 2024, Doshi and Hauser 2024, Lee and Chung 2024, Holzner et al. 2025). We stress, however, that these are plausible interpretive mechanisms rather than separately identified causal channels in our data.

Although we cannot cleanly identify the separate mechanisms through which signaling operates, Lofter's non-work posts (text-based posts where creators share thoughts and opinions) offer a useful window into how creators interpreted the pro-AI stance signaled by Lofter's AI-tool launch. To better understand the mechanisms driving creator responses and further validate our signaling interpretation, we analyze these posts expressing attitudes toward Generative AI.

We identify all posts containing the keyword "AI" and related terms (e.g., "Artificial Intelligence," "Generative AI"). A key challenge is that some posts mention these terms without expressing opinions about Generative AI (e.g., short improvised science-fiction vignettes or micro-stories). We therefore use an LLM (gpt-4o-mini) to separate posts expressing attitudes about Generative AI from those that do not, and classify the relevant posts by valence (anti-AI, pro-AI, neutral, mixed); about 75% express negative attitudes. We further use the LLM to identify the broad reasons creators give for resisting Generative AI and conduct a deeper analysis within the top three categories to recover more granular motives. Online Appendix H details

the full procedure.

To validate the LLM coding, we recruit human annotators to code three subsamples of posts and derive gold labels via majority voting. The LLM achieves high accuracy: 0.94 for attitude detection, 0.92 for valence classification, and 0.81 for primary-reason identification; these results are consistent with recent evidence that LLMs can approximate human judgments on perceptual and attitudinal constructs (Arora et al. 2025, Li et al. 2024, Goli and Singh 2024). Full procedure in Online Appendix H.

Table 3 reports the three most common broad reasons for resisting AI and their prevalence in our sample. The leading concern is risk of replacement (26.19%), followed by perceived low quality of AI-generated artwork (24.65%) and copyright infringement (18.90%). The remaining reasons are sparse, and even the most common among them has a prevalence below 2%.

These findings help explain the decrease in creator activity following Lofter's pro-AI rollout. The three concerns interlock into a coherent picture: creators see their own works potentially being used without consent (copyright) to train AI systems that produce lower-quality outputs (low quality) which nonetheless may substitute for human-made work in the marketplace (replacement). Beyond these material concerns, creators are highly attuned to the distinctions between human-made and AI-made works (process, effort, skill, authorship, and traceable provenance) and perceive AI outputs as lacking these hallmarks. They therefore interpret focal policy signals that blur these distinctions as threats to artistic identity, reputation, and fair attribution; a pro-AI rollout sends precisely this kind of signal. Together, these perceptions raise the expected costs of participation (policing infringement, defending authorship) and lower the expected returns (visibility, attribution, recognition for craft), making reductions in posting a rational response. Figure 4 further decomposes each of the three broad categories into more granular concerns.

Table 3: Top Three Reasons for Resisting Generative AI

Reason	Count	Prevalence (%)	Example (Translated from Chinese)
Replacement Risk	392	26.19	After seeing what AI can do in art, I'm worried an art major might not have a future.
Low Quality	369	24.65	The nonsensical brushstrokes and messy structures give me a headache. Is this what shortcut-taking artists in the AI era look like?
Copyright Infringement	283	18.90	Any organization or individual who uses any content I post (including but not limited to articles, images, comments, etc.) as AI training data will be considered infringing.

Note. The unit of analysis is the Lofter non-work post. Counts and prevalence are computed among 1,497 posts classified as expressing negative attitudes toward Generative AI; rows report the three most common primary reasons.

In a separate analysis, we restrict the sample to posts expressing negative attitudes during the period in March 2023 slightly before and after the Generative AI feature was removed, and classify these posts according to what they primarily target using rule-based text analysis. Specifically, we first identify *Direct Feature-Related* posts by searching for explicit references to the AI image generator itself, including its name (i.e., "Profile Picture Generator") and closely related expressions such as "this feature." Among the remaining posts, we classify as *Platform AI-Related* those that mention the platform by name (i.e., Lofter)

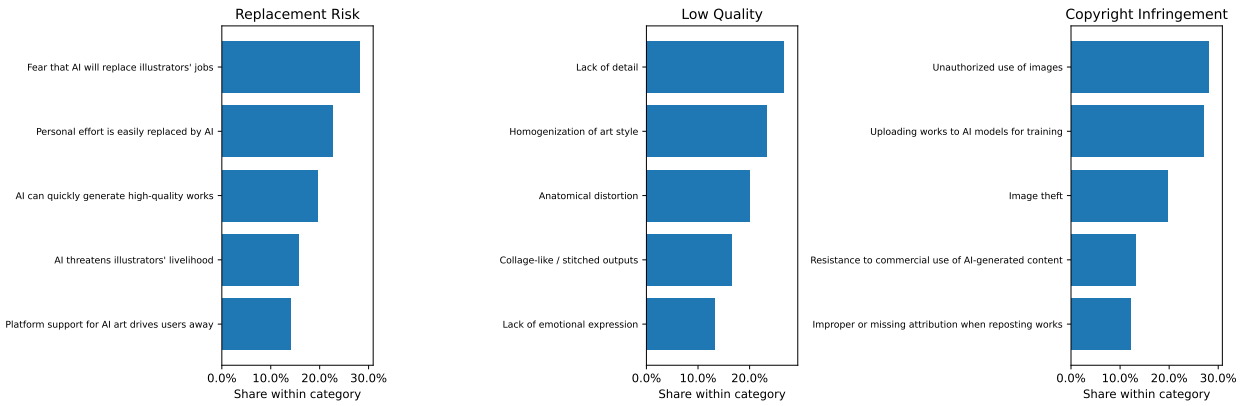


Figure 4: Detailed Reasons for Resisting Generative AI (Within the Top Three Broad Categories)

or through similar expressions such as "this platform." All remaining posts are classified as *Generative AI-Related or Other*. Because this categorization is based on transparent keyword and context rules rather than latent semantic judgment, we view it as a relatively objective descriptive decomposition. Table 4 reports the results: only about 7% of posts target the Generative AI feature itself, roughly half target Lofter's decision or the pro-AI stance it communicated, and the remaining 43% express broader anxieties about Generative AI not specific to either. These results reinforce our earlier conclusion that it is Lofter's decision and the stance it communicated, rather than the tool itself, that primarily drives the reductions in creator activity.

Table 4: Decomposition of Posts in March, 2023

Category	Count	Prevalence (%)	Example (Translated from Chinese)
Direct Feature-Related	27	7.18	I was alarmed by Lofter's AI drawing feature.
Platform AI–Related	187	49.73	I recently learned that Lofter permits AI-generated art, and I have decided to stop posting images on the platform.
General AI-Related or Other	162	43.09	It's everyone's responsibility to resist AI.

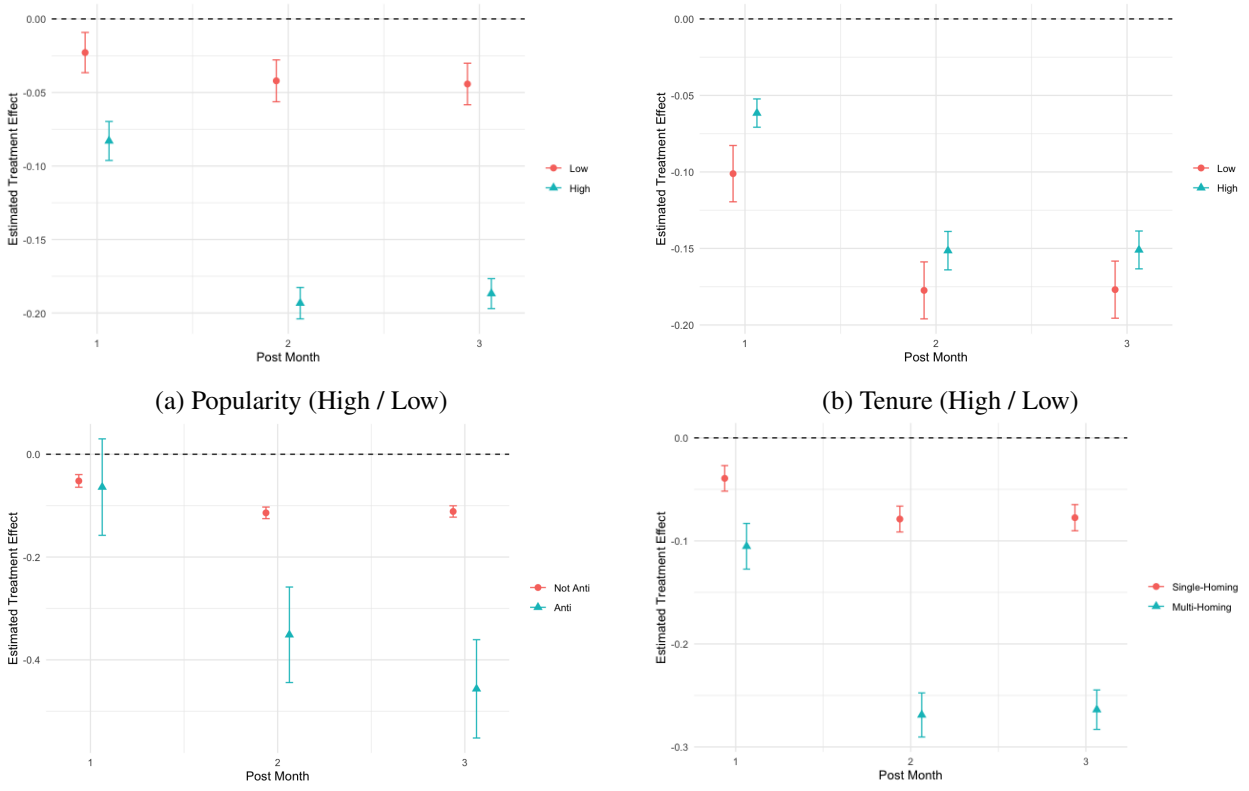
Note. The unit of analysis is the Lofter non-work post. Counts and prevalence are computed among 376 posts expressing negative attitudes toward Generative AI in March 2023 slightly before and after the feature was removed; categories are mutually exclusive.

5 Heterogeneity Analysis

Our main results indicate that the pro-AI stance signaled by Lofter's AI-tool launch generates negative signaling effects, whereas the anti-AI stance signaled by Graffiti Kingdom's prohibition can generate positive signaling effects, at least in the short run (see Table 2). Understanding which creator segments are most responsive to these stance signals is critical for platform AI policy design. In this section, we focus primarily on Lofter to examine how different types of creators respond to the AI stance signal communicated by the focal policy. We assess heterogeneity along four dimensions: (i) popularity, (ii) tenure, (iii) multi-homing status, and (iv) expressed AI sentiment.

We measure popularity as the average likes per work during the pre-treatment period, and infer multi-homing by examining posts and profile bios for mentions of accounts on other platforms.[18] Tenure refers to the time elapsed since a creator first posted a work, and we infer attitudes toward Generative AI by analyzing creators' non-work posts. We employ a median-split strategy (Lu et al. 2021): for popularity, we split treated creators into a *high* group (above-median) and a *low* group (below-median); for the binary moderators (multi-homing and expressed AI sentiment), value 1 forms the high group and value 0 the low group. We estimate treatment effects and standard deviations separately for each subgroup. Tenure cannot be split within cohort by construction, so we instead split cohorts into high- and low-tenure groups (above/below median) and aggregate within each.

Figure 5 summarizes the heterogeneity results for Lofter. More popular and multi-homing creators each reduce activity significantly more than their respective counterparts, while longer-tenure creators exhibit smaller reductions. AI-averse creators (those who had previously expressed negative attitudes toward Generative AI in their non-work posts) reduce activity more than other creators following Lofter's AI tool launch. This pattern is consistent with a key prediction of the signaling interpretation (policies that conflict with creators' prior views should generate stronger behavioral responses); the finding therefore strengthens the case for the signaling mechanism.



(a) Popularity (High / Low)

(b) Tenure (High / Low)

(c) AI-Averse or Not

(d) Multi-Homing or Not

Figure 5: Heterogeneity Analysis (Lofter)

[18] About 25% of Lofter users have shown multi-homing behavior on Lofter.

While the median split reveals that more popular creators respond more negatively, a binary comparison cannot detect potential non-linearities or competitive reallocation among less popular creators. Figure 6 presents results by popularity quartile. The pattern is monotonic: treatment effects become increasingly negative as popularity rises, with the most popular creators (Q4) experiencing the largest and most persistent declines. We find no evidence of competitive reallocation: even creators in the lowest quartile (Q1) show negative (though modest) point estimates rather than increases. The negative signal from Lofter's pro-AI policy decision therefore dominates any competitive benefits from reduced attention rivalry.

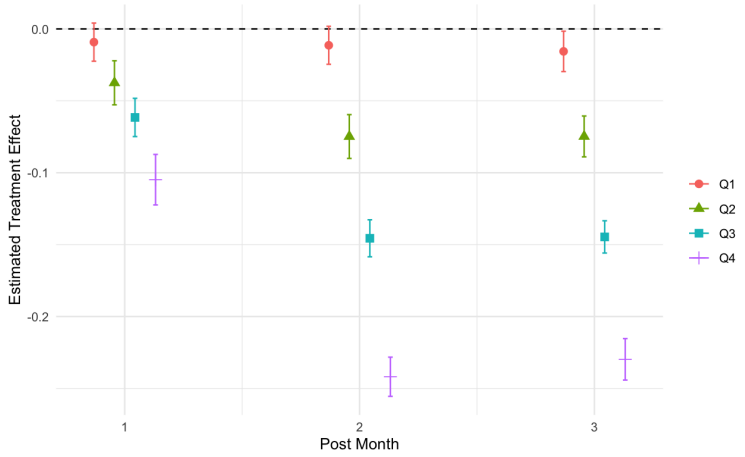
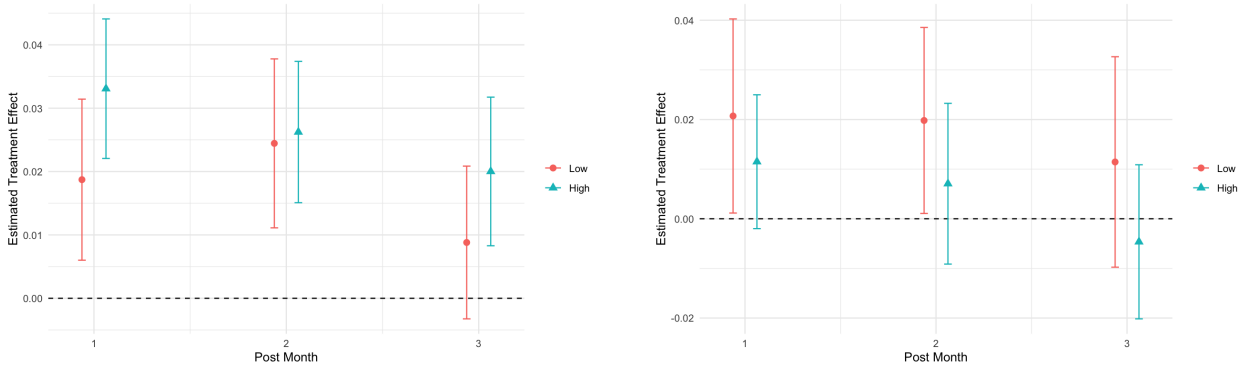


Figure 6: Heterogeneity Analysis by Creator Popularity: Quartile Results (Lofter)

In contrast to Lofter's clear heterogeneous patterns, Graffiti Kingdom exhibits limited evidence of heterogeneity. For Graffiti Kingdom, we focus only on popularity and tenure dimensions, as the platform does not offer a non-work post feature and thus we cannot infer creators' AI attitudes or reliably observe whether they maintain accounts on other platforms. Figure 7 shows that while the anti-AI stance signaled by Graffiti Kingdom's prohibition generates positive effects across creator segments, the differences between high- and low-popularity creators or between high- and low-tenure creators are not statistically distinguishable.



(a) Popularity (High / Low)

(b) Tenure (High / Low)

Figure 7: Heterogeneity Analysis (Graffiti Kingdom)

These results show that pro-AI policy signals can disproportionately alienate high-value but vulnerable

creator segments: popular creators (who account for an outsized share of content and traffic) and multi-homing creators (who face lower switching costs and can shift effort to competing platforms). Longer-tenure creators exhibit smaller reductions, suggesting that accumulated loyalty provides a buffer against negative AI-policy stance signals. The finding that AI-averse creators respond most strongly suggests that platforms can monitor expressed sentiment in advance to identify at-risk segments and anticipate churn before implementing AI-related policies.

6 Conclusion

This paper examines how the AI stances signaled by specific platform policy decisions influence creator behavior on visual arts platforms. We leverage two separate natural experiments: Lofter's launch of an AI image generator, which signaled a pro-AI stance, and Graffiti Kingdom's prohibition of AI-generated artwork, which signaled an anti-AI stance. These two policy decisions anchor opposite ends of the AI-stance spectrum, allowing us to offer a comprehensive picture of creator responses to contrasting AI-policy signals.

Our analysis yields three main findings. First, the pro-AI stance signaled by Lofter's AI-tool launch leads to significant and persistent declines in creator activity. Even after Lofter withdrew the tool and publicly apologized, Lofter creators continued to produce fewer works, suggesting that the pro-AI signal communicated by the focal launch decision had lasting effects on creator engagement. Second, the anti-AI stance signaled by Graffiti Kingdom's prohibition can strengthen creator activity. Graffiti Kingdom's prohibition yielded positive effects on creator output, demonstrating that platforms can credibly signal commitment to human creators and a protected environment for human-created work. Third, high-value creators are most sensitive to the AI stance signals communicated by platform policy decisions. Heterogeneity analysis reveals that more popular, multi-homing, and AI-averse creators reduced activity more on Lofter.

Where a platform sits along the output-centric versus creator-centric continuum likely shapes the magnitude and channels of these effects. On platforms where attention and economic value attach mainly to discrete creative works (and where concerns about replaceability, authorship, and intellectual property are particularly acute), the AI stance communicated by a platform's focal policy decision bears directly on the form of value creators supply and audiences consume. Such policy-communicated stance signals are especially likely to be interpreted as information about the future value of human creative labor, substitution risk from AI-generated outputs, and the platform's commitment to authorship and copyright norms. This interpretation is consistent with our findings: the estimated effects persist beyond the immediate mechanical consequences of the policies, and creators' expressed concerns center on replacement risk, low quality, and copyright infringement. As platforms move toward the creator-centric end of the continuum, where persona and ongoing audience relationships matter more, the same AI-related signals may still influence creator behavior, but their relative weight and the specific channels through which they operate may shift. In addition, since both empirical settings are Chinese visual-arts platforms, the observed responses may partly reflect China-specific cultural norms, creator-platform relationships, and the contemporaneous policy environment. Future research should examine comparable AI-policy changes across countries and regulatory regimes to assess the extent to which these findings generalize.

More broadly, our findings reveal that content creators face substantial uncertainty about their welfare in the current AI environment – uncertainty about replacement, intellectual property, and the value of their skills. The responses we observe reflect creators' rational reactions to these unresolved concerns under current conditions. Given the substantial potential of Generative AI to enhance creative productivity and expand creative possibilities, the lesson for platforms and policymakers should not be to exclude AI from creative platforms, but to create conditions that facilitate beneficial AI-human collaboration while protecting creator welfare.

Our findings offer several implications. At the tactical level, platforms can implement sentiment monitoring as an early warning system. Since creators who had previously expressed negative AI sentiment reduced activity more following Lofter's pro-AI policy signal, monitoring creator sentiment before implementing AI policies enables proactive identification of at-risk segments and targeted retention interventions. Platforms can also consider phased rollouts that target less-sensitive creator segments first. Since multi-homing creators (with lower switching costs) and shorter-tenure creators (with less accumulated loyalty) show larger reductions following pro-AI policy signals, platforms could initially introduce AI features to less-sensitive segments such as longer-tenure or single-homing creators, and monitor responses before broader deployment. Additionally, platforms can offer opt-out mechanisms that directly address the three primary concerns surfaced by our text analysis (replacement risk, perceived low quality, copyright infringement). Specific options include excluding works from AI training data, opting out of AI-assisted features, or clear labeling distinguishing human-created from AI-assisted content, which directly address replacement and copyright concerns while signaling respect for creator autonomy.

At the strategic level, the two cases together highlight a complementary principle:creators respond strongly to what focal AI-policy decisions communicate about AI's role on the platform, the platform's commitment to human creators, and the future competitive environment, not just to the mechanical availability of tools or the formal enforcement of rules. Lofter's pro-AI policy signal generated negative responses, while Graffiti Kingdom's anti-AI policy signal generated positive ones, and our text analysis confirms that only 7% of negative posts address the AI tool itself, while roughly 50% target the AI stance communicated by the platform's policy decision. This implies that credibly signaling commitment to creator interests can strengthen engagement, through measures such as creator-protective defaults, clear authorship and labeling standards, and transparent communication about how AI features will interact with human-created work. Platforms can leverage this responsiveness by making AI-related signaling intentional and consistent with creator interests.

For policymakers, our findings highlight the importance of establishing clear frameworks for AI-human collaboration in creative industries. The three primary creator concerns (replacement risk, quality standards, and copyright) all reflect gaps in the current regulatory and normative environment. Policymakers can help by clarifying intellectual property rights for AI-generated and AI-assisted content, establishing disclosure and labeling standards, and supporting mechanisms that allow creators to benefit from AI technologies. Importantly, this relationship is symbiotic: just as creators can benefit from AI, AI systems depend on high-quality human-created content for training and improvement. Well-designed mechanisms, such as compensation for training data contributions, attribution systems, and revenue-sharing arrangements, can

align incentives so that creators both benefit from and contribute to AI development. The ultimate goal should be an ecosystem where AI enhances human creativity while creators are fairly recognized and compensated for their role in advancing AI capabilities and creative expressions.

Funding and Competing Interests: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

References

Anderson A, Huttenlocher D, Kleinberg J, Leskovec J (2013) Steering user behavior with badges. *Proceedings of the 22nd international conference on World Wide Web*, 95–106.

Angrist JD, Pischke JS (2009) *Mostly harmless econometrics: An empiricist's companion* (Princeton university press).

Arora N, Chakraborty I, Nishimura Y (2025) Ai–human hybrids for marketing research: Leveraging large language models (llms) as collaborators. *Journal of Marketing* 89(2):43–70.

Babina T, Fedyk A, He A, Hodson J (2024) Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics* 151:103745.

Blaszczyk M (2025) Thaler v. perlmutter: Human authors at the center of copyright? .

Bleier A, Fossen BL, Shapira M (2024) On the role of social media platforms in the creator economy. *International Journal of Research in Marketing* 41(3):411–426.

Burtch G, Lee D, Chen Z (2023) The consequences of generative ai for ugc and online community engagement. *Available at SSRN 4521754* .

Castelo N, Bos MW, Lehmann DR (2019) Task-dependent algorithm aversion. *Journal of Marketing Research* 56(5):809–825.

Castelo N, Ward A (2016) Political affiliation moderates attitudes towards artificial intelligence. *ACR North American Advances* .

Cavusoglu H, Phan TQ, Cavusoglu H, Airoldi EM (2016) Assessing the impact of granular privacy controls on content sharing and disclosure on facebook. *Information Systems Research* 27(4):848–879.

Chen J, Roth J (2024) Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics* 139(2):891–936.

Chen Y, Ghosh A, McAfee RP, Pennock D (2008) Sharing online advertising revenue with consumers. *International workshop on internet and network economics*, 556–565 (Springer).

Connelly BL, Certo ST, Ireland RD, Reutzel CR (2011) Signaling theory: A review and assessment. *Journal of Management* 37(1):39–67.

Costabile C (2024) Digital platform ecosystem governance of private companies: Building blocks and a research agenda based on a multidisciplinary, systematic literature review. *Data and Information Management* 8(1):100053.

Crusey M (2025) Section 1202 (b) and ai: Implications for copyright infringement lawsuits and considerations for digital creators. *J. Copyright Soc'y* 72:480–485.

Dixon WJ, Massey Jr FJ (1951) Introduction to statistical analysis. .

Doshi AR, Hauser OP (2024) Generative ai enhances individual creativity but reduces the collective diversity of novel content. *Science advances* 10(28):eadn5290.

Duffy BE, Meisner C (2023) Platform governance at the margins: Social media creators' experiences with algorithmic (in) visibility. *Media, Culture & Society* 45(2):285–304.

Eichenbaum M, Godinho de Matos M, Lima F, Rebelo S, Trabandt M (2024) Expectations, infections, and economic activity. *Journal of Political Economy* 132(8):2571–2611.

Eloundou T, Manning S, Mishkin P, Rock D (2023) Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130* .

Epstein Z, Hertzmann A, of Human Creativity I, Akten M, Farid H, Fjeld J, Frank MR, Groh M, Herman L, Leach N, et al. (2023) Art and the science of generative ai. *Science* 380(6650):1110–1111.

Fleiss JL (1971) Measuring nominal scale agreement among many raters. *Psychological bulletin* 76(5):378.

Frazier K (2024) The past, present, and future of platform governance. *Journal of Regulatory Compliance* .

Goldberg S, Lam HT (2025) Generative ai in equilibrium: Evidence from a creative goods marketplace .

Goldberg SG, Johnson GA, Shriver SK (2024) Regulating privacy online: An economic evaluation of the gdpr. *American Economic Journal: Economic Policy* 16(1):325–358.

Goli A, Singh A (2024) Frontiers: Can large language models capture human preferences? *Marketing Science* 43(4):709–722.

Gourieroux C, Monfort A, Trognon A (1984a) Pseudo maximum likelihood methods: Applications to poisson models. *Econometrica: Journal of the Econometric Society* 701–720.

Gourieroux C, Monfort A, Trognon A (1984b) Pseudo maximum likelihood methods: Theory. *Econometrica: journal of the Econometric Society* 681–700.

Granger CW (1980) Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and control* 2:329–352.

Gray K (2017) Ai can be a troublesome teammate. *Harvard Business Review* 2:20–21.

Green A (2024) Artificial intelligence and the changing demand for skills in the labour market. Technical report, OECD Publishing.

Hanisch M, Goldsby CM, Fabian NE, Oehmichen J (2023) Digital governance: A conceptual framework and research agenda. *Journal of Business Research* 162:113777.

Hausman C, Rapson DS (2018) Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics* 10:533–552.

Holzner N, Maier S, Feuerriegel S (2025) Generative ai and creativity: A systematic literature review and meta-analysis. *arXiv preprint arXiv:2505.17241* .

Kathuria A, Karhade PP, Malik O, Baruri VP, Konsynski BR (2026) Ai governance and the decentralization of technology production: An investigation of ai-based ipa bots. *Information Systems Research* .

Khobzi H, Canhoto AI, Ramezani MS (2025) Content creators at a crossroads between decentralized and centralized social media. *Business Horizons* 68(1):109–120.

Kim JH, Kim M, Kwak DW, Lee S (2022) Home-tutoring services assisted with technology: investigating the role of artificial intelligence using a randomized field experiment. *Journal of Marketing Research* 59(1):79–96.

Kooperberg C, Stone CJ, Truong YK (1995) Hazard regression. *Journal of the American Statistical Association* 90(429):78–94.

Lambert D (1992) Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics* 34(1):1–14.

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *biometrics* 159–174.

Lechner M, et al. (2011) The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics* 4(3):165–224.

Lee BC, Chung J (2024) An empirical investigation of the impact of chatgpt on creativity. *Nature Human Behaviour* 8(10):1906–1914.

Lemley MA (2023) How generative ai turns copyright upside down. *Available at SSRN 4517702* .

Li P, Castelo N, Katona Z, Sarvary M (2024) Frontiers: Determining the validity of large language models for automated perceptual analysis. *Marketing Science* 43(2):254–266.

Li X, Liao C, Xie Y (2021) Digital piracy, creative productivity, and customer care effort: Evidence from the digital publishing industry. *Marketing Science* 40(4):685–707.

Liu Y, Feng J (2021) Does money talk? the impact of monetary incentives on user-generated content contributions. *Information Systems Research* 32(2):394–409.

Lloyd T, Gosciak J, Nguyen T, Naaman M (2025) Ai rules? characterizing reddit community policies towards ai-generated content. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–19.

Lu S, Rajavi K, Dinner I (2021) The effect of over-the-top media services on piracy search: evidence from a natural experiment. *Marketing Science* 40(3):548–568.

Lucchi N (2025) Generative ai and copyright: Training, creation, regulation .

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.

Menkhoff M (2025) Belief updating and ai adoption: Experimental evidence from firms .

Noy S, Zhang W (2023) Experimental evidence on the productivity effects of generative artificial intelligence. *Available at SSRN 4375283* .

Peng S, Kalliamvakou E, Cihon P, Demirer M (2023) The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590* .

Peres R, Schreier M, Schweidel DA, Sorescu A (2024) The creator economy: An introduction and a call for scholarly research.

Rousseau DM (1989) Psychological and implied contracts in organizations. *Employee responsibilities and rights journal* 2(2):121–139.

Rubin DB (1980) Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American statistical association* 75(371):591–593.

Sag M (2023) Fairness and fair use in generative ai. *Fordham L. Rev.* 92:1887.

Shan S, Cryan J, Wenger E, Zheng H, Hanocka R, Zhao BY (2023) Glaze: Protecting artists from style mimicry by {Text-to-Image} models. *32nd USENIX Security Symposium (USENIX Security 23)*, 2187–2204.

She S, Xu H, Wu Z, Tian Y, Tong Z (2020) Dimension, content, and role of platform psychological contract: based on online ride-hailing users. *Frontiers in psychology* 11:2097.

Shi Z, Liu X, Srinivasan K (2022) Hype news diffusion and risk of misinformation: the oz effect in health care. *Journal of Marketing Research* 59(2):327–352.

Shin M, Kim J, Shin J (2024) The adoption and efficacy of large language models: Evidence from consumer complaints in the financial industry. *Available at SSRN 5004194* .

Shin M, Kim J, van Opheusden B, Griffiths TL (2023) Superhuman artificial intelligence can improve human decision-making by increasing novelty. *Proceedings of the National Academy of Sciences* 120(12):e2214840120.

Silva JS, Tenreyro S (2006) The log of gravity. *The Review of Economics and statistics* 641–658.

Sobel B (2024) Elements of style: copyright, similarity, and generative ai. *Harvard Journal of Law & Technology, Forthcoming* 38.

Spence M (1973) Job market signaling. *The Quarterly Journal of Economics* 87(3):355–374.

Spitale G, Biller-Andorno N, Germani F (2023) Ai model gpt-3 (dis) informs us better than humans. *arXiv preprint arXiv:2301.11924* .

Sun M, Zhu F (2013) Ad revenue and content commercialization: Evidence from blogs. *Management Science* 59(10):2314–2331.

Toubia O, Stephen AT (2013) Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter? *Marketing Science* 32(3):368–392.

Turjeman D, Feinberg FM (2023) When the data are out: measuring behavioral changes following a data breach. *Marketing Science* .

Wittenberg C, Epstein Z, Péloquin-Skulski G, Berinsky AJ, Rand DG (2025) Labeling ai-generated media online. *PNAS nexus* 4(6):pgaf170.

Xu Y (2017) Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis* 25(1):57–76.

Yu H, Shen Z, Miao C, Leung C, Lesser VR, Yang Q (2018) Building ethics into artificial intelligence. *arXiv preprint arXiv:1812.02953* .

Zhang S, Mehta N, Singh PV, Srinivasan K (2021) Frontiers: Can an artificial intelligence algorithm mitigate racial economic inequality? an analysis in the context of airbnb. *Marketing Science* 40(5):813–820.

Zhou E, Lee D (2024) Generative artificial intelligence, human creativity, and art. *PNAS nexus* 3(3):pgae052.

Online Appendix A Platform AI Policies

This appendix documents the platform AI policies that ground our stance construct: the first subsection situates the two focal policy decisions within the broader cross-platform policy landscape, and the second reconstructs the timeline of AI-related actions on the two focal platforms.

A.1 Cross-Platform AI Policy Landscape

Tables A.1 and A.2 provide context for our two focal cases. Table A.1 situates the focal policies of Lofter and Graffiti Kingdom within the broader spectrum of AI-policy stance signals. Table A.2 documents the prevalence of the two specific intervention types we study (platform-developed AI tools and explicit AI prohibitions) across other comparable visual-art platforms.

Table A.1: Platform Policies and Policy-Communicated AI Stances: A Spectrum from Pro-AI to Anti-AI

Platform	Policy	Date	Stance (Pro-AI → Anti-AI)
Lofter	Developed a built-in Generative AI image generator.	2023-03	Active Embrace
Shutterstock	Contributor Fund compensating artists for data licensing; expanded six-year OpenAI partnership.	2023-07	Promotion / Partnership
CGTrader	No explicit policy; mostly silent on AI.	n/a	Permissive (Silent)
DeviantArt	Requires an explicit "Created with AI tools" label (auto-applied when metadata indicates AI).	2023-06	Permissive (Labeled)
Pixiv	Upload toggle for "AI-generated," user search filters, and separate rankings for AI works.	2022-10	Permissive (Segregated)
Cubebrush	AI content permitted only if labeled, but then hidden from search and storefronts—"essentially invisible."	2025-05	Restrictive (Labeled)
Dribbble	Community Guidelines: Do not share content that is wholly generated by AI. When creating content with AI assistance, it is recommended to acknowledge the contribution of AI in your Shot description.	n/a	Restrictive (Human-Primary)
Graffiti Kingdom	Community rules prohibit uploading AI-generated images. Any users who violate this will be punished.	2023-02	Active Restriction

Note. The unit of analysis is the platform-policy case. Policies are ordered from most pro-AI to most anti-AI, with the focal Lofter and Graffiti Kingdom policy decisions anchoring the opposite ends of the spectrum.

Table A.2: Platforms with Generative AI Policies Comparable to Lofter and Graffiti Kingdom

Platform	Region	Core Focus	AI-Policy Bucket	Key Policy Summary
Lofter	China	Broad creative community with strong visual-art, illustration, photography, and fandom subcommunities	Platform-developed AI feature	LOFTER introduced and tested an in-platform AI image generator/profile-picture tool (often described as a "Profile Picture Generator" or AI drawing feature), later narrowing or removing access after creator backlash.
DeviantArt	Global (U.S.-based)	Online art community and portfolio platform for digital/traditional artists	Platform-developed AI feature	DeviantArt operates DreamUp, its own text-to-image service and policy framework for AI image generation on-platform.
Picsart	Global	Creator platform for image/video editing and design with a large maker community	Platform-developed AI feature	Picsart is primarily aimed at human creators and integrates in-platform generative-AI tools, including an AI image generator and prompt-based editing features, into its creator workflow.
Clip Studio Paint	Japan / Global	Digital drawing and comics platform used by illustrators, manga artists, and studios	Platform-developed AI feature	Clip Studio Paint officially announced an experimental in-app 'Image Generator palette' based on Stable Diffusion for creators, and then withdrew it days later after backlash from artists.
Graffiti Kingdom	China	Original illustration / drawing community and artist platform	Explicit prohibition	Platform content rules state that AI-generated images may not be uploaded; repeat violators can have works removed and uploading privileges restricted.
Cara	Global	Artist social and portfolio platform	Explicit prohibition	AI-generated media are not allowed in portfolios on Cara; AI content may appear only for news/reporting/discussion outside portfolio content and must be accurately labeled.
Newgrounds Art Portal	Global (U.S.-based)	Creator community with dedicated portals for art, games, audio, and animation	Explicit prohibition	Newgrounds' Art Portal does not allow AI-generated art, aside from limited disclosed ancillary use cases (for example, an AI-generated background accompanying otherwise human-made character art).
INPRNT	Global (U.S.-based)	Curated art print marketplace and artist storefront platform	Explicit prohibition	INPRNT's content guidelines prohibit artworks generated completely via an automated AI or machine-learning process.
Fur Affinity	Global (U.S.-based)	Online art community and portfolio platform focused on furry art	Explicit prohibition	Fur Affinity's upload policy states that content made using AI, machine learning, or similar generators is not allowed, including derivative works created from it.

Note. The unit of analysis is the platform-policy case. The table summarizes comparable visual-art or creator platforms with either platform-developed Generative AI features or explicit prohibitions of AI-generated content.

A.2 Timeline of AI-Related Actions on the Focal Platforms

To provide a fuller account of the two policy episodes, we searched official platform pages, contemporaneous Chinese-language news and legal coverage, and the focal announcements retained by our author team. We included platform-level product actions, content rules, public explanations, and implementation measures that changed or explained how AI creation or AI-generated content would be governed; user commentary and third-party discussion were used only as corroborating evidence. Table A.3 reports the resulting chronology, covering the focal decisions and the directly related clarifications, restrictions, withdrawal, remediation, and implementation actions.

Table A.3: Timeline of Publicly Documented AI-Related Actions and Announcements on Lofter and Graffiti Kingdom, July 2022–June 2023

Platform	Date	Publicly Documented Action or Announcement	Role in the Focal Policy Episode	Evidence
Lofter	March 6, 2023	Publicly opened the "Profile Picture Generator"; the limited-test version had been called the "Lofter Drawing Machine." Users could enter text prompts to generate images. In statements issued that day, Lofter characterized the tool as an avatar-use feature, stated that it used open-source training data rather than Lofter user works, prohibited commercial use, and promised adjustments.	*Focal launch and immediate explanation.* The product launch constitutes the treatment. The same-day statements explained the intended use of the tool without reversing the launch.	Jiemian; Sohu/Bianews; Pandaily
Lofter	March 7, 2023	Announced that the feature entry would be moved to the Avatar Frame Center and that generated images would be limited to profile-picture use, without download or publication functions. Lofter also announced prospective governance measures, including a rule against presenting AI-generated content as original work, a related feedback channel, anti-AI-scraping capability, and product mechanisms to distinguish AI content from original work.	*Reactive scope restriction and prospective remediation.* The restrictions narrowed the use of the focal tool. The additional governance measures were announced as future responses to creator concerns rather than as separately initiated AI policies.	Archived official statement; JWView; Superpixel (archived)
Lofter	March 8, 2023	Lofter subsequently stated that it had taken the disputed image-generation feature offline.	*Withdrawal.* The action removed the focal product feature and restored the pre-launch product state; it did not establish a new affirmative AI-governance policy.	Blue Whale News; China Economic Net
Lofter	March 10, 2023	Published an in-app apology acknowledging creator dissatisfaction and announced a forthcoming "Creator Protection Plan." The plan outlined prospective measures, including capabilities intended to prevent AI-related misappropriation and malicious scraping, a dedicated infringement-reporting channel, restrictions on presenting AI-generated content as original work, stronger AI-content identification and feedback handling, a planned separate section for AI-generated content, and a future community-partner mechanism.[a]	*Apology and prospective remedial program.* The plan was presented as a response to the launch controversy. This row records the measures announced by Lofter and does not treat the listed measures as having been implemented at that time.	Sina; 21st Century Business Herald
Lofter	March 16, 2023	Issued a second public apology, reiterated that the disputed feature had been taken offline, and again stated that user works had not been used for AI training. Lofter reported that the rule against presenting AI content as original was in force, that an anti-misappropriation and anti-scraping system and a dedicated infringement-reporting channel were active, and that 1,148 infringement complaints had been processed by March 15.[b] It continued to describe the "Creator Protection Plan" as a forthcoming initiative.[a]	*Second apology and platform-reported follow-up.* The statement reported progress on selected elements of the remedial program within the same policy episode; it did not document implementation of every element announced on March 10.	Blue Whale News; Securities Daily/Tencent
Graffiti Kingdom	February 22, 2023	Announced that AI-generated artwork was prohibited, that previously uploaded works would be reviewed, and that violations would trigger sanctions. Platform staff subsequently clarified that a sanctioned user's account and previously uploaded non-AI works would be preserved.	*Focal prohibition.* The policy explicitly excluded AI-generated artwork and constitutes the treatment analyzed in the paper.	Announcement retained by the authors; author correspondence; official content rules; contemporaneous legal commentary
Graffiti Kingdom	February 23–June 30, 2023	Maintained and implemented the February 22 prohibition through review and sanctions for AI-generated uploads. Our search identified no separately initiated AI-specific product launch or policy change during the remaining window.	*Implementation and maintenance.* These actions operationalized the same prohibition rather than creating a separate contemporaneous AI-policy shock.	Author correspondence and platform records; official content rules

Note. The primary event-search window is July 1, 2022–June 30, 2023. Searches combined each platform's Chinese and English names with terms including AI, AIGC, AI drawing, AI-generated content, avatar generator, creator protection, anti-scraping, prohibition, and content rules. The table reports the platform-level actions and announcements identified in the public record and the focal materials retained by the authors; it does not rule out unobserved internal deliberations or actions that were not publicly documented. The original February 22 Graffiti Kingdom announcement is retained by the authors but is no longer publicly indexed; the current official rules and contemporaneous March 2023 commentary provide external corroboration. We found no additional independently initiated AI-specific platform action before the focal decisions or after the listed follow-up actions within the primary search window. Lofter's March 6–16 actions are presented separately to make the sequence visible, although the post-launch restrictions, withdrawal, apologies, and remedial announcements were reactive components of the same policy episode. All hyperlinks in this table were last accessed on July 22, 2026.

[a] For the prospective Lofter measures, we conducted an additional implementation search covering the year through March 16, 2024. We found no subsequent public announcement or independently verifiable evidence confirming that the planned separate AI-content section or the Lofter Partner Plan was launched. The absence of publicly available evidence does not rule out unpublicized, partial, or limited implementation.

[b] The March 16 implementation claims were made by Lofter and reported in contemporaneous news coverage. We found no independent technical audit of the anti-misappropriation or anti-scraping system. Moreover, the reported 1,148 complaints were described as infringement complaints generally, rather than specifically as AI-related complaints.

Online Appendix B Data Collection Details

Lofter

Lofter hosts various content formats (painting, writing, photography, filmmaking). We focus on visual artists because the AI image generator primarily affects this group. We compiled a list of visual-arts tags (e.g., *drawing*, *painting*). Among semantically similar tags, *painting* is most popular and provides exhaustive uploads for arbitrary time intervals. We retrieved all works tagged *painting* from January 2020 to December 2023, then visited each artist's homepage to collect their complete work history.

Graffiti Kingdom

Graffiti Kingdom maintains a "New Works Today" section indexing all daily uploads. We systematically scraped publicly reachable historical pages (approximately 40,000 pages), achieving near-complete coverage from 2019 onward. For each upload, we visited the artist's homepage and extracted all their previously uploaded works.

Full-Sample Descriptive Statistics

Table B.1 reports descriptive statistics for the full creator-month panel on each platform.

Table B.1: Full-Sample Descriptive Statistics of Key Variables

Variable	Definition	N	Mean	Std.Dev.	Min	Max
Panel A: Lofter						
Num of Works	Number of works uploaded during the month	7,510,542	1.84	18.35	0	14,726
Active	Equals 1 if creator uploads ≥1 work	7,510,542	0.05	0.22	0	1
review_avg	Average number of reviews received	7,510,542	0.98	9.41	0	8,141
like_avg	Average number of likes received	7,510,542	68.87	1,085.66	0	534,697
Panel B: Graffiti Kingdom						
Num of Works	Number of works uploaded during the month	1,822,186	1.01	4.64	0	1,464
Active	Equals 1 if creator uploads ≥1 work	1,822,186	0.16	0.37	0	1
review_avg	Average number of reviews received	1,822,186	0.07	0.85	0	167
like_avg	Average number of likes received	1,822,186	1.77	10.62	0	1,629

Note. The unit of analysis is the creator-month. The full sample covers Lofter from January 2020 to December 2023 and Graffiti Kingdom from January 2019 to December 2023; *N* denotes the number of creator-month observations.

Online Appendix C TCI Cohort Construction and Validation Tests

Tables C.1 and C.2 report, for each treated cohort on Lofter and Graffiti Kingdom, the matched control cohorts, the cohort sample sizes, and the per-pair Kolmogorov–Smirnov and Granger causality test results referenced in Section 4.

Table C.1: Validation Results for Temporal Causal Inference: Kolmogorov–Smirnov and Granger Causality Tests for Lofter

Treated cohort	Control cohorts	Treated creators	Control creators	Granger F-statistic	Granger p-value	Kolmogorov–Smirnov D-statistic	Kolmogorov–Smirnov p-value
8	11,12,13,14,15	6,512	28,973	1.9011	0.2400	0.1250	1.0000
9	12,13,14,15,16	5,647	27,981	3.1156	0.1378	0.1111	1.0000
10	13,14,15,16,17	5,238	25,976	8.3616	0.0341	0.1000	1.0000
11	14,15,16,17,18	5,200	24,174	4.3150	0.0924	0.0909	1.0000
12	15,16,17,18,19	6,121	23,034	1.0064	0.3618	0.0833	1.0000
13	16,17,18,19,20	5,742	21,859	1.3114	0.3040	0.0769	1.0000
14	17,18,19,20,21	5,882	23,493	0.0252	0.8802	0.0714	1.0000
15	18,19,20,21,22	6,028	23,196	1.0955	0.3432	0.0667	1.0000
16	19,20,21,22,23	4,208	21,971	56.0278	0.0007	0.1250	0.9999
17	20,21,22,23,24	4,116	19,586	2.2838	0.1911	0.1176	0.9999
18	21,22,23,24,25	3,940	17,129	1.6858	0.2508	0.1111	1.0000
19	22,23,24,25,26	4,742	13,833	0.5877	0.4779	0.0526	1.0000
20	23,24,25,26,27	4,853	13,319	0.0348	0.8593	0.0500	1.0000
21	24,25,26,27,28	5,842	13,812	0.5614	0.4874	0.0476	1.0000
22	25,26,27,28,29	3,819	13,807	11.4692	0.0195	0.0455	1.0000
23	26,27,28,29,30	2,715	13,724	4.6711	0.0831	0.0435	1.0000
24	27,28,29,30,31	2,357	13,311	0.7441	0.4278	0.0417	1.0000
25	28,29,30,31,32	2,396	12,236	0.8257	0.4052	0.0400	1.0000
26	29,30,31,32,33	2,546	10,617	0.3923	0.5586	0.0385	1.0000
27	30,31,32,33,34	3,305	10,771	0.0156	0.9056	0.0370	1.0000
28	31,32,33,34,35	3,208	10,906	0.4819	0.5185	0.0357	1.0000
29	32,33,34,35,36	2,352	11,214	1.0702	0.3483	0.0690	1.0000
30	33,34,35,36,37	2,313	12,484	0.1336	0.7297	0.0667	1.0000
31	34,35,36,37,38	2,133	15,505	0.5147	0.5052	0.0645	1.0000
32	35,36,37,38,39	2,230	18,748	0.7012	0.4406	0.0625	1.0000
33	36,37,38,39,40	1,589	22,013	0.6622	0.4528	0.0909	0.9995
34	37,38,39,40,41	2,506	23,081	2.6588	0.1639	0.0882	0.9996
35	38,39,40,41,42	2,448	21,329	3.9318	0.1042	0.0857	0.9997
36	39,40,41,42,43	2,441	18,043	0.7938	0.4138	0.0556	1.0000
37	40,41,42,43,44	3,500	13,661	0.8702	0.3937	0.0541	1.0000
38	41,42,43,44,45	4,610	9,409	0.0128	0.9142	0.0526	1.0000
39	42,43,44,45,46	5,749	7,260	1.4256	0.2860	0.0513	1.0000
40	43,44,45,46,47	5,713	7,790	0.4799	0.5193	0.0750	0.9999
41	44,45,46,47,48	3,509	8,218	0.0346	0.8597	0.0976	0.9914
42	45,46,47,48,49	1,748	8,137	1.3046	0.3051	0.0714	1.0000
43	46,47,48,49,50	1,324	7,694	8.3095	0.0345	0.0698	1.0000
44	47,48,49,50,51	1,367	7,372	0.1346	0.7287	0.0682	1.0000
45	48,49,50,51,52	1,461	6,132	0.1158	0.7474	0.0444	1.0000
46	49,50,51,52,53	1,360	5,619	0.2759	0.6218	0.0652	1.0000
47	50,51,52,53,54	2,278	5,652	0.5760	0.4821	0.0638	1.0000
48	51,52,53,54,55	1,752	5,546	0.2187	0.6597	0.0625	1.0000
49	52,53,54,55,56	1,286	5,298	5.1326	0.0728	0.0612	1.0000
50	53,54,55,56,57	1,018	5,105	3.0342	0.1420	0.0600	1.0000
51	54,55,56,57,58	1,038	4,693	3.0267	0.1424	0.0392	1.0000
52	55,56,57,58,59	1,038	4,252	0.0036	0.9543	0.0385	1.0000
53	56,57,58,59,60	1,239	4,646	0.0803	0.7883	0.0566	1.0000
54	57,58,59,60,61	1,319	4,921	0.0607	0.8152	0.0370	1.0000
55	58,59,60,61,62	912	4,866	2.0041	0.2160	0.0545	1.0000

Note. The table-level unit of analysis is a treated-cohort/matched-control-cohort pair; the underlying data are creator-month observations. Cohort i refers to creators whose first work was posted i months before Lofter's launch of the "Profile Picture Generator". Rows with Granger p-values below 0.05 fail the validation screen and are excluded from the subsequent Temporal

Table C.2: Validation Results for Temporal Causal Inference: Kolmogorov–Smirnov and Granger Causality Tests for Graffiti Kingdom

Treated cohort	Control cohorts	# Treated creators	# Control creators	Granger F-statistic	Granger p-value	Kolmogorov–Smirnov D-statistic	Kolmogorov–Smirnov p-value
7	10,11,12,13,14	925	5,582	0.0008	0.9791	0.1429	1.0000
8	11,12,13,14,15	1,041	5,951	0.0688	0.8101	0.1250	1.0000
9	12,13,14,15,16	1,200	5,848	0.1981	0.6864	0.1111	1.0000
10	13,14,15,16,17	963	5,720	11.7603	0.0416	0.1000	1.0000
11	14,15,16,17,18	1,091	5,656	4.3461	0.1284	0.0909	1.0000
12	15,16,17,18,19	1,200	5,582	0.8339	0.4285	0.0833	1.0000
13	16,17,18,19,20	1,188	5,260	1.0625	0.3785	0.1538	0.9992
14	17,18,19,20,21	1,140	4,966	2.9061	0.1868	0.1429	0.9996
15	18,19,20,21,22	1,332	4,498	5.5025	0.1007	0.1333	0.9998
16	19,20,21,22,23	988	3,892	3.2195	0.1706	0.0625	1.0000
17	20,21,22,23,24	1,072	3,367	0.8487	0.4249	0.1176	0.9999
18	21,22,23,24,25	1,124	2,862	0.7677	0.4454	0.1111	1.0000
19	22,23,24,25,26	1,066	2,717	34.5998	0.0098	0.1579	0.9781
20	23,24,25,26,27	1,010	2,488	1.3147	0.3347	0.1000	1.0000
21	24,25,26,27,28	694	2,428	9.0789	0.0571	0.0952	1.0000
22	25,26,27,28,29	604	2,283	4.3433	0.1285	0.0455	1.0000
23	26,27,28,29,30	518	2,162	4.4171	0.1264	0.0870	1.0000
24	27,28,29,30,31	541	1,986	12.4283	0.0388	0.0417	1.0000
25	28,29,30,31,32	505	1,995	6.1046	0.0900	0.0800	1.0000
26	29,30,31,32,33	549	2,086	4.8862	0.1141	0.0385	1.0000
27	30,31,32,33,34	375	2,455	1.4238	0.3185	0.0370	1.0000
28	31,32,33,34,35	458	2,902	2.3572	0.2223	0.0714	1.0000
29	32,33,34,35,36	396	3,376	0.6368	0.4832	0.0690	1.0000
30	33,34,35,36,37	384	4,113	2.8048	0.1926	0.1000	0.9988
31	34,35,36,37,38	373	4,542	0.0288	0.8759	0.1290	0.9634
32	35,36,37,38,39	384	4,564	29.3603	0.0123	0.0625	1.0000
33	36,37,38,39,40	549	4,112	7.4284	0.0722	0.0606	1.0000
34	37,38,39,40,41	765	3,515	11.2943	0.0437	0.0588	1.0000
35	38,39,40,41,42	831	2,663	1.1378	0.3643	0.0571	1.0000
36	39,40,41,42,43	847	1,968	0.4950	0.5324	0.0833	0.9998
37	40,41,42,43,44	1,121	1,481	0.6806	0.4699	0.1081	0.9846
38	41,42,43,44,45	978	1,457	9.5919	0.0534	0.1579	0.7379
39	42,43,44,45,46	787	1,474	3.1554	0.1738	0.1282	0.9114
40	43,44,45,46,47	379	1,518	0.7627	0.4468	0.1250	0.9188
41	44,45,46,47,48	250	1,578	0.0155	0.9088	0.0976	0.9914
42	45,46,47,48,49	269	1,546	0.3189	0.6117	0.0952	0.9926

Note. The table-level unit of analysis is a treated-cohort/matched-control-cohort pair; the underlying data are creator-month observations. Cohort i refers to creators whose first work was posted i months before Graffiti Kingdom's prohibition of AI-generated artwork. Rows with Granger p-values below 0.05 fail the validation screen and are excluded from the subsequent Temporal Causal Inference estimation; all listed pairs pass the Kolmogorov–Smirnov test.

Online Appendix D Identification Assumptions & Validation

Before presenting the estimation results, we discuss the identification assumptions following the Potential Outcome Framework (Rubin 1980) to validate the TCI framework in our context.

A1: SUTVA (Stable Unit Treatment Value Assumption)

The SUTVA assumption (Rubin 1980) consists of two conditions. First, there should be no interference between units. In our case, this means that whether one content creator is exposed to the treatment should not affect the outcomes of any other content creators on the platform. In other words, $Y_i^{(W_i)} \perp Y_j^{(W_j)}$ for any two content creators $i, j \in 1, 2, ..., N$. Under our TCI framework, the assignment to treatment or control groups is primarily based on the cohort to which the content creator belongs. This implies that the outcomes do not occur at the same real-world time (Turjeman and Feinberg 2023), meaning that no two users in different groups (i.e., control and treatment) can influence each other's treatment or lack thereof (i.e., $\forall i, j$, $Y_i^{(1)} \perp Y_j^{(0)}$).

An important concern here is the network effect. First, due to the potential network effect, $Y_i^{(1)} \perp Y_j^{(1)}$ and $Y_i^{(0)} \perp Y_j^{(0)}$ may not hold for some i, j pairs. For instance, users may increase or decrease their activity in response to the behavior of other users they know. However, since our objective is to measure how the AI stances signaled by platform policy decisions affect users' behavior, this potential dependency is a crucial aspect of estimating the treatment effect and should be considered in its assessment (Turjeman and Feinberg 2023).

Second, the potential network effect means that, as earlier users joining the platform, the activities of content creators in the control group may influence those in the treatment group. However, both platforms are stable and mature as they have been established for over ten years. Moreover, the treatment assignment is independent of this potential effect since the treatments (i.e., the platform policies) could not be anticipated. Even if interference does exist between units – where the behavior of content creators in the control group, who joined before the treatment group, influences those in the treatment group – a similar argument applies to the control group concerning their predecessors. This means the activities of content creators in the control group are also affected by those who joined earlier. Therefore, any similar effects, if they do exist, would be accounted for by simply comparing the control and treatment units (Turjeman and Feinberg 2023). Moreover, we apply the Granger causality test (Turjeman and Feinberg 2023, Granger 1980): For each treatment–control pair, the test evaluates whether behavior of the control cohort (earlier entrants) has predictive power for behavior in the treatment cohort. Evidence of predictive power implies a breach of SUTVA; accordingly, we classify that treatment–control pair as failing the Granger test, which are not used in our analysis.

Additionally, SUTVA requires that no hidden versions of treatments exist. In other words, nothing except the treatment itself can influence the outcome variables, i.e., $Y_i = Y_i^{(1)} W_i + Y_i^{(0)}(1 - W_i)$. We meticulously reviewed all previous announcements from the two platforms and confirmed that, at the time of the treatments, there were no other changes to the platforms, significant holidays, or confounding events apart from the treatments themselves. We also examined other prominent digital art platforms in China to verify that they did not take any actions related to Generative AI during the same period. In particular, we check the announcements of the following platforms: ZCOOL, Mihuashi, bcy.net, huaban.com, and duitang.com,

all of which are well-known Chinese platforms and can be considered competitors to Lofter and Graffiti Kingdom. Moreover, our TCI framework, which incorporates multiple cohorts into the control group, helps mitigate the impact of unobserved events on the outcome variable Y_i. Any potential effects from such events can be averaged out across all control cohorts, providing a more reliable control group baseline (Turjeman and Feinberg 2023). Furthermore, we apply the Kolmogorov-Smirnov (KS) test to demonstrate that there are no significant differences in behaviors between the control groups constructed by TCI and the treatment groups prior to the treatments. This is discussed in more detail in the next assumption.

A2: Conditional Independence Assumption

This assumption states that, the assignment of treatment or control group should be (approximately) random, conditional on the control variables X_i, ensuring the comparability between treatment and control units:

$$Pr(W_i|X_i, Y^{(0)}, Y^{(1)}) = Pr(W_i|X_i)$$

While the treatments on both platforms can be considered exogenous since the policy shocks are unanticipated, one potential concern is that the joining time may be confounded with user activities. Under the TCI framework, the constructed control groups consist of content creators who joined before those in the treatment groups. Therefore, if unobserved factors influence both the joining time and user activities on the platforms, our estimated treatment effects would be biased.

To address this concern, we apply the Kolmogorov-Smirnov (KS) Test to determine whether the control and treatment groups differ in their activities at the same tenures (Dixon and Massey Jr 1951, Turjeman and Feinberg 2023). Specifically, for each treatment-control pair, we calculate the cumulative sum of the average number of works uploaded to the platforms before the treatment. We then divide this by the maximum cumulative sum of the groups, which create a CDF-like timeline that can be tested using the KS Test (Turjeman and Feinberg 2023).

Only treatment-control pairs that pass both KS and Granger tests are included in our subsequent analysis. Columns (5)-(8) of Tables C.1 and C.2 show the KS Test and Granger Causality Test's results for Lofter and Graffiti Kingdom, separately. We can see, all the cohorts (on both platforms) have passed the KS test while a few cohorts do not pass the Granger test, and hence will not be used in our later analysis. Since only a small number of cohorts fail the Granger test and therefore excluded from the analysis, the impact on our sample size is minimal.

Last but not least, to test the parallel trends assumption for the DiD specification, we run an event study for each cohort and aggregate the resulting coefficients across cohorts using the same inverse-cohort-size weighting described above. We normalize the period immediately before treatment as the base period and interpret all event-time coefficients relative to this pre-treatment benchmark. As seen in Figure 3 and its corresponding Table D.1, the pre-period coefficients do exhibit some mild directional drift visually, but they are not statistically different from zero, indicating no meaningful pre-trends.[19]

A3: Overlap Assumption

[19]The wider confidence intervals for Lofter in later event-time periods do not reflect compositional change. In our setting, users who become inactive remain in the sample, with their number of works recorded as zero, so the estimates are not driven by attrition of inactive users. Rather, the widening confidence intervals primarily reflect greater heterogeneity in creators' post-treatment behavior.

Table D.1: Parallel-Trends Test under Temporal Causal Inference

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	log(Num of Works + 1)	
Treat × Pre Month 7	-0.0033	0.0045
	[0.0076]	[0.0095]
Treat × Pre Month 6	-0.0085	0.0035
	[0.0057]	[0.0088]
Treat × Pre Month 5	-0.0106	0.0041
	[0.0067]	[0.0088]
Treat × Pre Month 4	-0.0094	0.0042
	[0.0065]	[0.0079]
Treat × Pre Month 3	-0.0048	-0.0012
	[0.0047]	[0.0079]
Treat × Pre Month 2	0.0007	0.0019
	[0.0041]	[0.0072]
Treat × Pre Month 1	0	0
	[—]	[—]
Treat × Post Month 1	-0.0894***	0.0169**
	[0.0077]	[0.0074]
Treat × Post Month 2	-0.1707***	0.0146*
	[0.0079]	[0.0081]
Treat × Post Month 3	-0.1707***	0.0055
	[0.0081]	[0.0091]
Creator Fixed Effects	Yes	Yes
Tenure Fixed Effects	Yes	Yes
Number of Creators	140,061	24,178

Note. The unit of analysis is the creator-month; Pre Month 1 is the omitted reference period and its coefficient is normalized to zero. Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

This assumption requires that, for each content creator, her propensity to be treated should be strictly between 0 and 1:

$$0 < P(W_i = 1|X_i = x) < 1$$

Due to the nature of the exogenous treatments, every content creator who joined the platforms before the treatment has a positive chance of being treated. In fact, all content creators in our main sample were eventually treated, but at different tenures on the platforms.

Finally, Table 1 reports summary statistics for the final estimation sample. This sample retains only observations from cohorts that pass both the Granger causality test and the KS test. In addition, as mentioned earlier, we exclude users who joined more than 55 months earlier before the focal date on Lofter and more than 42 months on Graffiti Kingdom, so as to restrict attention to cohorts that are more comparable in their pre-treatment behavioral dynamics. We note that the reduction in sample size is somewhat larger for Graffiti Kingdom than for Lofter. One reason is that Graffiti Kingdom has relatively earlier-entry cohorts. Because these cohorts correspond to users who had been on the platform longer, they contribute more observations to the panel. As a result, excluding them removes a larger share of the final sample.

Online Appendix E Hazard Regression & Results

To assess whether policy timings are strategically aligned with creators' activities on the platforms, we estimate a discrete-time hazard model on platform–month aggregates prior to treatments ($t \leq 1$, with $t = 1$ denoting the happening months of the treatments). Let Posts_{pt} (Num of Works), Likes_{pt} (Num of Likes), and Reviews_{pt} (Num of Reviews) be monthly totals. Here, the number of likes and the number of reviews denote the total counts received by works uploaded in month t. Because we do not observe time-varying data on the likes and reviews that each work receives, we use these variables as proxies for the actual number of likes and the actual number of reviews on the whole platform during month t. We construct lagged levels $\text{lvl\_x}_{pt} = \text{x}_{p,t-1}$, six–month rolling trends slp_x_{pt} (OLS slope within the past window), and volatilities vol_x_{pt} (within–window standard deviation), for $x \in \{\text{Posts}, \text{Likes}, \text{Reviews}\}$. The adoption indicator is defined as

$$\text{adopt}_{pt} = \mathbb{I}\{t = 1\},$$

which equals 1 in the month when the treatment occurs and 0 otherwise.

We then estimate a hazard regression to assess whether pre-event dynamics predict adoption timing, with the model specification as follows:

$$\Pr(\text{adopt}_{pt} = 1 \mid \mathcal{H}_{p,t-1}) = \text{logit}^{-1}(Z_{pt}),$$

where $Z_{pt} = \beta_0 + \beta_1 \text{lvl\_posts}_{pt} + \beta_2 \text{slp\_posts}_{pt} + \beta_3 \text{vol\_posts}_{pt} + \beta_4 \text{lvl\_likes}_{pt} + \beta_5 \text{slp\_likes}_{pt} + \beta_6 \text{vol\_likes}_{pt} + \beta_7 \text{lvl\_reviews}_{pt} + \beta_8 \text{slp\_reviews}_{pt} + \beta_9 \text{vol\_reviews}_{pt} + \gamma \text{t}$. Table E.1 reports the estimates. None of these coefficients is statistically significant, providing no evidence that pre-event platform-level dynamics predict policy adoption timing.

Table E.1: Discrete-Time Hazard Regression Results for Policy Adoption Timing

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	Adopt	
(Intercept)	7.6121	-7.6985
	[16.6743]	[13.1650]
lvl_posts	-0.0001	0.0001
	[0.0001]	[0.0005]
slp_posts	0.0001	-0.0026
	[0.0002]	[0.0040]
vol_posts	0.0001	0.0018
	[0.0002]	[0.0018]
lvl_likes	0.00001	-0.0001
	[0.0001]	[0.0002]
slp_likes	-0.0001	0.0007
	[0.0001]	[0.0010]
vol_likes	0.00001	-0.0003
	[0.0001]	[0.0003]
lvl_reviews	0.0001	0.0010
	[0.0003]	[0.0025]
slp_reviews	0.0001	-0.0122
	[0.0004]	[0.0153]
vol_reviews	-0.0001	0.0061
	[0.0002]	[0.0091]
time_trend	0.1389	-0.0840
	[0.3396]	[0.2369]
Number of Observations	31	31

Note. The unit of analysis is the platform-month. The dependent variable equals one in the policy-adoption month and zero otherwise; each column uses 31 monthly observations ending in that month (Lofter: March 6, 2023; Graffiti Kingdom: February 22, 2023). The covariates are lagged levels (`lvl`), six-month rolling slopes (`slp`), and six-month rolling volatilities (`vol`) of monthly platform-level totals of works/posts, likes, and reviews. Robust standard errors clustered at the month level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Online Appendix F Long-Term Signaling Effect of Lofter's Pro-AI Policy Signal

Given the estimates in Table 2, we observe negative and significant effects for Lofter and positive and significant effects for Graffiti Kingdom in the first three post-treatment months. This naturally raises the question of whether these effects persist over a longer horizon.

Accordingly, we extend the post-treatment window to six months and re-estimate the TCI for both platforms. As reported in Table F.1, the estimated coefficients remain statistically significant in all six post-treatment months for both Lofter and Graffiti Kingdom. For Lofter, all coefficients remain negative, consistent with a persistent negative signaling effect of Lofter's pro-AI policy signal on creators' behavior. For Graffiti Kingdom, all coefficients remain positive, consistent with a positive effect that holds over the six-month window. It is worth noting that, because the post-treatment window now spans six months, we exclude control cohorts that entered less than six months before the treated cohorts, so that the control group remains uncontaminated over the full post-treatment horizon.

Table F.1: Six-Month Temporal Causal Inference Results Using Difference-in-Differences

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	log(Num of Works + 1)	
Treat × Post Month 1	-0.0935***	0.0283***
	[0.0075]	[0.0085]
Treat × Post Month 2	-0.1170***	0.0274***
	[0.0065]	[0.0085]
Treat × Post Month 3	-0.0964***	0.0202**
	[0.0067]	[0.0084]
Treat × Post Month 4	-0.0906***	0.0206**
	[0.0066]	[0.0096]
Treat × Post Month 5	-0.0832***	0.0251***
	[0.0067]	[0.0097]
Treat × Post Month 6	-0.0771***	0.0315***
	[0.0053]	[0.0094]
Estimation Method	Difference-in-Differences	
Creator Fixed Effects	Yes	Yes
Tenure Fixed Effects	Yes	Yes
Number of Creators	140,061	27,897

Note. The unit of analysis is the creator-month; Post Month 1–6 denote the first six post-treatment periods. Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Online Appendix G Robustness Checks

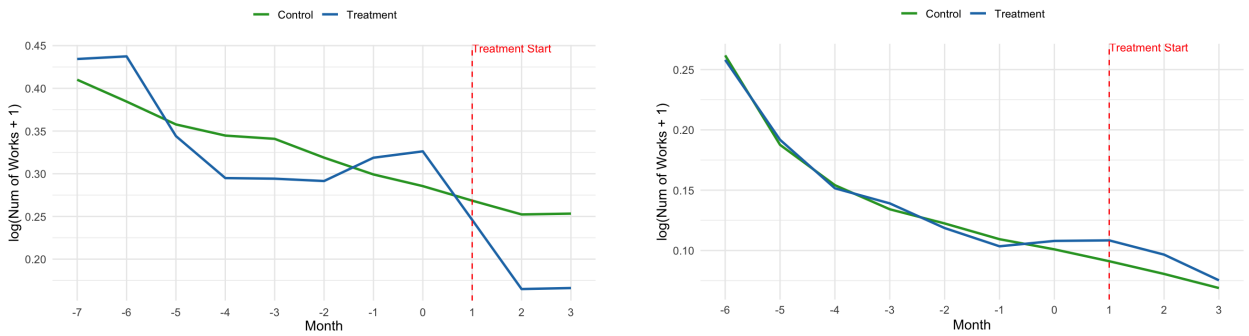
Our main results rely on the Temporal Causal Inference framework with Difference-in-Differences estimation and log-transformed outcome variables. To ensure robustness, we conduct several sets of robustness checks that address potential concerns about estimation method, identification strategy, outcome measurement, functional form, and sample selection. All robustness checks yield results consistent with our main findings, which strengthens confidence in our conclusions.

First, we apply Generalized Synthetic Control (GSC) as an alternative estimator within the TCI framework. Second, we implement a Regression Discontinuity in Time (RDiT) design (Cavusoglu et al. 2016, Hausman and Rapson 2018, Shi et al. 2022), which exploits the discontinuity at the policy cutoff rather than tenure-based control groups. Third, we use a Panel Difference Estimator (Goldberg et al. 2024), which compares year-over-year changes for the same creators. Fourth, motivated by concerns about log-like transformations $\log(Y + 1)$ (Chen and Roth 2024), we use an alternative binary outcome variable ("active" indicator). Last, we replace the linear specification with Poisson pseudo-maximum likelihood (PPML) estimation. These checks provide strong support for our main findings.

G.1 Generalized Synthetic Control

As an alternative to DiD, Generalized Synthetic Control (GSC) (Xu 2017) constructs a synthetic control group as a weighted combination of control units that best matches the treatment group's pre-treatment trajectory. While DiD assumes parallel trends between treatment and control groups, GSC relaxes this assumption by allowing for interactive fixed effects. We apply GSC within the same TCI framework, using the same treatment–control pairs constructed by cohort matching.

Table G.1 presents GSC estimates. The results are directionally consistent with our DiD estimates in Table 2: negative treatment effects for Lofter and positive effects for Graffiti Kingdom. Figure G.1 visualizes the treatment and synthesized control group trends. We note that the pre-treatment fit is imperfect. For Lofter, the treatment and control groups show some divergence in pre-treatment levels; for Graffiti Kingdom, a small gap appears in period 0, the month before the policy was implemented. Despite these limitations, the directional consistency between GSC and DiD estimates provides reassurance for our main findings.



(a) Lofter (b) Graffiti Kingdom

Figure G.1: GSC Trends: Treatment vs. Synthesized Control

Table G.1: Temporal Causal Inference Results Using Generalized Synthetic Control

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	log(Num of Works + 1)	
Treat × Post Month 1	-0.0280***	0.0173***
	[0.0065]	[0.0049]
Treat × Post Month 2	-0.0916***	0.0160***
	[0.0058]	[0.0050]
Treat × Post Month 3	-0.0910***	0.0064
	[0.0056]	[0.0043]
Estimation Method	Generalized Synthetic Control	
Creator Fixed Effects	Yes	Yes
Tenure Fixed Effects	Yes	Yes
Number of Creators	140,061	24,178

Note. The unit of analysis is the creator-month; Post Month 1–3 denote the first three post-treatment periods. Standard errors computed using nonparametric bootstrapping are in brackets. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

G.2 Regression Discontinuity in Time

TCI relies on tenure-based control groups and requires that creators at the same tenure exhibit comparable behavior across cohorts. RDiT offers an alternative identification strategy that does not require tenure matching. Instead, RDiT exploits the discontinuity at the policy implementation date and differences out unobservable factors that evolve smoothly over time.

We restrict the sample to three months before and three months after the policy change – a window sufficiently wide to observe effects yet narrow enough to enhance comparability.[20] Following Cavusoglu et al. (2016), we estimate:

$$\begin{aligned} Y_{it} = \lambda_i + \beta_1 \cdot AfterPolicy_t + \beta_2 \cdot Month_t \\ + \beta_3 \cdot AfterPolicy_t \times Month_t + \epsilon_{it}. \end{aligned} \tag{G.1}$$

where i indexes content creators and t indexes months (not tenure). The outcome variable Y_{it} is the log of the number of works uploaded by content creator i to the platform during month t. We code $Month_t$ as the number of months before or after the policy change, which serves as the running variable in the RDiT model (Cavusoglu et al. 2016, Hausman and Rapson 2018). $AfterPolicy_t = 1$ if month t falls within the post-treatment period, and 0 otherwise; λ_i represents the fixed effects of individual creator i, and ϵ_{it} is the error term. Our primary coefficient of interest, β_1, captures the magnitude of the discontinuous change in the outcome variable at the cutoff month. The interaction term $AfterPolicy_t \times Month_t$ allows us to fit different time trends on each side of the cutoff point (Cavusoglu et al. 2016).

Table G.2 presents RDiT estimates. For Lofter, the discontinuous drop at the policy cutoff is -0.1144 ($p < 0.001$). For Graffiti Kingdom, the discontinuous jump is $+0.0103$ ($p < 0.001$). These estimates align

[20] We therefore exclude users who created their accounts less than three months before the policy change, i.e., those who had not joined the platform before the start of our three-month pre-policy window.

closely with our TCI results in Table 2, which confirm negative signaling effects from Lofter's pro-AI policy signal and positive signaling effects from Graffiti Kingdom's anti-AI policy signal.

Table G.2: Regression Discontinuity in Time Results

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	log(Num of Works + 1)	
AfterPolicy	-0.1144***	0.0103***
	[0.0020]	[0.0029]
Month	0.0128***	-0.0050***
	[0.0008]	[0.0011]
AfterPolicy × Month	-0.0569***	-0.0098***
	[0.0011]	[0.0016]
Creator Fixed Effects	Yes	Yes
Adj. R-squared	0.0258	0.0015
Number of Creators	186,660	37,576
Observation Window Length	6	6
Number of Observations	1,119,960	225,456

Note. The unit of analysis is the creator-month. The Regression Discontinuity in Time analysis uses a six-month window, with three monthly periods before and three monthly periods after each policy cutoff (Lofter: March 6, 2023; Graffiti Kingdom: February 22, 2023). Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

G.3 Panel Difference Estimator

As a second alternative to TCI, the Panel Difference Estimator uses prior-year observations for the same creators as controls. This design absorbs seasonal variation (e.g., holiday effects) while remaining insulated from policy changes, which occurred only in 2023. The method is analogous to Difference-in-Differences and has been widely used in economics (Goldberg et al. 2024, Eichenbaum et al. 2024, Babina et al. 2024) and information systems (Burtch et al. 2023).[21]

We focus on creators observed in at least two consecutive years (2022 and 2023), using 2022 observations as controls and 2023 observations as treatment, aligned by calendar month. We estimate:

$$Y_{it} = \mu_i + \kappa_t + \sum_{m=1}^{3} \tau_m Treat_i \times PostMonth_{mt} + \epsilon_{it}. \tag{G.2}$$

where i indexes creators and t indexes calendar months; Y_{it} is the log of works uploaded; $Treat_i = 1$ for 2023 observations (zero for 2022); $PostMonth_{mt}$ is an indicator equal to one if period t corresponds to post-policy month $m \in \{1, 2, 3\}$ and zero otherwise; μ_i are individual fixed effects; κ_t are calendar month fixed effects; and ϵ_{it} is the error term. The coefficients τ_1, τ_2, and τ_3 capture the time-varying treatment effects in the three post-policy months.

Table G.3 presents estimates using both DiD and GSC. Results align closely with our TCI estimates in Table 2, again confirming negative signaling effects on Lofter and positive signaling effects on Graffiti Kingdom.

G.4 Alternative Outcome Measures and Functional Forms

Our main results use $\log(Y + 1)$ transformations. Chen and Roth (2024) show that if treatment operates on the extensive margin, re-scaling Y before the log transformation can yield treatment effects of any desired size. We address this concern using three approaches.

First, we conduct a scaling-sensitivity analysis using alternative offsets in the log transformation, replacing $\log(Y + 1)$ with $\log(Y + 2)$, $\log(Y + 3)$, and $\log(Y + 4)$. The corresponding estimates are reported in Table G.4. Reassuringly, the qualitative conclusions remain unchanged across all specifications: the estimated effects retain the same signs, remain statistically significant, and are broadly similar in magnitude to those reported in Table 2. These results suggest that our main findings are not driven by the particular choice of the +1 offset in the baseline log transformation.

Second, we replace the continuous outcome with a binary "active" indicator (= 1 if creator uploads at least one work in month t). We re-estimate Equation (1) under the TCI framework with this alternative outcome. Table G.5 reports the results, which show patterns similar to Table 2.

Last but not least, following Chen and Roth (2024), we replace the linear model with Poisson regression estimated via Poisson pseudo-maximum likelihood (PPML). PPML does not require the full Poisson distributional assumption (e.g., $Var(Y \mid X) = \mathbb{E}[Y \mid X]$) (Gourieroux et al. 1984b,a) and is recommended as a

[21] Also called the "long-difference estimator" in financial economics and macroeconomics.

Table G.3: Panel Difference Estimator Results Using Difference-in-Differences and Generalized Synthetic Control

	Lofter		Graffiti Kingdom	
	(1)	(2)	(3)	(4)
Dependent Variable	log(Num of Works + 1)			
Treat × Post Month 1	-0.0476**	-0.0482***	0.0628**	0.0568***
	[0.0207]	[0.0023]	[0.0238]	[0.0035]
Treat × Post Month 2	-0.1048***	-0.1320***	0.0903***	0.0568***
	[0.0234]	[0.0024]	[0.0295]	[0.0035]
Treat × Post Month 3	-0.0782***	-0.1395***	0.1476***	0.0555***
	[0.0234]	[0.0025]	[0.0295]	[0.0043]
Estimation Method	Difference-in-Differences	Generalized Synthetic Control	Difference-in-Differences	Generalized Synthetic Control
Creator Fixed Effects	Yes	Yes	Yes	Yes
Calendar Month Fixed Effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.4754		0.3762	
Number of Creators	156,841	156,841	39,799	39,799
Observation Window Length	26	26	26	26
Number of Observations	4,077,866	4,077,866	1,034,744	1,034,744

Note. The unit of analysis is the creator-month. The Panel Difference Estimator compares policy-year observations in 2023 with aligned calendar-month observations in 2022; Post Month 1–3 denote the first three post-policy months and their aligned 2022 comparison periods. For Difference-in-Differences, robust standard errors clustered at the creator level are in brackets; for Generalized Synthetic Control, bootstrapped standard errors are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

substitute for log-like transformations (Chen and Roth 2024, Silva and Tenreyro 2006). The specification is:

$$\mathbb{E}\left[Y_{it} \mid \text{Treat}_i, \text{PostMonth}_{mt}, \mu_i, \lambda_t\right] = \exp\left(\mu_i + \lambda_t + \sum_{m=1}^{3} \tau_{jm}\, \text{Treat}_i \times \text{PostMonth}_{mt}\right). \tag{G.3}$$

with variable definitions identical to those in Equation (1). Table G.6 reports PPML estimates. Results broadly align with our main findings in Table 2, indicating that log-like transformation concerns do not drive our results.

In summary, all robustness checks – alternative estimation methods (GSC), alternative identification strategies (RDiT, PDE), alternative outcomes (binary active indicator), and alternative functional forms (PPML) – yield results broadly consistent with our main findings.

G.5 Sample Selection

Finally, we conduct four additional sample-selection and comparison-group robustness checks, and report the corresponding regression results in Table G.7, which are qualitatively consistent with our main estimation results. First, to address the concern that our results may be driven by a small number of highly productive users, we re-estimate the baseline specification after excluding *super-users*, defined as the top 1% of users in each cohort ranked by total work volume (`work_sum`). We chose this cutoff because the data cover the entire

Table G.4: Scaling-Sensitivity Analysis under Temporal Causal Inference Using Difference-in-Differences

	Lofter			Graffiti Kingdom		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	log(Num of Works + offset)					
Log Offset	+2	+3	+4	+2	+3	+4
Treat × Post Month 1	-0.0379***	-0.0307***	-0.0261***	0.0192***	0.0155***	0.0131***
	[0.0046]	[0.0037]	[0.0032]	[0.0037]	[0.0030]	[0.0026]
Treat × Post Month 2	-0.0832***	-0.0669***	-0.0567***	0.0184***	0.0149***	0.0127***
	[0.0043]	[0.0035]	[0.0030]	[0.0037]	[0.0030]	[0.0025]
Treat × Post Month 3	-0.0814***	-0.0655***	-0.0554***	0.0108***	0.0086***	0.0073***
	[0.0042]	[0.0034]	[0.0029]	[0.0034]	[0.0028]	[0.0024]
Estimation Method	Difference-in-Differences					
Creator Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Tenure Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Creators	140,061	140,061	140,061	24,178	24,178	24,178

Note. The unit of analysis is the creator-month. Post Month 1–3 denote the first three post-treatment periods. Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table G.5: Temporal Causal Inference Results Using Difference-in-Differences with Active Outcome

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	Active	
Treat × Post Month 1	-0.0280***	0.0156***
	[0.0034]	[0.0028]
Treat × Post Month 2	-0.0646***	0.0143***
	[0.0031]	[0.0030]
Treat × Post Month 3	-0.0657***	0.0086***
	[0.0029]	[0.0026]
Estimation Method	Difference-in-Differences	
Creator Fixed Effects	Yes	Yes
Tenure Fixed Effects	Yes	Yes
Number of Creators	140,061	24,178

Note. The unit of analysis is the creator-month; Post Month 1–3 denote the first three post-treatment periods. Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table G.6: Temporal Causal Inference Results Using Poisson Pseudo-Maximum Likelihood

	Lofter	Graffiti Kingdom
	(1)	(2)
Dependent Variable	Num of Works	
Treat × Post Month 1	-0.2618***	0.1701**
	[0.0270]	[0.0673]
Treat × Post Month 2	-0.6067***	0.1235*
	[0.0322]	[0.0690]
Treat × Post Month 3	-0.5839***	-0.0736
	[0.0375]	[0.0715]
Estimation Method	Poisson Pseudo-Maximum Likelihood	
Creator Fixed Effects	Yes	Yes
Tenure Fixed Effects	Yes	Yes
Number of Creators	140,061	24,178

Note. The unit of analysis is the creator-month; Post Month 1–3 denote the first three post-treatment periods. Robust standard errors clustered at the creator level are in brackets. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

user population, most of whom post little or nothing in a given month. With so much of the distribution concentrated at the bottom, even the top 1% threshold reaches down to a modest 24 works a month on Lofter and 18 on Graffiti Kingdom, so a wider cut would remove ordinary active creators rather than additional super-users. Reassuringly, excluding the top 1% of users leaves our main results qualitatively unchanged, indicating that the estimated treatment effects are not driven by a small group of dominant contributors.

Second, we examine whether our findings are sensitive to the precise choice of nearby control cohorts. In the baseline design, for each treated cohort i, the control group is constructed from cohorts $i + 3$ to $i + 7$. As a robustness check, we instead use cohorts $i + 6$ to $i + 10$, namely cohorts that joined 6 to 10 months earlier, as the control group. This alternative specification allows us to assess whether our conclusions depend on the exact set of nearby control cohorts used in the baseline TCI construction.

Third, we address the concern that, under the baseline TCI design, treated and control cohorts may share a common calendar-time period during which both are active. Although such overlap does not contaminate the baseline design mechanically, it raises the possibility that the activity of one group may affect the other through channels not fully absorbed by tenure fixed effects. To address this concern, we construct an alternative control group that is further removed in entry time: for treated cohort i, we use cohorts $i + 11$ to $i + 15$, namely cohorts that joined 11 to 15 months earlier. This alternative comparison removes the common period between treated and control cohorts.

Last but not least, we examine whether the comparability of cohorts weakens when treated cohorts are matched to much earlier cohorts that may have faced different expectations, norms, or competitive conditions. To reduce exposure to such long-run calendar-time heterogeneity, we exclude early cohorts and restrict the analysis to the 24 cohorts from the most recent two years that pass the Granger and KS tests.

Taken together, these additional exercises strengthen the plausibility of the identifying assumptions underlying our TCI design and show that the main findings are robust to alternative sample restrictions and control-group constructions.

Table G.7: Sample-Selection Robustness Checks under Temporal Causal Inference Using Difference-in-Differences

	Lofter				Graffiti Kingdom			
	(1) Exclude Superusers	(2) Control Cohorts (i + 6 to i + 10)	(3) Control Cohorts (i + 11 to i + 15)	(4) Exclude Early Cohorts	(5) Exclude Superusers	(6) Control Cohorts (i + 6 to i + 10)	(7) Control Cohorts (i + 11 to i + 15)	(8) Exclude Early Cohorts
Dependent Variable	log(Num of Works + 1)							
Treat × Post Month 1	-0.0841***	-0.0843***	-0.0906***	-0.0717***	0.0239***	0.0251***	0.0320***	0.0295***
	[0.0073]	[0.0146]	[0.0234]	[0.0090]	[0.0047]	[0.0098]	[0.0114]	[0.0071]
Treat × Post Month 2	-0.1622***	-0.1613***	-0.1538***	-0.1283***	0.0221***	0.0221***	0.0271**	0.0286***
	[0.0071]	[0.0153]	[0.0261]	[0.0087]	[0.0043]	[0.0094]	[0.0111]	[0.0063]
Treat × Post Month 3	-0.1617***	-0.1552***	-0.1462***	-0.1243***	0.0119***	0.0099	0.0125	0.0184***
	[0.0071]	[0.0174]	[0.0270]	[0.0086]	[0.0042]	[0.0079]	[0.0101]	[0.0055]
Estimation Method	Difference-in-Differences							
Creator Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Creators	138,825	112,429	101,215	69,697	23,985	17,156	13,264	15,127

Note. The unit of analysis is the creator-month. Cohort i denotes the treated cohort in the control-window labels. The "Exclude Superusers" specifications drop the top 1% of creators in each cohort by total work volume, and the "Exclude Early Cohorts" specifications keep the 24 most recent validated treated cohorts. Robust standard errors clustered at the creator level are in brackets. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Online Appendix H Text Analysis: Methodology and Validation

General Text Analysis Procedure

The textual analysis of Lofter posts consists of three main steps, summarized as follows. First, we collect all posts from Lofter and then use an LLM (gpt-4o-mini) to (i) distinguish posts that express attitudes or opinions about Generative AI artworks from those that are irrelevant, and (ii) classify the attitude expressed in those posts that do discuss Generative AI. We implement this step using the following prompts:

```
SYSTEM_PROMPT = """\
You are a precise classifier for AI-generated content discussions.
Goal: Decide whether a text expresses an *opinion/judgment* about *AI creation* (
    especially AI art/AI drawing), and determine stance.

Definitions:
- OPINION_ON_AI_CREATION: The text conveys the author's opinion or evaluation (support/
    oppose/mixed/neutral) about AI creation. It may discuss ethics, copyright, plagiarism
    , creativity, skill, job market, fairness in contests, model/regulation, value
    judgment, personal attitude.
- NON_OPINION_AI: AI-creation-related but no opinion: tutorials, factual Q&A, news relay,
     tool setup, pure description without stance.
- IRRELEVANT_OR_FICTION: Unrelated to AI creation; or clear fiction/novel/fanfic/roleplay
    /character narratives, even if "AI" appears incidentally.

Output JSON fields (STRICT):
- label: one of ["OPINION_ON_AI_CREATION","NON_OPINION_AI","IRRELEVANT_OR_FICTION"]
- stance: one of ["pro","anti","neutral","mixed","n/a"]  # 'n/a' if not
    OPINION_ON_AI_CREATION
- rationale_zh: <= 40 chars, brief reason
- opinion_snippet: a short quote (<= 40 chars) that best evidences the judgment; empty if
     none

Be conservative: if mainly fiction/novel, choose IRRELEVANT_OR_FICTION.
If AI-related but purely instructional/news without stance, choose NON_OPINION_AI.
"""
```

Based on the outputs, we can classify the relevant posts into four categories: (i) dislike AI; (ii) pro AI; (iii) neutral; and (iv) mixed feelings. Then we use the LLM again to extract the detailed reasons why the content creators dislike Generative AI, with the following prompts:

```
SYSTEM_PROMPT = """\
You are a precise annotator. The input texts come from people who DISLIKE or OPPOSE AI
    creation (especially AI art/drawing).
Task: Extract the KEY REASONS (free-form, no fixed taxonomy).

Output STRICT JSON:
```

```
- reasons: array of 1-5 concise English phrases (2-8 words each) capturing the main
    reasons (e.g., copyright infringement, job displacement, unfair contests, low quality
    , privacy/portrait rights, opaque training, deepfake risk, censorship, high learning
    cost). Avoid redundancy.
- evidence_snippet: <= 40 characters quoting the best evidence from the text.
- confidence: number in [0,1].

Guidelines:
- Include only reasons for DISLIKING/OPPOSING AI creation. Be specific but concise.
- Prefer stable, generalizable phrases (avoid overly specific names/events).
- If no clear reason, return [] with low confidence.
"""
```

Last but not least, for each of the top three broad reasons (Replacement Risk, Low Quality, and Copyright Infringement), we apply the following prompt to extract more detailed and concrete underlying reasons:

```
SYSTEM_PROMPT = """\
You are a rigorous annotator. Focusing only on the given broad category {category},
    please extract no more than six specific reasons (short Chinese phrases) from the
    following posts, and provide the occurrence count (an integer) for each reason.

Return only a strict JSON object.
"""
```

Categorizing Posts by Target (March 2023 Sub-Analysis)

For the separate analysis discussed in Section 4.3, we restrict the sample to posts expressing negative attitudes toward Generative AI during the period in March 2023 slightly before and after the Generative AI feature was removed and classify them into three categories by what they primarily target:

- **(A) Directly feature-related**: posts that explicitly mention the "Profile Picture Generator" or refer to it through indirect phrases like "this tool" when the surrounding context clearly refers to the Generative AI feature. We rely on keyword searches for explicit references such as "profile picture generator" and "drawing machine", together with transparent keyword and context rules for indirect references such as "this tool" when the surrounding context clearly refers to the Generative AI feature.
- **(B) Platform-related but not feature-specific**: posts that mention the platform (using keywords such as "Lofter," or implicit expressions such as "the platform" when the discussion clearly concerns Lofter's deployment of the AI feature) but do not mention the feature itself.
- **(C) General negative attitudes**: posts that express general negative attitudes toward Generative AI without referring to either the feature or the platform.

When a post mentions both the tool and the platform in the same discussion, we classify it as category (A) rather than (B). Any post that does not meet the criteria for (A) or (B) is classified as (C).

Validation against Human Annotators

To verify the accuracy of our LLM-based classification, we recruited eleven human annotators to code a sample of posts drawn from NetEase Lofter. All annotators are native Chinese speakers and graduate students enrolled in top art academies in China, and they are active users of Lofter who are familiar with the platform's communities, norms, and terminology. The three annotation tasks are designed to validate our LLM-based classifications. The annotation tasks were organized into three datasets (each containing approximately 100 posts), and annotators were instructed to assign exactly one label per post unless otherwise specified. The coding instructions for each dataset are summarized below.

Dataset 1: Whether the post expresses an attitude toward AI art

Task. Determine whether the post explicitly expresses an attitude toward "AI Drawing/AI Art".

Labels.

- **1 = Yes**: The post clearly expresses a view or attitude toward AI Drawing/AI Art (e.g., liking, supporting, opposing, or worrying about AI Drawing/AI Art).
- **0 = No**: The post may mention "AI" or related content, but does not express an attitude toward AI Drawing/AI Art, or is unrelated to AI Drawing/AI Art.

Dataset 2: Overall stance toward AI Drawing/AI Art

Task. All posts in this dataset are confirmed to be about AI Drawing/AI Art. Judge the overall affective stance toward AI Drawing/AI Art.

Labels.

- **1 = Positive**: Clearly expresses liking, support, approval, anticipation, or other positive attitudes.
- **2 = Negative**: Clearly expresses dislike, opposition, concern, dissatisfaction, or other negative attitudes.
- **3 = Mixed**: Contains both clearly positive and clearly negative evaluations so that it cannot be simply classified as positive or negative.
- **4 = Neutral**: Describes facts, asks questions, or discusses issues in an objective tone without a clear affective stance.

Dataset 3: Specific reasons for disliking AI Drawing/AI Art

Task. All posts in this dataset express dislike of or opposition to AI Drawing/AI Art. Identify the *primary* reason for disliking or opposing AI Drawing/AI Art in the post.

Labels.

- **1 = Replacement Risk**: Concern that AI Drawing/AI Art will replace human artists, affecting jobs, income, or survival space.
- **2 = Low quality**: Belief that AI works are of poor quality (e.g., stiff style, anatomical errors, lack of emotion).

- **3 = Copyright infringement**: Concern or accusations that AI training or generation involves infringement (e.g., unauthorized use of others' works, "art theft", misattribution, or difficulty protecting original rights).

- **4 = Other**: Reasons that do not fall into the above categories or are too vague/ambiguous to classify.

After we collect the annotated data, we first assess agreement across the eleven human annotators and, for each dataset, derive a single gold label for every post using a simple majority-voting rule. We then compare the LLM's predictions to these human gold labels.

For Datasets 1–3, Fleiss' κ for the human annotators is 0.77, 0.55, and 0.68, respectively, indicating moderate to substantial agreement among coders on these subjective classification tasks. Relative to the resulting gold labels (Fleiss 1971, Landis and Koch 1977), the LLM achieves accuracies of 0.94, 0.92, and 0.81 in Datasets 1–3, respectively. Thus, for the first two tasks the LLM's classifications are very close to the consensus human coding, and even for the more fine-grained reason taxonomy in Dataset 3, the LLM attains a level of agreement with the gold standard that is comparable to typical human–human reliability in similar multi-category content-analysis settings.