Viewer Traffic Allocation for Small Creator Development: Experimental Evidence from Short-Video Platforms

Qinlu Hu
Chinese University of Hong Kong

Ni (Nina) Huang
University of Miami

Renyu (Philip) Zhang
Chinese University of Hong Kong

Abstract

To maximize viewer engagement, online short-video-sharing platforms like TikTok often prioritize the established, star creators in viewer traffic allocation. However, small creators, who have a smaller number of followers but constitute the majority of the creator pool, are vital to these platforms, bringing diverse content and fresh perspectives. In collaboration with a leading online short video sharing platform, we leverage two large-scale randomized field experiments to evaluate the value of small creators to viewer engagement and examine the efficacy of viewer traffic allocation on small creator development. In the viewer-side experiment, 50% of the videos from small creators were removed from the recommendation algorithm's candidate pool for treatment viewers, but the control viewers received no intervention. We find that, on average, reducing the supply of small creators' content benefited viewer engagement on the platform. Furthermore, such treatment increased video watching time for less active viewers but decreased it for active viewers. In the creator-side experiment, the viewer traffic for treatment creators was boosted via recommendation algorithms, but not for control creators. Compared to the control creators, the treated creators produced 5.87% more videos without compromising content quality, exerted more effort in video production and self-engagement with their content, and enjoyed a 289.81% increase in the new follower count growth. The effect of additional viewer traffic was more pronounced for more experienced, popular small creators. Altogether, our research offers actionable insights into viewer traffic allocation that balances viewer engagement and creator development on short-video platforms.

Keywords: Online video sharing platforms, traffic allocation, digital content creators, content creation, creator development, viewer engagement, field experiments

"*For the first time in history, hits and niches are on equal economic footing, both just entries in a database called up on demand, both equally worthy of being carried.*"

– Chris Anderson (2006) *The Long Tail: Why the Future of Business is Selling Less of More*

1 Introduction

Short-video sharing platforms such as TikTok, Triller, and Snapchat Spotlight have become increasingly popular in recent years (Shutsko, 2020). For example, as of 2024, TikTok has built an extensive global user base, boasting over 1.5 billion users worldwide, with approximately 1.1 billion monthly active users.[1] TikTok users on average dedicate more than one and a half hours (i.e., 95 minutes) per day to the app, and the platform records billions of video views per day, reflecting its exceptional ability to attract and maintain user attention and engagement.[2]

Central to the success of online short video sharing platforms is an effective viewer traffic allocation mechanism, enabled by personalized recommendation algorithms, which effectively allocates viewer traffic to user-generated video content with matched behavioral characteristics and individual preferences.[3] However, given the primary objective being maximizing user engagement (Besbes et al., 2016; Rafieian, 2023), these recommendation algorithms tend to disproportionately favor established creators who have already amassed a large number of followers (Qian and Xie, 2022), often at the expense of less known talents in the long tail (Anderson, 2006) who seek to organically accumulate audience bases.

This phenomenon, often referred to as the "popularity bias", creates a two-sided challenge for both viewers and creators (Abdollahpouri et al., 2019). For viewers, while consuming more content from established and popular creators may maximize their engagement, yet such arrangement hinders their content discovery and broadening of interests, potentially diminishing the viewers' utility in the long run. This is because the recommended content may not fulfill their complete spectrum of tastes (Nguyen et al., 2014; Moeller et al., 2018). For the creators, there is intense competition for viewer traffic, with the recommender system frequently channeling a larger share of views to established and popular creators. As a consequence, emerging and small creators face significant barriers in gaining visibility and building their audience (Pallais, 2014).

Following the long tail theory that points to the potential value of the non-popular mass (Anderson, 2006;

[1] "TikTok Revenue and Usage Statistics (2024)," Business of Apps
[2] "TikTok Statistics For 2024," Demandsage
[3] "How TikTok Reads Your Mind," The New York Times

Elberse, 2008), we focus on the understanding of less-known, small creators who face higher barriers than the star creators to break through the virtual attention clutter. We define *small creators* as those have historically attained fewer than one thousand followers, constituting about 99% of the creator population on our partner platform. Understanding small creators is crucial for content platforms for several reasons. First, small creators introduce competition to the content supply of the platform, thus preventing the over-domination of a few established stars and encouraging innovation (Dewi, 2021). Second, the long-tail small creators produce niche content and, therefore, enhance the content diversity of the platform, thereby creating values particularly for the viewers who prefer a broader range of content (Zhou et al., 2010a). Finally, compared to established stars, small creators often have a greater growth potential with a higher return on investment (ROI) (Jennings, 2022). By investing in and supporting small creators, platforms can identify and nurture the future stars, ensuring a healthy stream of new talent and high-quality content. Putting everything together, this work seeks to answer the following research questions:

- *What is the value of the content from lesser-known, small creators, in terms of user engagement, to the content platform?*
- *What is the efficacy of viewer traffic allocation in motivating and supporting small creators, to enhance their visibility and foster their development?*

To answer these research questions, we collaborate with a leading online short-video sharing platform in Asia (hereafter referred to as the *partner platform*). The partner platform systematically conducted two randomized field experiments in 2022. In this paper, we analyze the archival data provided by the partner platform, and report the findings from these experiments.

The first experiment was on the viewer side of the platform. In this experiment, the partner platform randomly assigned viewers into two groups: A treatment group, where 50% of the videos from small creators were removed from their candidate video pool feeding the recommendation algorithm, and a control group without such intervention (see Figure 2). Our main analysis reveals that, on average, removing 50% of the content produced by small creators significantly increased the viewers' video watching time by 0.33% in the treatment group, compared to those in the control group. In other words, reducing the supply of content from small creators appears to be beneficial for viewer engagement on the platform. What's more, limiting small creators' content supply had heterogeneous effects on viewer engagement with respect to different viewer activeness (quantified by the total number of video views one week prior to the experiment). Specifically, as a result of reducing content supply from small creators, the video watching time of less active viewers will increase but that of active viewers will decrease. Further investigation reveals that such a divergence in treatment effects can be attributed to the preference of active viewers for content diversity, who are, therefore, more adversely affected by the reduction of content supply by small creators. Our results suggest that exclusively focusing on delivering the content by popular big creators better engages the less active viewers, but hurts those who are dedicated and active. The active viewers contribute to most of a platform's video viewership and are, therefore, critical to platform success.

The second experiment was on the creator side of the platform. In this experiment, the creators on the partner platform were randomly divided into two groups: A treatment group receiving boosted viewer traffic with the enhanced weight of their content for more prominent rankings in the recommendation algorithm, and a control group without such an intervention. Our results show that the small creators in the treatment group, who received 265.65% more exposure to viewers than the control creators, significantly increased their video production quantity by 5.87%, without compromising the quality of the videos produced. Additionally, the small creators in the treatment group exerted more effort in content production, with a 2.89% increase in the number of days a video was uploaded, a 36.65% increase in self-reply rate, and an 18.16% increase in self-comment rate. Furthermore, the treated small creators enjoyed a 289.81% higher new follower count increase compared to the control creators. Moreover, the effect of increasing viewer traffic varies by creator popularity (measured by the number of content viewership prior to the experiment). The positive impacts of additional viewer traffic on creators' content production quantity, content creation effort, and new follower growth were more salient for more popular small creators.

Our study contributes to the literature on digital content creation by focusing on the small, emerging creators who received limited research attention in the literature (Qian and Xie, 2022). Previous studies have examined various factors that motivate digital content creation (Tang et al., 2012; Zeng et al., 2023), the heterogeneity among creators (Carroni et al., 2024; Foerderer et al., 2023), and the operational mechanisms of online video platforms (Jain and Qian, 2021; Qian and Jain, 2024). Furthermore, the existing literature has primarily focused on the star creators with a substantial number of followers and high visibility (Qian and Xie, 2022). However, the skewed attention towards star creators creates a self-reinforcing loop that exacerbates the disparities between star and small creators. This imbalance may intensify platforms' reliance on star creators and stifle the growth and visibility of non-star creators (Qian and Xie, 2022). Moreover, prior analytical work suggests that investing in smaller creators can yield long-term benefits for platforms, potentially leading to increased market thickness (Bhargava, 2022). Our study fills this gap in the literature by enhancing the understanding of small creators' value in viewer engagement and examining how viewer traffic allocation helps motivate them. Additionally, our explorations of heterogeneous treatment effects (HTE) deliver actionable insights for balancing viewer engagement and creator development in user traffic allocation.

This work also contributes to the growing body of work on the effective management of algorithms (Zhou et al., 2010a; Hosanagar et al., 2014; Qian and Jain, 2024). While previous research has demonstrated the value of personalized recommendation algorithms in enhancing user engagement and content consumption (Claussen et al., 2019; Holtz et al., 2020), it has also documented the unintended consequences of algorithmic personalization, such as creating "filter bubbles" and "popularity bias" (Anderson et al., 2020; Abdollahpouri et al., 2019). To address these concerns, scholars have focused on developing and refining recommendation algorithms (Zhou et al., 2010b; Chen et al., 2024), as well as employing analytical models to study various factors related to recommender systems (Li et al., 2018; Qian and Jain, 2024). Furthering this line of research, our study provides novel field evidence on the effectiveness of viewer traffic allocation on balancing viewer engagement and creator development in the context of online video sharing platforms, contributing to the empirical investigations of algorithm management (Srinivasan and Sarial-Abi, 2021; Chen et al., 2023). In particular, we focus on improving our understanding of how to manage recommendation algorithms to allocate viewer traffic by only adjusting their input data (i.e., the accessibility and weight of the content produced by small creators), without directly modifying the underlying algorithmic structures.

The findings of this work also offer actionable, practical implications for the managers, creators, and viewers of online content platforms. Platform managers can directly implement the findings of our study to support small creator development. Specifically, by segmenting viewers based on their activeness and tailoring content recommendations accordingly, the platform can improve the overall viewer engagement. Furthermore, small creators can benefit from the increased visibility and platform support through the recommendation algorithms boosting viewer traffic allocation. In turn, relatively more experienced and popular small creators should focus on creating engaging content to attract and retain viewers, capitalizing on the boosted viewer traffic and growth opportunities. Continuously monitoring and optimizing viewer traffic allocation based on viewer and creator feedback can help maintain a balance between short-term viewer engagement and long-term creator and platform growth. Our work offers insights into developing more diverse, engaging, and inclusive ecosystems for online content platforms.

2 Related Literature

Our paper is primarily connected to two streams of research in the literature: (a) digital content creation and (b) algorithm management. We contribute to the intersection of these literature streams by unveiling new insights into how algorithmic viewer traffic allocation helps balance viewer engagement and creator development on short-video platforms.

2.1 Digital Content Creation

Our work is related to the body of research on digital content creation. Prior Information Systems (IS) literature has extensively explored content creation in text-based contexts, such as, online communities, online reviews websites and knowledge-sharing platforms. For example, Feng et al. (2019) demonstrated that firms adjust pricing based on online reviews and word of mouth. Liu et al. (2017) analyzed how online reviews and past sales jointly influence consumer decisions. Sun et al. (2019) assessed the impact of review informativeness on perceived helpfulness. Zhang and Zhu (2011) investigated group size effects on contributions to Wikipedia, and Greenstein et al. (2021) studied ideological moderation in U.S. political articles on Wikipedia.

Meanwhile, there is a growing interest in understanding digital content creation in the format of online videos. Specifically, extant work recognizes the distinct characteristics of video content, such as its rich media (Zhao et al., 2021), creation effort (Zhao et al., 2023), and algorithm-based recommendation (Qian and Jain, 2024), etc., which set it apart from text-based content. Prior literature on video content creation can be grouped into three main themes: Motivating content creation, understanding creator heterogeneity, and platform mechanism design.

To start with, scholars explored various factors in motivating video content creation. For instance, Liu and Feng (2021) build a theoretical model to comprehensively investigate the underlying mechanisms of monetary incentive for online content platforms. Tang et al. (2012) found that shared advertising revenue effectively motivates video contributions. However, Liu et al. (2022) suggested that financial incentives do not consistently improve video quality or platform profitability. Additionally, Zeng et al. (2023) reported that ”social nudges,” where users encourage content creators, can significantly boost video production by fostering a sense of community and appreciation.

Next, a group of studies seek to understand the heterogeneity of creators and its implications to the content platforms. Carroni et al. (2024) developed an analytical model distinguishing superstar creators from amateurs in two-sided markets, examining the impact of exclusive premium content on platform competition, complementor participation, and market outcomes. In addition, Qian and Xie (2022) create a theoretical model to analyze competition between star and non-star creators in digital markets, empirically testing it using Twitch.tv data. They found that a star creator's exit reduces content provision, consumption, and non-star creators' effort levels, with effects lasting several months. Similarly, Foerderer et al. (2023) study how a star creator's departure affects remaining creators' content production. Using a quasi-experimental design, they found a 16.7-20.4% reduction in peer creators' contributions, driven by downsizing rather than leaving the platform. The effect was weaker for creators with diversified content portfolios or larger followings.

Furthermore, another body work employed a market design perspective, using analytical models to investigate the mechanisms of online video platforms. For example, Bhargava (2022) explore the economics of three-sided platforms connecting consumers, content creators, and advertisers. They analyzed how factors like creator heterogeneity, revenue sharing rates, advertising policies, and platform design choices influence outcomes such as content supply, consumer demand, advertising revenue, creator participation, and market concentration. Jain and Qian (2021) examine how platforms like YouTube and Twitch should design revenue-sharing plans to incentivize high-quality content creation, concluding that optimal strategies depend on competition among creators, user base size, and user donations. Qian and Jain (2024) investigate how recommender systems can motivate independent content creators to produce high-quality content. They suggested that platforms benefit from biasing recommendations towards high-quality content, even if it mismatches some users' preferences, as it can intensify competition and improve content quality.

Adding to the prior research on digital content creation, particularly in the context of video content platforms, this study focuses on the small, less known creators, who have received limited research attention in the literature. Prior analytical work suggests that investing in small creators can have long-term benefits for the platform, as it can lead to increased market concentration (Bhargava, 2022). Yet, as pointed out by Qian and Xie (2022), the disproportionately imbalanced attention towards star creators establishes a self-reinforcing loop, exacerbating the disparity between stars and non-stars. And this inherent skewness in the digital content market may intensify its reliance on star creators, potentially hindering and impeding the growth of non-star creators. Therefore, our work seeks to address this gap in the literature by improving our understanding on the value of small creators, as well as the effectiveness of motivating small creators via viewer traffic allocation.

2.2 Algorithm Management

This study also connects to the stream of work on the management of algorithms, particularly, recommender systems. Online content platforms typically operate on an advertising-based revenue model – offering free content to users while generating revenue from advertisers (Bhargava, 2022). In this ad-based revenue model, recommendation algorithms play a crucial role in driving content consumption and user engagement, which in turn, directly impacts the platform's ability to attract advertisers and generate revenue (Qian and Jain, 2024).

Previous research has demonstrated the efficacy of personalized recommendation algorithms in enhancing the content consumption of users on digital platforms. For instance, Zhou et al. (2010a) find that YouTube's recommender system increases the diversity of video views. Claussen et al. (2019) observe higher user clicks with personalized recommendations compared to human-curated ones, indicating that algorithmic recommendations better capture user preferences. Additionally, Holtz et al. (2020) report a significant increase in podcast streams after switching from popularity-based to personalized recommendation algorithms.

Meanwhile, scholars have been seeking to investigate unintended consequences of personalized recommendation algorithms, such as algorithmic biases. Lambrecht and Tucker (2019) find gender-neutral STEM career ads were shown more to men due to algorithms optimizing for cost-effectiveness, not intentional discrimination. Other studies, like Brown et al. (2022), discuss how algorithms contribute to polarization and "echo chambers" by recommending content that aligns with users' existing beliefs. Anderson et al. (2020) conduct a field experiment on Spotify, revealing that personalized recommendations reinforced users' existing preferences, reducing content diversity and exploration.

To address these challenges, scholars have actively engaged in research on enhancing the efficacy and mitigating the drawback of these algorithms. A stream of literature has focused on the development and refinement of recommendation algorithms to promote more diverse and balanced content exposure. For example, Chen et al. (2024) introduce a deep learning approach for recommending background music for short videos, using user-music and video-music matching with attention-based aggregation, outperforming existing models.

Besides the technical work on algorithm development, another body of research employ theoretical, analytical approaches and explore the management aspect of algorithms. Those studies focus on optimizing various factors related to existing recommender systems to better serve the interests of different stakeholders, such as platforms, content creators, and users, without altering the underlying algorithms. For instance, in the context of e-commerce platforms, Li et al. (2018) create an analytical model for e-commerce platforms, suggesting that a mildly profit-oriented recommendation strategy balances the platform's financial interests and consumer welfare. Additionally, Qian and Jain (2024) develop a game-theoretic model to study how recommender systems and revenue-sharing plans impact content creation, profits, and welfare on online short video platforms. They found that biased recommendations favoring high-quality content incentivize creators to invest in quality, resulting in a win-win scenario for both the platform and creators.

Extending the prior technical and analytical work, some scholars report empirical investigations on algorithm management that accounts for the various constraints, uncertainties, and contextual factors present in practical, real-world settings. For example, Cui et al. (2018) use Facebook data to show that incorporating social media information enhances sales forecast accuracy and provided recommendations for integrating such data into forecasting. Similarly, Sun et al. (2022) examine the impact of incorporating human deviations into bin packing algorithms on packing time and operational efficiency. They conducted a large-scale randomized field experiment with the Alibaba Group and found that including human deviations improved the optimization algorithms. Additionally, Bai et al. (2022) assess the impact of algorithmic versus human-based task assignment on fairness perceptions and productivity in another Alibaba experiment, finding that algorithmic assignment improves both. These studies contribute empirical evidence on effectively managing recommendation algorithms to maintain viewer engagement and support small creators.

3 Research Context

Our research context is a prominent online video-sharing platform in Asia (hereafter referred as "the partner platform"),[3]which has a substantial user base, including hundreds of millions of daily active users (DAU) and close to a billion monthly active users (MAU). The partner platform primarily features short videos, with durations ranging from a few seconds up to a few minutes. Our study focuses on the algorithm-driven "recommendation page" (see Figure 1) [4].This page displays a seamless scrolling feed of videos tailored to viewers based on their interests, watch history, and engagement behaviors, helping users discover new videos and creators.

The platform functions as a two-sided market with content creators and viewers. It has millions of active creators producing a wide range of videos. A main goal of content creators is to generate viewership and engagement for their content, and the partner platform shares revenue with them based on consumption metrics. Viewers access the partner platform freely to consume extensive video content spanning entertainment, education, sports, lifestyle, and other verticals. At the same time, the platform employs a state-of-the-art recommendation algorithm to distribute content to viewers, matching viewer interests with content, based on past viewing behavior, search queries, engagement and related data.[5] By matching viewers and creators, The partner platform generates revenue from video consumption through advertising and sponsorship. A portion of the platform revenue gets shared with creators, incentivizing them for continued content production.

We obtained and analyzed proprietary data from two randomized field experiments conducted by the partner platform. Specifically, our results encompass findings from: a) A viewer-side experiment, wherein the partner platform randomly assigned and implemented treatments to the viewers, and b) a Creator-side experiment, in which the partner platform randomly assigned and implemented interventions to the creators. These complementary field experiments offer valuable insights into our research questions on: i) the value of small creators to viewer engagement on the platform, and ii) the efficacy of leveraging viewer traffic allocation in fostering creator development.

[3]The name of the partner platform is concealed due to none-disclosure agreement (NDA).

[4]To protect the identity of the partner platform, we use TikTok's user interface (UI) as an example. Figure 1 presents a screenshot of the UI on TikTok, with key UI features labeled in the figure. The partner platform's UI features are very similar to those of TikTok.

[5]The partner platform deploys an industry-standard recommendation algorithm, which is commonly used by short video sharing platforms. Appendix A illustrates further details of the recommendation algorithm on the platform.

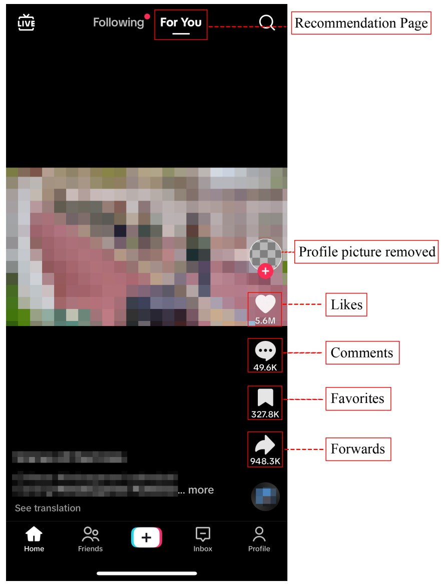

Figure 1: Screenshot of the User Interface (UI) on TikTok

4 Viewer-Side Study

4.1 Experimental Design

The partner platform performed the viewer-side experiment on December 2-12, 2022. The experiment employed a between-subjects design. The experimented viewers were randomly assigned into two groups: A treatment group, where 50% of the videos produced by small creators were blocked from the candidate video pool for each viewer, and a control group without such intervention. In other words, the treatment manipulation involves the partner platform randomly blocking content from the pool of candidate videos that could have been recommended to treated viewers.[6] To elaborate, the video supply from small creators, who have cumulatively attained fewer than one thousand followers, were randomly blocked from entering the pool of candidate videos that could be recommended to treated viewers. This approach ensured that the blocked content was removed before being distributed by recommendation algorithm. In this way, the treatment intervention does not directly modify the recommender system itself but only pertains to manipulating the video supply. We summarize the design of the viewer-side experiment in Table 1. We also present the flow of the viewer-side experiment in Figure 2.

[6]The content creators and viewers were unaware of this manipulation on the partner platform.

Table 1: Design of the Viewer-side Experiment

Experimental Group	Manipulation	Number of Observations
Treatment	Block 50% of small creators' video supply from the candidate video pool of each viewer.	4,532,431
Control	N/A	9,065,114

Notes: Sample size across the treatment and control groups $N > 13 million$

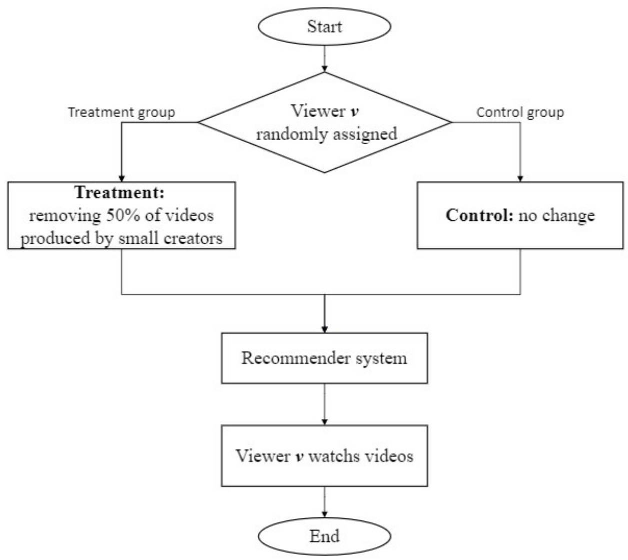

Figure 2: Flow of the Viewer-Side Experiment

4.2 Data

We acquired the viewer-level data on the user profile and behavioral characteristics of the viewers who participated in the viewer-side experiment. The viewer-level data contain the viewers' demographic and behavioral records aggregated for each individual during the experiment. Table 2 presents the definitions of the key variables and Table 3 shows the descriptive statistics of these variables in our viewer-level data. In addition, we report the randomization checks in Appendix B.

4.3 Main Analysis

We first test the main effect of removing the content from small creators. To recap, the treatment involves, for each viewer, excluding 50% of the videos produced by small creators from entering into the viewer's recommendation pool. Our main outcome of interest is the viewers' the total video watching time during the experiment ($WatchingTime$), as a critical measurement of viewer engagement for the online short-video

Table 2: Variables Description of the Viewer-Level Data in the Viewer-side Experiment

Variable	Description
Experimental Group	
Treatment	A viewer's assignment of the experimental group (0 = Control, 1 = Treatment)
Viewer Demographics	
Gender	A gender indicator (0 = Unkonwn,1= Male, 2 = Female)
Age	A numerical integer that captures a viewer's self-reported age.
CityLevel	The tier of a viewer's located city (0 = Unknown, 1 = First-tier city, 2 = Second-tier city, 3 = Third-tier city, 4 = Fourth-tier city, 5 = Fifth-tier and over Fifth-tier city).
NumFriend	The total number of friends that a viewer has on the platform.
NumFollower	The total number of followers that a viewer has on the platform.
Active	A binary indicator on whether a viewer is considered active (0 = Less Active, 1 = Active). "Active viewers" are defined as individuals whose number of video views in the week preceding the start of the experiment exceeds a certain threshold.
Behavioral Data at the Viewer level	
WatchingTime	The total duration of video content watched by a viewer during the experiment (in seconds).
NumView	The total number of videos watched by a viewer during the experiment.
NumComment	The total number of comments a viewer posted on videos during the experiment.
NumForward	The total number of videos a viewer forwarded to others during the experiment.
NumLike	The total number of likes a viewer left on videos during the experiment.
NumCategories	The total number of categories of videos watched by a viewer during the experiment.

platform.

We run an ordinary least squares (OLS) regression to estimate the treatment effect of removing 50% of the content produced by small creators, following the model specification Eqn. (1). where $Treatment_v$ is a binary variable indicating whether viewer v was in the treatment (*vs.* control) group. Table 4 presents the regression results.

$$WatchingTime_v = \beta_0 + \beta_1 \times Treatment_v + \varepsilon_v, \tag{1}$$

As shown in Table 4, we observe a significant and positive treatment effect – removing 50% of the content

Table 3: Descriptive Statistics of the Viewer-Level Data in the Viewer-side Experiment

Variable	Obs.	Mean	Median	Std	Min	Max
Viewer Demographic						
Gender	13,597,545	1.48	1.00	0.52	0	2
Age	13,597,545	42.65	42.00	14.20	18	80
CityLevel	13,597,545	3.40	4.00	1.29	0	5
NumFriend	13,597,545	456.80	225.00	704.23	0	22,288
NumFollower	13,597,545	0.40	0.00	6.24	0	4,787
Active	13,597,545	0.04	0.00	0.20	0	1
Behavioral Data at the Viewer level						
WatchingTime	13,597,545	7,659.90	309.50	19,364.27	0	637,114
NumView	13,597,545	333.60	22.00	901.29	0	61,923
NumComment	13,597,545	0.22	0.00	3.45	0	550
NumForward	13,597,545	1.00	0.00	6.95	0	550
NumLike	13,597,545	2.24	0.00	13.47	0	550
NumCategories	13,597,545	11.82	8.00	11.69	0	38

Notes: (a) The partner platform scaled the raw values of the variables ($NumFriend$, $NumFollower$, $WatchingTime$, $NumView$, $NumComment$, $NumForward$, $NumLike$, $NumCategories$) to protect data confidentiality. As a result, the numerical values presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) In compliance with the partner platform's policy and to ensure responsible and ethical data practices, we have obtained data from viewers between the ages of 18 and 80. (c) All variables are measured at the viewer level.

produced by small creators, on average, increases viewers' total video watching time during the experiment ($\hat{\beta}_1 = 24.906$, $p < 0.05$), suggesting that less (more) supply of the content from small creators is beneficial (detrimental) to viewer engagement. Small creators on the partner platform were likely inexperienced and less-skilled to create high-quality videos that appeal to a broad audience. By removing 50% of the videos produced by these creators from the recommendation pool, the platform potentially redirected viewer traffic to more engaging content produced by more popular and experienced star creators. This reallocation of viewer traffic led to a .33%[7] increase in average video watching time compared to that in the control group without intervention. It is worth noting that a 0.33% increase in video watching time can be quite meaningful for a large-scale short-video platform. DataProt (2023) reports that TikTok users, on average, engages with the app for approximately 95 minutes (i.e., 5,700 seconds) daily. Therefore, a 0.33% increase in video watching time translates to an additional 18.81 seconds per day. Given that there are more than 300 million

[7] 0.33% = (24.906 / 7651.597) * 100%

daily active users (DAU) on the partner platform, the lift in total video watching time of the platform amounts to more than 179 years per day.

Table 4: Treatment Effect on Total Video Watching Time in the Viewer-side Experiment

Variable	***WatchingTime*** (1)
Intercept	7651.597***
	(6.480)
Treatment	24.906**
	(11.216)
Relative effect size (%)	0.33**
Observations	13,597,545
F test	4.931**

Notes: (a) The partner platform scaled the raw values of the outcome variable ($WatchingTime$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) The relative effect size represents the percentage change in the outcome variable due to the treatment. (c) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.4 Analysis of Heterogeneous Treatment Effects

Our main analysis suggests that, on average, the content produced by small creators appears to negatively impact viewer engagement. We next explore the potential heterogeneity of the main effect – whether different types of viewers will respond to the change of content supply from small creators differently. Previous research indicates that online platform users differ in their activeness and exhibit different behavioral patterns accordingly (Heinonen, 2011). Adapting these prior considerations to the context of the online video-sharing platform, we hypothesize that active viewers, who frequently engage with the platform and consume a greater amount of content, may have different preferences and responses compared to inactive viewers, who use the platform less frequently. Thus, we test how active viewers' response towards the treatment differs from that of inactive viewers.

We measure the activeness of a viewer using his/her total number of video views one week prior to the experiment (November 25, 2022 - December 1, 2022). Specifically, we define "active viewers" (resp. "less active viewers") as those whose number of video views in this week exceeds (resp. does not exceed) a certain threshold.[8] It is important to note that while active users constitute only 1.80% of the total viewer

[8] The partner platform did not share the exact raw value of the threshold due to data sensitivity.

population, they account for 49.5% of the total viewership and 42.09% of the total video watching time.

We estimate Eqn. (2) to quantify how active viewers' response towards the treatment differs from that of less active viewers, where $Active_v$ is a binary variable classifying whether viewer v was an active (*vs.* less active) viewer one week prior to the experiment.[9] The estimation results, reported in Column (1) of Table 5, reveal interesting insights: Blocking half of small creators' videos from the recommendation pool has different effects for active and less active viewers, increasing the watching time for the latter ($\hat{\beta}_1 = 24.691$, $p < 0.05$) and decreasing that of the former ($\hat{\beta}_3 = -86.045$, $p < 0.05$).

$$WatchingTime_v = \beta_0 + \beta_1 \times Treatment_v + \beta_2 \times Active_v + \beta_3 \times Treatment_v \times Active_v + \varepsilon_v, \quad (2)$$

Table 5: Heterogeneity by Viewer Activeness in the Viewer-side Experiment

Outcome Variable	***WatchingTime*** (1)	***NumCategories*** (2)
Treatment	24.691***	-0.002
	(8.831)	(0.006)
Active	58,431.645***	22.634***
	(24.597)	(0.016)
Treatment Active*	-86.045**	-0.059**
	(42.588)	(0.028)
Observations	13,597,545	13,597,545
F test	2,819,883.754***	924,462.065***

Notes: (a) The partner platform scaled the raw values of the outcome variables ($WatchingTime$, $NumCategories$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Removing small creators' videos and replacing them with content from established creators better engages the less active viewers but may have an opposite effect on active viewers. Less active viewers prefer high-quality content from well-known creators, while active viewers appreciate discovering and engaging with smaller creators. Therefore, to maximize viewer engagement, a platform should support and promote small creators to keep active viewers engaged and attract and retain less active viewers with the popular high-quality content from star creators. These findings highlight the importance of managing recommendation algorithms

[9]The continuous measurement of viewer activeness yields qualitatively consistent estimates.

to cater to varying viewer activeness and preferences.

4.5 Plausible Mechanisms

To uncover the underlying mechanisms for the observed HTE with respect to viewer activeness, we interviewed the product managers and data scientists of the partner platform. They commented that small creators typically produce more diverse and innovative content, catering to various viewer tastes, while popular star creators focus on the mainstream appeal. A product manager noted that, “Small creators are likely to publish innovative content that diverges from the platform’s mainstream trends.” Another added, “Small creators often venture into niche areas, unlike popular creators who target mass appeal.” The head of the Data Analytics team observed that, “The diversity of content from small creators likely surpasses that of other creators.”

Based on these insights from practitioners, we provide some plausible mechanisms for our empirical findings. First, the content of small creators might be perceived as low quality, explaining the positive effect of removing it on the overall viewer engagement. Second, small creators are more likely to produce diverse content. Removing such content reduces consumption diversity, which may affect the active viewers more than the less active ones. Thus, while the perceived lower quality of small creators may drive the overall negative engagement effect, the higher content diversity appeals more to active viewers. Consequently, removing small creators’ content disproportionately impacts the active (*vs.* the less active) viewers, explaining the observed heterogeneous treatment effects.

To test the aforementioned mechanisms, we obtained a dataset containing a random sample of 188,374,718 creators from the partner platform, observing the creators’ characteristics and content production behaviors over a 7-day period. This 7-day dataset allows us to compare the attributes of small creators with those of other creators. We first examine whether the content produced by small creators has lower perceived quality compared to that produced by other creators on the platform. To verify this, we use $CompleteViewRate$, defined as the proportion of the video views that finish watching at least 90% of the video’s total duration, as a measure of content quality. We present two pieces of evidence. First, a comparison of the $CompleteViewRate$ over the 7-day dataset indicates that small creators have a significantly lower $CompleteViewRate$ (15.70%) than other creators (36.02%) ($p < 0.01$). Second, an experimental comparison between the treatment and control groups shows that the treatment group has a significantly higher $CompleteViewRate$ (40.27%) than the control group (40.16%) ($p < 0.01$). These finding indicate that the video quality of small creators is likely lower than that of other creators, lending supporting evidence for the content quality mechanism.

Next, we investigate whether the diversity of the content produced by small creators differs from that of more established creators. To measure content diversity, we follow the prior literature and adopt the Herfindahl–Hirschman Index (HHI)[10] of video topics as the key metric (Chen et al., 2023). The HHI measures the concentration of the videos produced by a creator across different topics. A lower HHI value indicates a higher level of content diversity for the creator. Conversely, a higher HHI value suggests that a creator's content is more focused on a subset of topics, indicating lower content diversity. We then conducted a two-tailed t-test to compare the HHI of small creators to that of others, and the results show that the average HHI index of small creators (HHI = 0.10) is significantly lower ($p < 0.01$) than that of other creators (HHI = 0.15). This finding suggests that small creators contribute to the overall diversity of content on the platform by producing videos that cover a wider variety of topics, supporting the content diversity mechanism.

Furthermore, given that small creators produce content spanning a more diverse range of topics compared to other creators, we seek to explain why blocking small creators is detrimental to active viewers by examining whether the active viewers appreciate diverse content more than the less active ones. Thus, we analyze the viewer-side experiment data and compare the viewers' consumption diversity metrics between the treatment and control groups during the experiment. The outcome of interest is $NumCategories$, capturing the total number of categories of videos watched by a viewer during the experiment. We estimate the following Eqn. (4) to investigate how content consumption diversity varies between the active and less active viewers in the experiment:

$$NumCategories_v = \beta_0 + \beta_1 \times Treatment_v + \beta_2 \times Active_v + \beta_3 \times Treatment_v \times Active_v + \varepsilon_v, \tag{4}$$

where $Treatment_v$ is a binary variable indicating whether viewer v was in the treatment (*vs.* control) group. We report the estimation results in Column (2) of Table 5. We find that, for the less active viewers ($Active$ = 0), blocking content produced by small creators does not significantly influence the diversity of their content

[10]The HHI is calculated as follow:

$$HHI_c = \sum_{j=1}^{M} p_j^2 \tag{3}$$

where HHI_c indicates is the Herfindahl-Hirschman Index (HHI) for a content creator c, M is the total number of unique video topics produced by creator c, and p_j represents the proportion of videos that creator c has produced on topic j, relative to their total video production.

consumption. However, the coefficient of the interaction term, $Treatment \times Active$, is significant and negative ($\hat{\beta}_3 = -0.059$, $p < 0.05$), suggesting that the treatment leads to a decrease in content consumption diversity for the active viewers. In other words, by limiting the active viewers' access to small creators, the treatment also constrained the diversity of the content these active users consume. This finding helps explain why the treatment leads to decreases in viewer engagement for the active viewers – When these viewers are exposed to less diverse content due to the blocking of content from small creators, they may find the platform less engaging.

To conclude, the empirical results and mechanism explorations of our viewer-side experiment suggest that small creators are likely to produce lower-quality content, which could negatively impact the engagement of general audience. However, small creators also contribute to the platform by producing more diverse content. While such content may not appeal to the broader audience, it often provides a wider range of topics, perspectives and styles. Moreover, our findings reveal that the active viewers appear to especially appreciate content diversity. They value the opportunity to discover new voices and engage with content that diverges from the mainstream. For these active viewers, the presence of small creators enhances their overall experience on the platform and keeps them engaged. Bearing the above in mind, it is important for platforms to support and retain small creators. We next report a Creator-side experiment that examines the efficacy of viewer traffic allocation as a possible strategy to foster the development of smaller creators.

5 Creator-side Study

5.1 Experimental Design

The creator-side experiment sought to evaluate the impact of viewer traffic allocation, in the form of increased content exposure, on the performance and development of small creators. The partner platform performed the creator-side experiment from March 18 to April 1, 2022. The experiment focused on the small creators and employed a between-subjects design. The experiment participants were randomly assigned into the treatment and control groups. In the treatment group, the small creators received additional viewer traffic/content exposure through an enhanced weighting of the their content in the recommendation algorithm. Specifically, the treatment involves the partner platform adjusting parameters to increase the prominence of treated small creators' videos in the algorithm's ranking system, aiming to boost the exposure of treated small creators' videos, potentially leading to higher viewership. Conversely, the control group creators did not receive such an intervention, serving as a baseline for comparison. Table 6 summarizes the experimental design of the creator-side experiment, and Figure 3 presents the flow of the experiment.

Table 6: Design of the Creator-side Experiment

Experimental Group	Manipulation	Number of Observations
Treatment	Allocate additional viewer traffic/content exposure to small creators	480,048
Control	N/A	480,353

Notes: Sample size across the treatment and control groups $N > 960,000$

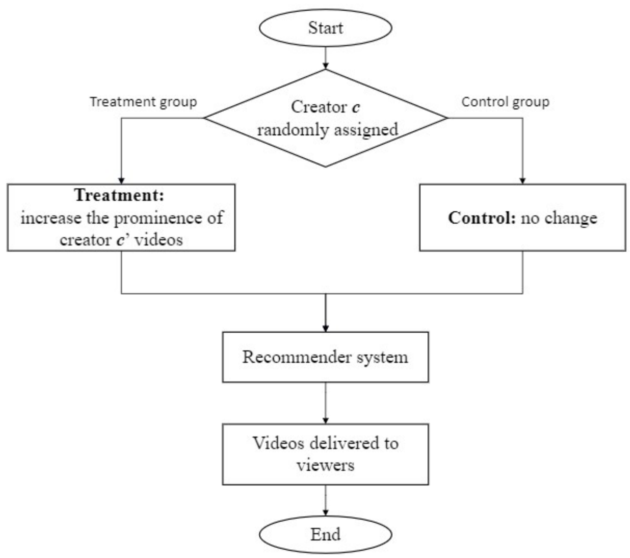

Figure 3: Flow of the Creator-side Experiment

5.2 Data

We obtained the data from the creator-side experiment (480,353 creators in the control group and 480,048 in the treatment group). The dataset contains the creators' demographic and behavioral records aggregated for each participant in the experiment. Table 7 presents the definitions of the main variables in the creator-side experiment, whereas Table 8 shows the descriptive statistics of these variables in the experiment data. Additionally, we report the randomization checks in Appendix B.

5.3 Main Analysis

The creator-side experiment enables us to evaluate the effect of allocating additional traffic/exposure to small creators. We perform the analysis from multiple perspectives, examining the impact of additional viewer

Table 7: Variables Description of the Creator-Level Data in the Creator-side Experiment

Variable	Description
Experimental Group	
Treatment	A creator's assignment of the experimental group (0 = Control, 1 = Treatment)
Creator Demographics	
Gender	A gender indicator (0 = Unknown, 1 = Male, 2 = Female)
Age	A numerical integer that captures a creator's self-reported age.
CityLevel	The tier of a creator's located city (0 = Unknown, 1 = First-tier city, 2 = Second-tier city, 3 = Third-tier city, 4 = Fourth-tier city, 5 = Fifth-tier and over Fifth-tier city).
NumFollower	The total number of followers that a creator has on the platform.
Popular	A binary indicator on whether a creator is considered popular (0 = Less Popular, 1 = Popular). "Popular creators" are defined as those who received historical viewer traffic exceeding a certain threshold defined by the platform prior to the experiment.
Content Production Quantity and Quality	
NumUploadVideos	The total number of videos uploaded by a creator during the experiment.
LikeRate	The ratio of likes to valid views, representing the percentage of viewers who liked a creator's videos during the experiment.
CommentRate	The ratio of comments to valid views, representing the percentage of viewers who commented on a creator's videos during the experiment.
ForwardRate	The ratio of forwards to valid views, representing the percentage of viewers who shared a creator's videos during the experiment.
Creator Effort	
ProducedVideoLength	The average length of the video produced by a creator during the experiment.
DaysOfProduction	The total number of days a creator produced videos during the experiment.
SelfReplyRate	The ratio of the total number of comment replies a creator makes to their own content to the total number of videos produced by the creator during the experiment.
SelfForwardRate	The ratio of the total number of times a creator forwards their own content to the total number of videos they produce during the experiment.
SelfCommentRate	The ratio of the total number of comments a creator makes on their own content to the total number of videos they produce during the experiment.
Creator Growth	
NumNewFollowers	The total number of new followers gained by a creator during the experiment.

Notes: (a) For content production quality ($LikeRate$, $CommentRate$, $ForwardRate$), a view is considered valid if the viewing duration surpasses a specified threshold determined by the partner platform. The exact raw value of the threshold was not shared by the partner platform due to data sensitivity. And the number of valid views accounts for about 78% of the total views.

Table 8: **Descriptive Statistics of the Creator-Level Data in the Creator-side Experiment**

Variable	Obs.	Mean	Median	Std	Min	Max
Creator Demographic						
Gender	960,401	1.63	2.00	0.48	0	2
Age	960,401	37.00	32.00	11.46	18	70
CityLevel	960,401	3.44	4.00	1.39	0	5
NumFollower	960,401	10.31	2.00	44.92	0	999
Popular	960,401	0.01	0.00	0.11	0	1
Content Production Quantity and Quality						
NumUploadVideos	960,401	3.06	0.00	17.88	0	4,690
LikeRate	182,753	0.07	0.00	0.18	0	1
CommentRate	182,753	0.00	0.00	0.04	0	1
ForwardRate	182,753	0.03	0.00	0.11	0	1
Creator Effort						
ProducedVideoLength	182,753	345.61	261.70	348.37	0	9,426
DaysOfProduction	182,753	11.07	5.23	10.89	5	73
SelfReplyRate	182,753	0.02	0.00	0.12	0	1
SelfForwardRate	182,753	0.37	0.00	0.44	0	1
SelfCommentRate	182,753	0.02	0.00	0.10	0	1
Creator Growth						
NumNewFollowers	960,401	9.58	0.00	1,571.53	-1,795	658,610

Notes: (a) The partner platform scaled the raw values of the variables ($NumUploadVideos$, $LikeRate$, $CommentRate$, $ForwardRate$, $ProducedVideoLength$, $DaysOfProduction$, $SelfReplyRate$, $SelfForwardRate$, $SelfCommentRate$, $NumNewFollowers$) to protect data confidentiality. As a result, the numerical values presented in the table do not hold any practical meaning and should not be interpreted as absolute values. (b) In compliance with the partner platform's policy and to ensure responsible and ethical data practices, we have obtained data from creators between the ages of 18 and 70. The upper age limit reflects considerations of creators' productivity, ensuring our analysis remains coherent. (c) For variables associated with creator's production effort ($ProducedVideoLength$, $DaysOfProduction$, $SelfReplyRate$, $SelfForwardRate$, $SelfCommentRate$) and production quality ($LikeRate$, $CommentRate$, $ForwardRate$), we only consider creators who produced at least one video during the experiment, so the sample size is smaller (182,753). (d) A negative $NumNewFollowers$ indicates a decrease in the number of followers. (e) All variables are measured in creator level.

traffic on small creators': (a) content production quantity and quality, (b) content creation efforts, and (c) creator growth. To estimate the average treatment effect of additional viewer traffic on these outcomes, we performed ordinary least squares (OLS) regression, following specification Eqn. (5), where $Treatment_c$ is a binary variable indicating whether creator c was in the treatment (*vs.* control) group. $OutcomeVariable_c$

represents the outcome variables of interest in our study.

$$OutcomeVariable_c = \beta_0 + \beta_1 \times Treatment_c + \varepsilon_c, \tag{5}$$

5.3.1 Content Production Quantity and Quality

We estimate Eqn. (5) to examine the impact of traffic allocation on the quantity of content produced by small creators and report the results in Table 9. Column (1) of Table 9 shows that allocating additional traffic to small creators has a significant and positive effect on their content production quantity ($\hat{\beta}_1 = 0.174$, $p < 0.01$). This result suggests that allocating extra viewer traffic to small creators by enhancing the weight of their content for more prominent rankings in the recommendation algorithm (the treated creators receive 265.65% more exposures than the control ones) can significantly lift their content production quantity by 5.87%.[11]

Table 9: Treatment Effect on Content Production Quantity in the Creator-side Experiment

Variable	***NumUploadVideos*** (1)
Intercept	2.964***
	(0.026)
Treatment	0.174***
	(0.036)
Relative effect size (%)	5.87***
Observations	960,401
F test	22.844***

Notes: (a) The partner platform scaled the raw values of the outcome variable ($NumUploadVideos$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) The relative effect size represents the percentage change in the outcome variable due to the treatment. (c) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Next, we test the treatment effect of additional traffic on small creators' content production quality, measured by $LikeRate$, $CommentRate$, and $ForwardRate$. Specifically, $LikeRate$ is the total number of likes a creator receives divided by the number of valid views for this creator.[12] Valid views are important for

[11] 5.87% = (0.174 / 2.964) * 100%

[12] A view is considered valid if the viewing duration surpasses a specified threshold determined by the partner platform. The exact raw value of the threshold was not shared by the partner platform due to data sensitivity.

the short-video platform, which represents the its viewer engagement. Valid views are the viewer impressions accurately allocated by the recommender system to appropriate content, reflecting the genuine viewer interests. In contrast, views falling short of the duration threshold are inaccurately recommended. Thus, $LikeRate$ effectively measures the interests a creator garners from genuinely engaged viewers. Similarly, the $CommentRate$ and $ForwardRate$ are the proportion of valid views at which viewers comment on and share the creator's content, respectively. These metrics objectively assess the creator's content appeal and the resonant value of their work among the audience, aiming for more comprehensive evaluations of content quality. We estimate Eqn. (5) and report the results in Table 10, which suggest that allocating additional traffic to small creators does not affect content quality, as indicated by the statistically insignificant coefficients ($p > 0.10$) of $LikeRate$ and $CommentRate$ in Columns (1) and (2), as well as a positive and marginally significant coefficient ($p > 0.05$) of $ForwardRate$ in Columns (3).[13]

These results collectively suggest that allocating more viewer traffic to small creators increases their content production quantity without compromising content quality. Our findings challenge the common belief that higher output may compromise quality, as our evidence shows that small creators are capable of maintaining their production standards even when the visibility is significantly enhanced. It is possible that, upon receiving more viewer traffic, the small creators with motivation and passion about their content are able to scale up their production without sacrificing quality. The treatment appears to unleash their untapped creative potential, allowing them to produce more content while still maintaining the same level of engagement from their audience.

5.3.2 Creator Effort

We employ five metrics are employed to measure the content production effort of a creator: (a) $ProducedVideoLength$, the average length of the videos produced by a creator during the experiment; (b) $DaysOfProduction$, the total number of days a creator produced videos during the experiment; (c) $SelfReplyRate$, the number of comment replies by the creator him/herself per video during the experiment; (d) $SelfForwardRate$, the number of times a creator forwarded his/her own content per video during the experiment; and (e) $SelfCommentRate$, the number of comments a creator makes on his/her own content per video during the

[13]We focus on evaluating the creators' content production quality; hence our analysis is restricted to creators who have produced at least one video during the experiment.

Table 10: **Treatment Effect on Content Production Quality in the Creator-side Experiment**

Variable	***LikeRate*** (1)	***CommentRate*** (2)	***ForwardRate*** (3)
Intercept	0.070***	0.004***	0.029***
	(0.000)	(0.000)	(0.000)
Treatment	0.001	0.000	0.001*
	(0.000)	(0.001)	(0.000)
Relative effect size (%)	0.81	-0.37	3.36*
Observations	182,753	182,753	182,753
F test	2.444	2.330	4.498*

Notes: (a) The partner platform scaled the raw values of the outcome variables ($LikeRate$, $CommentRate$, $ForwardRate$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) The relative effect size represents the percentage change in the outcome variable due to the treatment. (c) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

experiment.

We estimate Eqn. (5) using these metrics as outcome variables and report the results in Table 11. Although the treatment effects on $ProducedVideoLength$ and $SelfForwardRate$ are insignificant ($p > 0.10$), as shown in Columns (1) and (4), the effects of allocating viewer traffic on creators' content production effort are generally significant and positive ($p < 0.05$), as evidenced in Columns (2), (3) and (5). Specifically, allocating additional viewer traffic to small creators, on average, increases the number of production days by 2.89%[14], the self-reply rate by 36.65%[15], and the self-comment rate by 18.16%[16].[17] The increased number of production days suggests that creators are dedicating more time to creating content, possibly due to the increased motivation and perceived value of their work resulting from the extra exposure. Also, the substantial increases in self-reply and self-comment rates imply that creators are more actively engaging with their own content, potentially in an effort to foster viewer engagement with the created content. Altogether these results suggest that additional exposure of small creators' content to viewers serves as a powerful motivator, prompting them to spend more time and effort on content creation and actively engaging with their audience.

[14] 2.89% = (0.315 / 10.907) * 100%

[15] 36.65% = (0.006 / 0.016) * 100%

[16] 18.16% = (0.003 / 0.017) * 100%

[17] Our analysis focuses on assessing creators' content production effort; hence the sample is restricted to creators who have produced at least one video during the experiment.

Table 11: Treatment Effect on Creator Effort in the Creator-side Experiment

Variable	*ProducedVideoLength* (1)	*DaysOfProduction* (2)	*SelfReplyRate* (3)
Intercept	378.755***	10.907***	0.016***
	(1.156)	(0.036)	(0.000)
Treatment	-1.939	0.315***	0.006***
	(1.724)	(0.051)	(0.001)
Relative effect size (%)	-0.51	2.89***	36.65***
Observations	182,753	182,753	182,753
F test	1.265	38.279***	120.309***

Variable	*SelfForwardRate* (4)	*SelfCommentRate* (5)
Intercept	0.374***	0.017***
	(0.001)	(0.000)
Treatment	-0.002	0.003***
	(0.002)	(0.000)
Relative effect size (%)	0.53	18.16***
Observations	182,753	182,753
F test	0.912	28.646***

Notes: (a) The partner platform scaled the raw values of the outcome variables ($ProducedVideoLength$, $DaysOfProduction$, $SelfReplyRate$, $SelfForwardRate$, $SelfCommentRate$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) The relative effect size represents the percentage change in the outcome variable due to the treatment. (c) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5.3.3 Creator Growth

Next, we investigate how allocating additional viewing traffic to small creators influences their growth on the platform. We estimate Eqn. (5), using $NumNewFollowers$, the total number of new followers gained by a creator during the experiment, as the outcome variable. As shown in Table 12 Column (1), we observe a significant and positive average treatment effect on the number of new followers ($\hat{\beta}_1 = 11.334$, $p < 0.01$), a 289.81%[18] increase for the treated creators, compared to the control group. The substantial growth in new followers has several implications for small creators. Firstly, it motivates these creators to continue producing high-quality content, as they see a tangible reward for their efforts to expand the follower base. This positive feedback loop can encourage creators to invest more effort into content creation, potentially

[18] 289.81% = (11.334 / 3.911) * 100%

further improving the quantity and quality of their output. Additionally, the follower increase can also lead to greater engagement and interaction between creators and their audience, which can provide valuable feedback and support for content creators, helping them refine their content and tailor it to the preferences of their target audience. Furthermore, the increase in new followers can open up potential opportunities for small creators, such as sponsorships and partnerships with advertisers, providing them with avenues for future growth and monetization.

Table 12: Treatment Effect on Creator Growth in the Creator-side Experiment

Variable	***NumNewFollowers*** (1)
Intercept	3.911*
	(2.267)
Treatment	11.334***
	(3.207)
Relative effect size (%)	289.81***
Observations	960,401
F test	12.489***

Notes: (a) The partner platform scaled the raw values of the outcome variable ($NumNewFollowers$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) The relative effect size represents the percentage change in the outcome variable due to the treatment. (c) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5.4 Analysis of Heterogeneous Treatment Effects

Our main results show that allocating extra viewer traffic to small creators significantly boosts their content production quantity, content production effort, and new follower count. Next, we explore the heterogeneous treatment effects with respect to creators' past viewership.

On one hand, for the small creators with low past viewership, additional traffic can be particularly encouraging, motivating them to create more and improve content quality as they see their number of followers grow. Conversely, creators with consistently low viewership might not benefit as much, because they may lack the motivation or skills to capitalize on the increased exposure. On the other hand, for small creators with higher past viewership, the marginal impact of additional traffic might be less pronounced since they are used to having their content visible. However, these creators, who likely have confidence and experience in content creation, may still be motivated to produce more high-quality content with increased visibility. In short, the HTE of allocating more viewer traffic to small creators with respect to creator past viewership is unclear and requires further investigations.

To empirically test the aforementioned counterarguments, we use $Popular_c$ to classify creator c's past viewership. Specifically, we define *popular creators* (i.e., $Popular_c = 1$) as those who received a historical viewer traffic exceeding a certain threshold during the two weeks prior to the experiment (March 3, 2022 - March 17, 2022).[19] It is important to note that while popular small creators constitute only 0.83% of our sample, yet they contribute to 79.07% of the total viewership traffic and 17.85% of the total videos uploaded.

We estimate Eqn. (6) to investigate how popular creators' responses to the treatment differ from those of less popular creators. In Eqn. (6), $Treatment_c$ is the treatment indicator denoting whether a small creator c was in the treatment (*vs.* control) group, $OutcomeVariable_c$ represents the outcome variables of interest in our study, and $Popular_c$ is a binary variable indicating whether a small creator c was a popular (*vs.* less popular) one.[20]

$$OutcomeVariable_c = \beta_0 + \beta_1 \times Treatment_c + \beta_2 \times Popular_c + \beta_3 \times Treatment_c \times Popular_c + \varepsilon_c, \tag{6}$$

Table 13 reports the estimation results of the HTE on small creators' content production quantity with respect to creator popularity. Column (1) of Table 13 shows that the coefficient of the interaction term between $Treatment$ and $Popular$ is positive and significant ($\hat{\beta}_3 = 1.012$, $p < 0.01$). This indicates that the positive effect of allocating additional viewing traffic on content production quantity is more salient for popular small creators compared to their less popular counterparts.[21]

Next, we examine the potential HTE on creators' content creation effort by creator popularity. Table 14 shows that the coefficients of the interaction between $Treatment$ and $Popular$ are significant and positive ($p < 0.05$) in Columns (3) and (5), yet remain insignificant in Columns (1), (2) and (4). The results indicate that popular small creators, when given additional exposure, are more likely to increase their effort in terms of their engagement with their own content through self-replies and self-comments.

[19]To protect the sensitive data, the partner platform did not share the exact raw value of the threshold.
[20]The continuous measurement of creator popularity yields qualitatively consistent estimates.
[21]We do not observe significant heterogeneity in the treatment effects on the creators' production quality with respect to creator popularity.

Table 13: Heterogeneity by Creator Popularity on Production Quantity in the Prod.-Side Exp.

Variable	***NumUploadVideos*** (1)
Treatment	0.159***
	(0.036)
Popular	32.285***
	(0.228)
*Treatment*Popular*	1.012***
	(0.322)
Observations	960,401
F test	13845.611***

Notes: (a) The partner platform scaled the raw values of the outcome variable ($NumUploadVideos$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Moreover, to investigate the heterogeneous treatment effect on small creator growth by creator popularity, we report the results in Table 15. We find that the coefficient of the interaction term between $Treatment$ and $Popular$ is significant and positive ($\hat{\beta}_3 = 454.865$, $p < 0.01$), as shown in Columns (1). This finding indicates that the effect of additional viewing traffic on creator growth in the form of acquiring new followers is more salient for popular small creators. This result suggests that popular small creators are able to capitalize on the increased exposure more effectively than less popular creators. When popular small creators receive additional traffic, they are more likely to convert those viewers into new followers.

To sum, the results of the HTE analysis provide useful insights into the impact of traffic incentives on small creators on the platform. The main findings indicate that allocating more traffic towards small creators leads to an increase in their video production volume without compromising the quality of the content produced. Additionally, small creators acquire more new followers and enhance their content creation efforts as a result of this increased viewer traffic. Notably, the benefits are more salient for relatively popular small creators. Therefore, for the platform, strategically allocating traffic to small creators, especially those who are relatively more popular, constitutes an effective strategy for the platform to foster creator growth and development.

Table 14: Heterogeneity by Creator Popularity on Creator Effort in the Creator-side Experiment

Variable	***ProducedVideoLength*** (1)	***DaysOfProduction*** (2)	***SelfReplyRate*** (3)
Treatment	-3.652**	0.314***	0.006***
	(1.669)	(0.050)	(0.001)
Popular	-17.704***	14.899***	0.023***
	(5.482)	(0.164)	(0.002)
*Treatment*Popular*	9.624	-0.162	0.006**
	(7.707)	(0.231)	(0.003)
Observations	182,753	182,753	182,753
F test	5.509***	5514.549***	182.113***

Variable	***SelfForwardRate*** (4)	***SelfCommentRate*** (5)
Treatment	-0.002	0.002***
	(0.002)	(0.000)
Popular	0.054***	0.01***
	(0.007)	(0.002)
*Treatment*Popular*	0.001	0.005**
	(0.010)	(0.002)
Observations	182,753	182,753
F test	42.114***	50.425***

Notes: (a) The partner platform scaled the raw values of the outcome variable ($ProducedVideoLength$, $DaysOfProduction$, $SelfReplyRate$, $SelfForwardRate$, $SelfCommentRate$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

6 Discussions and Conclusion

6.1 Key Findings

Delving into the long tail of creators on content platforms, this work empirically investigates the value of small creators' content to viewer engagement and the efficacy of viewer traffic allocation on small creator development. In collaboration with a leading online short-video sharing platform, we report findings from two field experiments. The viewer-side experiment shows that the content generated by small creators, on average, negatively impact viewer engagement, likely due to the perceived low content quality. However, small creators also contribute to the platform by producing more diverse content, which is especially appreciated by active viewers. The presence of small creators appear to enhance active viewers' viewing experience and keeps them engaged by providing a wider range of perspectives and styles. Next, the Creator-

Table 15: **Heterogeneity by Creator Popularity on Creator Growth in the Creator-side Experiment**

Variable	*NumNewFollowers*
	(1)
Treatment	5.632*
	(3.226)
Popular	116.267***
	(20.455)
*Treatment*Popular*	454.865***
	(28.885)
Observations	960,401
F test	276.357***

Notes: (a) The partner platform scaled the raw values of the outcome variable ($NumNewFollowers$) to protect data confidentiality. As a result, the coefficients presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

side experiment revealed that allocating additional viewer traffic to small creators leads to increases in their video production quantity without compromising quality. Additionally, small creators enhance their content creation efforts and acquire more new followers. The benefits of traffic allocation on increases in content production quantity, creator effort, and new follower numbers are more salient for relatively popular small creators.

6.2 Implications to Research

Our study contributes to the literature on digital content creation. Prior research has explored various aspects of content creation, including motivating factors (Tang et al., 2012; Zeng et al., 2023), creator heterogeneity (Carroni et al., 2024; Foerderer et al., 2023), and platform mechanisms (Jain and Qian, 2021; Qian and Jain, 2024). However, much of the prior work has focused on "star creators" with substantial followings (Qian and Xie, 2022), with limited understanding on the value and behavior of less-known, small creators. The long-tail of small creators hold significant potential for content platforms as they can lead to increased market concentration (Bhargava, 2022) and help mitigate the imbalanced viewer attention that exacerbates disparities between star and non-star creators (Qian and Xie, 2022). Our work aims to fill this research gap in the literature by focusing on small creators, defined as those with fewer than one thousand followers, who make up about 99% of the creator pool. We investigate the value of small creators' content for viewer engagement and evaluate the efficacy of viewer traffic allocation on their development. Our findings highlight the importance of small creators in catering to niche interests and enhancing content diversity. Additionally, we reveal meaningful differences among small creators in their ability to leverage increased exposure and improve performance.

This work also contributes to research on algorithm management, especially for recommender systems on online video-sharing platforms. Prior research has demonstrated the effectiveness of personalized recommendation algorithms in enhancing user engagement and content consumption (Claussen et al., 2019; Holtz et al., 2020). However, these algorithms have also been associated with unintended consequences, such as creating "filter bubbles" and "popularity bias" (Anderson et al., 2020; Abdollahpouri et al., 2019). To address these challenges, scholars have focused on developing and refining algorithms to promote more diverse and balanced content exposure. For example, Zhou et al. (2010b) proposed a topic diversification method to increase the diversity of recommendation lists while maintaining accuracy, and Chen et al. (2024) introduced a deep learning approach for recommending background music for short videos, outperforming existing models. Additionally, researchers have also explored theoretical and analytical approaches to optimize recommender systems. For instance, Li et al. (2018) suggest that the optimal recommendation strategy balances the platform's financial interests and consumer welfare. Similarly, Qian and Jain (2024) proposed that biased recommendations favoring high-quality content incentivize creators to invest more in quality, benefiting both the platform and creators. Building upon the extant literature, our study provides novel empirical evidence on the efficacy of viewer traffic allocation in influencing viewer engagement and creator development. We show that targeted adjustments to recommendation algorithm inputs can achieve desirable outcomes without modifying the underlying technical structure.

6.3 Implications to Practice

Our findings offer crucial implications for industry practitioners, particularly regarding traffic allocation in content-sharing platforms, which is vital for creator and user retention and long-term platform profitability. This work enhances understanding of small creators' value and behaviors, leading to an actionable traffic allocation strategy. Our viewer-side experiment shows that active viewers prefer content from small creators, while less active viewers favor established creators. The creator-side experiment reveals that directing additional traffic to small creators, especially the more popular ones, encourages them to produce more content, exert greater effort, and grow their follower base. Combining the insights from both experiments, we propose a viewership traffic allocation strategy that aims to balance viewer engagement and small creator development: Boosting the visibility of content from small creators, especially those who are relatively more popular, for the active viewers; and limiting the exposure of small creators' content to less active viewers in the recommendation algorithm of the platform. This strategy elevates the engagement of both active and less active viewers, and improves the productivity and growth of small creators. Therefore, the proposed approach has the potential to solve the "popularity bias" problem, and embody a desirable tripartite win-win-win outcome, indicative of a Pareto improvement that simultaneously benefits the viewers, creators, and platform.

6.4 Limitations and Future Research

This work has several limitations that present opportunities for future research. First, the short-term observation windows of our field experiments prevent examining long-term effects on viewer engagement and creator development (Ataman et al., 2010). Future studies should conduct longitudinal research to explore these impacts over time. Second, we focused on content diversity as the main mechanism driving viewer engagement, finding that small creators' diverse content appeals more to active viewers. Future research should investigate alternative mechanisms to better understand content consumption patterns. Third, we did not adjust for potential violations of the Stable Unit Treatment Value Assumption (SUTVA) (Pashley and Bind, 2023), considering the experimental traffic was less than 1% of the total platform traffic, thus minimizing significant breaches. However, future studies should conduct two-sided online experiments to comprehensively understand interactions between viewers and creators and potential spillover effects. Finally, while our results suggest an optimal strategy for leveraging viewer traffic allocation, future research should empirically evaluate this strategy in real-world settings to provide concrete evidence and guide platform managers in making informed traffic allocation decisions.

References

Abdollahpouri, H., Mansoury, M., Burke, R., and Mobasher, B. (2019). The unfairness of popularity bias in recommendation. *arXiv preprint arXiv:1907.13286*.

Anderson, A., Maystre, L., Anderson, I., Mehrotra, R., and Lalmas, M. (2020). Algorithmic effects on the diversity of consumption on spotify. In *Proceedings of The Web Conference 2020*, pages 2155–2165. ACM.

Anderson, C. (2006). *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion.

Ataman, M. B., Van Heerde, H. J., and Mela, C. F. (2010). The long-term effect of marketing strategy on brand sales. *Journal of Marketing Research*, 47(5):866–882.

Bai, B., Dai, H., Zhang, D. J., Zhang, F., and Hu, H. (2022). The impacts of algorithmic work assignment on fairness perceptions and productivity: Evidence from field experiments. *Manufacturing & Service Operations Management*, 24(6):3060–3078.

Besbes, O., Gur, Y., and Zeevi, A. (2016). Optimization in online content recommendation services: Beyond click-through rates. *Manufacturing & Service Operations Management*, 18(1):15–33.

Bhargava, H. K. (2022). The creator economy: Managing ecosystem supply, revenue sharing, and platform design. *Management Science*, 68(7):5233–5251.

Brown, M. A., Bisbee, J., Lai, A., Bonneau, R., Nagler, J., and Tucker, J. A. (2022). Echo chambers, rabbit holes, and algorithmic bias: How youtube recommends content to real users. *SSRN Electronic Journal*. Available at SSRN: https://ssrn.com/abstract=4088828.

Carroni, E., Madio, L., and Shekhar, S. (2024). Superstar exclusivity in two-sided markets. *Management Science*, 70(2):991–1011.

Chen, G., Chan, T. Y., Zhang, D., Liu, S., and Wu, Y. (2023). The effects of diversity in algorithmic recommendations on digital content consumption: A field experiment. *SSRN Electronic Journal*.

Chen, J., He, L., Liu, H., Yang, Y., and Bi, X. (2024). Background music recommendation on short video sharing platforms. *Information Systems Research*.

Claussen, J., Peukert, C., and Sen, A. (2019). The editor vs. the algorithm: Returns to data and externalities in online news. *CESifo Working Paper Series*, 8012.

Cui, R., Gallino, S., Moreno, A., and Zhang, D. J. (2018). The operational value of social media information. *Production and operations management*, 27(10):1749–1769.

DataProt (2023). Average time spent on tiktok. https://dataprot.net/statistics/average-time-spent-on-tiktok/. Accessed: 2023-04-01.

Dewi, Y. R. (2021). Creating customer engangement and customer value within 15 second: how tiktok works for content marketing. *AMAR (Andalas Management Review)*, 5(1):33–45.

Elberse, A. (2008). Should you invest in the long tail? *Harvard business review*, 86(7/8):88–96.

Feng, J., Li, X., and Zhang, X. (2019). Online product reviews-triggered dynamic pricing: Theory and evidence. *Information Systems Research*, 30(4):1107–1123.

Foerderer, J., Gutt, D., and Greenwood, B. N. (2023). Star wars: An empirical investigation of star performer turnover and content supply on multi-sided streaming platforms. *George Mason University School of Business Research Paper*.

Greenstein, S., Gu, G., and Zhu, F. (2021). Ideology and composition among an online crowd: Evidence from wikipedians. *Management Science*, 67(5):3067–3086.

Heinonen, K. (2011). Consumer activity in social media: Managerial approaches to consumers' social media behavior. *Journal of consumer behaviour*, 10(6):356–364.

Holtz, D., Carterette, B., Chandar, P., Nazari, Z., Cramer, H., and Aral, S. (2020). The engagement-diversity connection: Evidence from a field experiment on spotify. *Proceedings of the 21st ACM Conference on Computer-Supported Cooperative Work and Social Computing*.

Hosanagar, K., Fleder, D., Lee, D., and Buja, A. (2014). Will the global village fracture into tribes? recommender systems and their effects on consumer fragmentation. *Management Science*, 60(4):805–823.

Jain, S. and Qian, K. (2021). Compensating online content producers: A theoretical analysis. *Management Science*, 67(11):7075–7090.

Jennings, E. M. P. (2022). Short video marketing: A good strategy for small businesses on tiktok? Research paper, George Mason University School of Business.

Lambrecht, A. and Tucker, C. (2019). Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science*, 65(7):2966–2981.

Li, L., Chen, J., and Raghunathan, S. (2018). Recommender system rethink: Implications for an electronic marketplace with competing manufacturers. *Information Systems Research*, 29(4):1003–1023.

Liu, X., Mei, S., and Zhong, W. (2022). Video-sharing platform's optimal monetary incentive decisions considering motivation crowding-out effect. *Managerial and Decision Economics*, 44(1):371–387.

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

Liu, Y., Feng, J., and Liao, X. (2017). When online reviews meet sales volume information: Is more or accurate information always better? *Information Systems Research*, 28(4):723–743.

Moeller, J., Helberger, N., et al. (2018). Beyond the filter bubble: Concepts, myths, evidence and issues for future debates. *Amsterdam. University of Amsterdam. Retrieved January*, 19:2022.

Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., and Konstan, J. A. (2014). Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*, pages 677–686.

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, 104(11):3565–3599.

Pashley, N. E. and Bind, M.-A. C. (2023). Causal inference for multiple treatments using fractional factorial designs. *Canadian Journal of Statistics*, 51(2):444–468.

Qian, K. and Jain, S. (2024). Digital content creation: An analysis of the impact of recommendation systems. *Management Science*.

Qian, K. and Xie, Y. (2022). The power of star creators: Evidence from the live streaming industry. *Available at SSRN 4123516*.

Rafieian, O. (2023). Optimizing user engagement through adaptive ad sequencing. *Marketing Science*, 42(5):910–933.

Shutsko, A. (2020). User-generated short video content in social media. a case study of tiktok. In *Social Computing and Social Media. Participation, User Experience, Consumer Experience, and Applications of Social Computing: 12th International Conference, SCSM 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22*, pages 108–125. Springer.

Srinivasan, R. and Sarial-Abi, G. (2021). When algorithms fail: Consumers' responses to brand harm crises caused by algorithm errors. *Journal of Marketing*, 85(5):74–91.

Sun, J., Zhang, D. J., Hu, H., and Van Mieghem, J. A. (2022). Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations. *Management Science*, 68(2):846–865.

Sun, X., Han, M., and Feng, J. (2019). Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. *Decision Support Systems*, 124:113099.

Tang, Q., Gu, B., and Whinston, A. B. (2012). Content contribution for revenue sharing and reputation in social media: A dynamic structural model. *Journal of Management Information Systems*, 29(2):41–76.

Zeng, Z., Dai, H., Zhang, D. J., Zhang, H., Zhang, R., Xu, Z., and Shen, Z.-J. M. (2023). The impact of social nudges on user-generated content for social network platforms. *Management Science*, 69(9):5189–5208.

Zhang, X. and Zhu, F. (2011). Group size and incentives to contribute: A natural experiment at chinese wikipedia. *American Economic Review*, 101(4):1601–1615.

Zhao, K., Lu, Y., Hu, Y., and Hong, Y. (2023). Direct and indirect spillovers from content providers' switching: Evidence from online livestreaming. *Information Systems Research*, 34(3):847–866.

Zhao, K., of Houston, U., Hu, Y., Hong, Y., Westland, J. C., and of Illinois, U. (2021). Understanding characteristics of popular streamers on live streaming platforms: Evidence from twitch.tv. *Journal of the Association for Information Systems*, 22(4):1076–1098.

Zhou, R., Khemmarat, S., and Gao, L. (2010a). The impact of youtube recommendation system on video views. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 404–410.

Zhou, T., Kuscsik, Z., Liu, J.-G., Medo, M., Wakeling, J. R., and Zhang, Y.-C. (2010b). Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences*, 107(10):4511–4515.

Online Appendices

A Recommender System

The partner platform utilizes a state-of-the-art, deep-learning based recommendation algorithm in its recommender system, aligning with industry standards, to deliver video recommendations to viewers. As illustrated in Figure A.1, this system operates through a sequential three-phase process: retrieval, ranking, and re-ranking, to effectively navigate through a vast video pool and recommend a curated selection of videos (Chen et al., 2023). In the retrieval phase, the recommender system initially selects a broad set of videos (ranging from 500 to 10,000) from the video pool for each user. The selection process involves computing the proximity between user and video embeddings, prioritizing videos with closer distances indicative of user interest and preference. In the subsequent ranking phase, the algorithm ranks the retrieved set of videos based on their anticipated match value to the user, shortlisting the top videos (approximately 100 to 300) for further processing. This ranking is accomplished by deploying multiple deep neural networks designed to predict user behaviors. At last, in the re-ranking phase, the system re-ranks the previously ranked videos to determine the final recommendations (between 10 and 20 videos). The selected videos are then sorted based on user preferences and sequentially presented to the users on the platform's user interface. After the user consumes these 10 to 20 recommended videos, if they continue to engage with the platform, the recommender system dynamically excludes the previously recommended videos and repeat the aforementioned phases to refresh the recommendations. This iterative process ensures that the user is continually presented with fresh, relevant, and engaging content tailored to their evolving interests and preferences.

B Randomization Checks

In the viewer-side experiment, we conducted randomization checks on all observable covariates across treatment and control groups. Specifically, we compared viewer demographics from one day before the experiment (December 1, 2022) and viewers' behavioral characteristics from the prior week (November 25 - December 1, 2022). As shown in Table B.1, two-tailed t-test results indicate no significant differences in observable covariates between the groups, demonstrating the successful randomization and comparability of experimental groups. Hence, any outcome differences after the experiment started shall be attributed to the treatment assignment.

In the creator-side Experiment, we conducted randomization checks on all observable covariates between

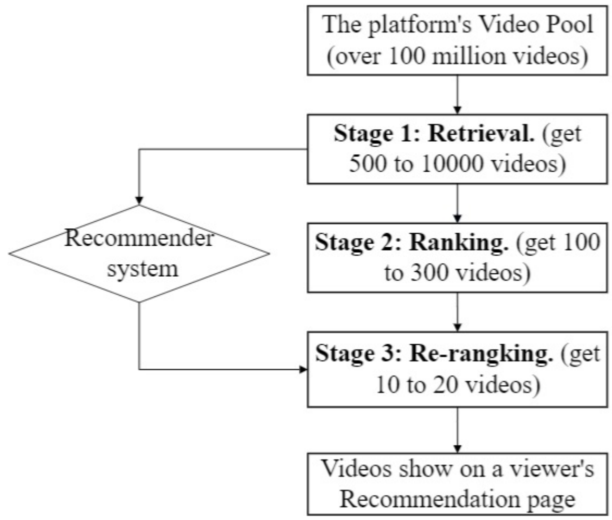

Figure A.1: Workflow of the Recommender System

Table B.1: Randomization Checks of the Viewer-side Experiment

Variable	Treatment	Control	*P*-Value
Viewer Demographics			
Gender	1.48	1.48	0.11
Age	42.60	42.60	0.10
CityLevel	3.40	3.40	0.68
NumFriend	456.42	456.99	0.16
NumFollower	0.40	0.40	0.32
Active	0.04	0.04	0.54
Behavioral Characteristics			
WatchingTime	5,417.88	5,404.59	0.10
NumView	229.28	228.81	0.19
NumComment	0.15	0.15	0.91
NumForward	0.65	0.65	0.41
NumLike	1.49	1.49	0.51
NumCategories	11.26	11.25	0.11

Notes: (a) The partner platform scaled the raw values of the variables ($NumFriend$, $NumFollower$, $WatchingTime$, $NumView$, $NumComment$, $NumForward$, $NumLike$, $NumCategories$) to protect data confidentiality. As a result, the numerical values presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) All variables are measured at the viewer level. (c) We report the *p*-value of two-tailed *t*-tests in the randomization checks. (d) The randomization checks on viewers' behavioral characteristics utilized data from the week before the experiment. In contrast, Table 3 reports data observed during the experimental period. This results in a difference in variable scales across the two.

the treatment and control groups. Specifically, we compared creator demographics from one day before the experiment (March 17, 2022) and behavioral characteristics from the prior two weeks (March 3 - March 17, 2022). As shown in Table B.2, two-tailed t-test results indicate no significant differences in the observable covariates between the groups, demonstrating successful randomization and comparability of experimental groups. Any outcome differences after the experiment started can be attributed to the treatment assignment.

Table B.2: Randomization checks of the Creator-side Experiment

Variable	Treatment	Control	*P*-Value
Creator Demographic			
Gender	1.63	1.63	0.63
Age	37.00	37.00	0.33
CityLevel	3.44	3.44	0.63
NumFollower	10.35	10.26	0.32
Popular	0.01	0.01	0.71
Content Production Quantity and Quality			
NumUploadVideos	3.02	2.98	0.19
LikeRate	0.05	0.05	0.73
CommentRate	0.00	0.00	0.10
ForwardRate	0.02	0.02	0.17
Creator Effort			
ProducedVideoLength	178.84	178.87	0.98
DaysOfProduction	6.96	6.90	0.29
SelfReplyRate	0.01	0.01	0.13
SelfForwardRate	0.20	0.20	0.31
SelfCommentRate	0.01	0.01	0.46
Creator Growth			
NumNewFollowers	2.27	2.24	0.77

Notes: (a) The partner platform scaled the raw values of the variables ($NumUploadVideos$, $LikeRate$, $CommentRate$, $ForwardRate$, $ProducedVideoLength$, $DaysOfProduction$, $SelfReplyRate$, $SelfForwardRate$, $SelfCommentRate$, $NumNewFollowers$) to protect data confidentiality. As a result, the numerical values presented in the table do not hold any intrinsic meaning and should not be interpreted as absolute values. (b) All variables are measured in creator level. (c) We report the p-value of two-tailed t-tests in the randomization checks. (d) The randomization checks on creators' behavioral characteristics used data from two weeks before the experiment. Meanwhile, Table 8 reports data observed during the experimental period, which also spans two weeks.