Algorithmic Attention and Content Creation on Social Media Platforms*

Yi Chen† Fei Li‡ Marcel Preuss§

June 11, 2025

Abstract

This paper develops a theoretical framework to examine how a social media platform allocates attention through recommendation algorithms and how this in turn shapes content creation and consumption. Creators and viewers, as the two sides of the algorithm, fall into different categories based on interest. Creators are also heterogeneous in ability. We show that a platform, to maximize advertisement revenue, optimally filters out low-ability creators, restricts the reach of medium-ability creators to relevant audiences only, and propagates viral content for high-ability ones at the expense of relevance. The attention a creator receives grows disproportionally in his ability and the popularity of his category. We show the source of the inefficiencies of the algorithm by contrasting it with a welfare-maximizing benchmark. We additionally study the effect of monetary transfers in the algorithm. Our framework offers insights into content production and matching in digital markets, giving rise to potential regulatory interventions.

Keywords: Social Media; Platform; Content; Attention; Mechanism; Mismatch
JEL Codes: D81, D82, D83

*We are grateful to Arjada Bardhi, Özlem Bedre-Defolie, Gabriel Carroll, Yeon-Koo Che, Alexandre de Corniere, Tom Conningham, Mira Frick, Chiu-Yu Ko, Wei He, Justin Johnson, Tony Ke, Ron Siegel, Yangbo Song, Kai Hao Yang, Hanzhe Zhang, Jidong Zhou and seminar or conference participants for their helpful comments and valuable discussions. We thank Jiaming Wei for providing an excellent review of the marketing literature for us. All remaining errors are our own.

†SC Johnson Graduate School of Management, Cornell University. yc2535@cornell.edu.
‡University of North Carolina. lifei@email.unc.edu.
§SC Johnson Graduate School of Management, Cornell University. preuss@cornell.edu.

Contents

1 Introduction

Today, there are estimated to be more than 5 billion people globally who use social media, accounting for more than 62 percent of the world population. In the US, the share is even higher, at around 90 percent. The intensity at which people use social media is just as staggering as the number of users, if not more so. Americans, on average, clock in more than 2 hours per day. Teenagers hit almost 5 hours, with 90 percent of their time spent on Youtube, TikTok and Instagram. Accordingly, advertisers spend more than 100 billion USD in the United States alone on social media and influencer advertising.

Social media started out as a network in which what users saw was dependent on who they are connected with. Over the past years, however, many social media platforms have shifted to recommended content, letting algorithms decide what users see and how much attention content creators receive.[1] At the same time, social media platforms make money from advertising. A user scrolling through their Instagram feed inevitably encounters sponsored content regularly, that is, posts (or reels) which are only shown because an advertiser paid Instagram to target the user. To maximize advertising revenue, however, the platform cannot show viewers just ads. Rather, the algorithm needs to blend ads with high quality content of creators that viewers like, or else viewers would leave the platform. To incentivize content creators, in turn, the algorithm must allocate them a certain level of viewers' attention, or else creators would not spend effort to make content. Thus, a social media platform faces a two-sided mechanism design problem.

In this article, we study how a social media platform solves this mechanism design problem and characterize the platform-optimal recommendation algorithm (mechanism). In addition, we consider the welfare-optimal algorithm, allowing us to identify the distortions and inefficiencies created by the platform's profit-maximization incentive. The latter improves our understanding of which regulation and behavioral measures can raise the welfare of viewers and content creators. Finally, we study how introducing monetary transfers between the platform and creators affects the optimal recommendation algorithm. Youtube has been relying on creator payments for a long time, and Tiktok has recently experimented with this too (see WSJ, 2024).

Our analysis reveals three main findings. First, the profit-maximizing algorithm crowds out low-ability content creators, forces intermediate-ability creators to exert too much ef-

[1]On Instagram, the Reels and the Explore tab are exclusively for recommended content. Even the user's feed, which before 2022 was exclusively reserved for followed content, "will have a mix of content from the accounts you've chosen to follow, recommended content from accounts we think you'll enjoy and ads" an offcial instagram blogpost explains (Instagram Announcement, 2022). On Tiktok, recommended content is even more prominent.

fort, and induces viewers to spend excessive time on the platform. Our analysis thus makes an important contribution to the public debate about whether people spend too much time on social media.[2] Second, our analysis provides a profit-maximization rationale of the recent phenomenon of social media content "going viral," which took off after said platforms switched to an algorithm-led approach. Specifically, we show that to maximize profits from advertising, the platform optimally chooses to show already popular content to disproportionally many viewers, including viewers not interested in the content. Third, transfers from the platform to content creators fully eliminate irrelevant content from the viewers' feed if and only if selling ads is sufficiently lucrative for the platform (compared to the value content creators derive from receiving attention).

Our analysis builds on a novel model of social media that allows content creators to choose the effort they put into producing content and viewers to decide whether to pay attention to the recommended content. In the model, an algorithm is a mechanism that determines the content that each viewer sees (the viewer's "feed"), and the attention that each content creator receives. Content creators are horizontally differentiated in that creators focus on different topics, which vary in popularity among viewers. In addition, content creators differ in their ability to create high quality content. Content creators care about receiving attention from viewers.[4] Viewers vary in what topic they are interested in. We accordingly quantify the popularity of a topic by the mass of viewers interested in it. The platform wants to maximize the total attention paid to ads, which it sells for a fixed price in the advertising market. Ads can be blended in the viewer's feed together with the creators' content. Both producing content and paying attention to it is costly. Therefore, the platform needs to satisfy the following two sets of obedience constraints: (i) each viewer's feed yields the viewer weakly positive utility so that viewers pay attention to it, and (ii) each creator receives sufficient attention to motivate his production of high quality content.[5]

[2]A large and growing literature in social sciences studies the effect of social media on well-being. Allcott et al. (2020) find that disconnecting from social media in an experiment improves subjective well-being. Braghieri et al. (2022) find similar results, and further suggest that one of the mechanisms is related to social comparisons (the fear of receiving fewer likes than others).[3] In another controlled experiment by Collis and Eggers (2022), the authors do not find any effect of reducing social media usage on well-being. However, in that study subjects in the treatment group used other apps on their smartphone more heavily (instant messaging), which could be detrimental for well-being as well.

[4]This assumption reflects intrinsic or extrinsic motivation from monetizing attention via deals with advertisers. However, Toubia and Stephen (2013) show that intrinsic motivation plays an important role for social media users. Similarly, Lindström et al. (2021) show that the desire for attention on social media follows a pattern of "reward learning, comparable to the behavior of animals in seeking rewards.

[5]In the model, the platform perfectly observes quality as well as each viewer's and creator's type. In practice, social media platforms learn quality through experimentation, and user characteristics from machine learning. For example, to determine the quality of content, social media platforms reportedly show it only to a small group of users, whose engagement with said content is used as a proxy for quality.

Intuitively, if the platform has access to higher quality content by creators, this relaxes the viewer's obedience constraint. As a result, the platform can blend more ads into the viewer's feed. To incentivize the costly production of high quality content, however, the platform needs to reward producers with more attention. The platform thus faces a trade-off: allocate attention directly to ads or to content producers to incentivize the production of higher quality content, which then allows it to show more ads. For creators who are sufficiently popular or high in ability, the platform finds it profitable to do the latter. This relaxes the obedience constraint of a larger mass of viewers, eventually allowing the platform to show even more ads.

This implies distortions on both user sides of the platform. Some content creators exert more effort than they would if their content were recommended only to viewers who are in fact interested in their content. Conversely, viewers see content they are not interested in – in addition to ads – thus lowering their overall utility from consuming their feed. Thus, the fact that social media platforms earn money from selling ads not only affects viewers because it means viewers are exposed to ads, but it also because it distorts the platform's recommendation algorithm toward recommending irrelevant content.

That the platform thrives on irrelevant content recommendations also explains why certain content is made viral, i.e., shown to all viewers regardless of the viewers' interests. While showing the content of popular creators to all viewers lowers some viewers' utility from reading their feed, this boosts the utility of popular creators and allows the algorithm to extract more effort from them.[6] This, in turn, relaxes the obedience constraint of the viewers interested in those creators' topics, of whom there are many since the platform makes mostly creators with popular topics go viral. Those viewers, therefore, can then be shown more ads, whereas the platform shows more irrelevant content to a smaller mass of viewers interested in less popular topics. In other words, the irrelevant content has a non-linear effect on profit after this feedback loop. As a result of this feedback loop, the number of ads shown to a viewer depends on their type. Viewers interested in less popular topics see fewer ads because the platform finds it more profitable to steer their attention toward content creators of popular categories to raise the effort of these creators. Due to the increased quality of content by popular creators, the platform can then expose viewers interested in that popular content to more ads.

Finally, we highlight that if the value of ads is high, then inefficient recommendations of irrelevant content arise only if the platform is unable to pay creators. Compared to the algorithm that directs some of a viewer's attention to irrelevant (from a certain viewer's

Our analysis abstracts from this experimentation phase.

[6]Note that popular here means that the creator focuses on a popular topic. That those creators become eventually popular in the true sense is endogenous to the algorithm and not assumed ex ante.

point of view) but popular content creators, showing the viewer an ad instead means less total attention that the platform can allocate to those popular creators. With transfers, however, the platform can provide alternative incentives. In particular, if the platform earns more for directing a unit of attention to an ad than the content creators value one unit of attention, then it is more profitable to incentivize the content creator with payments rather than attention from viewers who deem their content irrelevant.

In sum, our model rationalizes common behaviors related to social media. First, with the advent of TikTok in the US, which pioneered the heavy use of recommended content on users' feeds, a new phenomenon started on college campus (and other places): groups of young adolescents spending hours together to film content for the platform, hoping the algorithm will help their video go viral.[7] Our model predicts that the profit-maximizing algorithm crowds out participation from low- to medium-ability content creators in favor of higher average content quality and increased advertisement exposure. Second, allocating too much attention to already popular content is part of the optimal algorithm. This explains the growing number of content creators whose content is distributed to millions of viewers as well as the phenomenon of viral content in general. Third, the profit-maximizing algorithm leads to excessive time spent on the social media platform, which appears in line with the high reported average daily social media usage of 4.8 hours (among teenagers) in the US. As our analysis shows, the main reason for these distortions lies in the advertising-funded nature of social media. As it tries to show viewers more ads, social media needs to also ramp up the production of content, which, in turn, requires it to harvest even more attention from viewers.

We contribute to the growing literature that studies competition for attention on social media. Relatively early work by Iyer and Katona (2016) studies a model of social media in which the platform cannot directly control the flows of content and attention. Rather, the key feature of social media in their model is that a message can be sent to multiple receivers. The authors show that increasing the number of recipients (growing the social media platform) drives up effort of senders, but also leads to fewer people choosing becoming senders. Ghosh and McAfee (2011), by contrast, consider a platform that can design an algorithm to incentivize content production. Specifically, they allow for algorithms that exclude low quality content from producers who exert too little effort.[8] Ben-Porat and Tennenholtz (2018) study recommender systems with strategic content creators

[7]Besides, this anecdotal evidence, data show that the share of teenagers (its heaviest users) creating content on tiktok indeed very high at almost 80 percent .

[8]The platform relies on collecting engagement data from users to learn about the quality of content. They derive an algorithm which maximizes the average quality on the platform while keeping the number of instances when users see a low quality post for the purpose of learning the post's quality as small as possible.

as well but consider different objectives of the algorithm (e.g., fairness). Importantly, in those papers, the recommendation algorithm does not affect the total supply of attention, nor is the goal to maximize profits from advertising.

Closest to our work is the analysis by Qian and Jain (2024). They study the interaction of social media recommendation algorithms with endogenous content creation and revenue-sharing plans between influencers and the platform. Their main result is that the platform may want to bias its recommendation in favor of high-quality content even if it is less relevant. There are several differences between their analysis and ours. First, they fix the number of content creators and put a cap on content consumption, whereas we allow the algorithm to endogenously choose the mass of creators, which has welfare implications. Second, our model features rich heterogeneity among viewers and content creators, enabling us to characterize users' equilibrium participation, production and consumption for different types of creators and viewers. This allows us to assess the distributional impact of targeted regulatory interventions.

More generally, we contribute to the literature on advertising-funded media platforms (see, e.g., Anderson and Coate, 2005; Peitz and Valletti, 2008). More broadly, our research is related to two-sided markets research (see Jullien et al., 2021, for an excellent overview). The two-sided market literature typically considers platforms that charge at least one group of users for access to the other side of the market (e.g., advertisers). Since social media platform also charges advertisers, this work is closely related to ours. However, our focus lies on moderating the exchange between content creators and viewers, neither of which pay a monetary fee. Notable exceptions are from Bhargava (2022) and Ren (2024), who explicitly model the three-sided nature of social media platforms. Bhargava (2022) analyzes the optimal level of ads permitted by the platform and the optimal revenue sharing mechanism given endogenous content supply decisions. Relatedly, Ren (2024) studies advertising policies on decentralized content creation and examine its implications on designing advertising and revenue-sharing. However, these authors do not study the design of the optimal algorithm.[9]

Other work on social media includes Filippas et al. (2023), who consider a model of social media platforms (e.g., Facebook or the previously Twitter) where users, rather than an algorithm, have full control over which users they interact with. They show that users form strategic links to attain more attention, and characterize patterns of such link formation between users. Those papers differ from ours in that we focus on the case where the platform can control the flow of attention between users through the design of a rec-

[9]Rather than modeling an algorithm explicitly, they assume, for example, that utility of viewers increases in the average effort and mass of content creators joining the platform, or that higher quality content is always prioritized by default.

ommendation algorithm. Although not our focus, others have also studied the spread of misinformation through social media users and the ensuing polarization (e.g., Berman and Katona, 2020; Acemoglu et al., 2024).

In our model, some content creators endogenously emerge as influencers due to the algorithm and we are agnostic about whether those influencers desire attention for monetary incentives or are intrinsically motivated. A small but growing literature focuses more on how influencers make money, and the trade-off the face when showing organic versus sponsored content, including Fainmesser and Galeotti (2021) and Mitchell (2021).

Notably, sellers, two-sided platforms, and intermediaries in general, frequently use recommendation algorithms in e-commerce setups as well. Bergemann and Bonatti (2024) study a platform that uses data to match heterogeneous consumers with multi-product sellers and consumers can purchase the product both on and off the platform. A common finding in this literature is that the profit-maximizing algorithm not always recommends the best product to consumers. For example, Hagiu and Jullien (2011) argue that an information intermediary diverts consumer search to gain higher consumer traffic and influence sellers pricing. Teh and Wright (2022) show that an intermediary has the incentive to steer the recommendation to influence the competition of upstream sellers on prices and commissions. Choi and Jeon (2023) analyze the platforms' design biases in a two-sided market. Peitz and Sobolev (2025) show when an intermediary recommends a welfare-reducing bad match to facilitate better surplus extraction from sellers, and Bar-Isaac and Shelegia (2022) consider when an intermediary steers consumers to more profitable products. Janssen et al. (2023) study the profit-maximizing ranking algorithm of a search platform when consumers face search costs to inspect all options and find that the platform obfuscates the search results. De Corniere and Taylor (2019), Aridor and Gonçalves (2022) and Chen and Tsai (2024) study how an intermediary leverages biased recommendations to favor its own products when competing with third-party sellers. Ichihashi and Smolin (2023) and Condorelli and Szentes (2023) examine how recommendation algorithms can enhance consumers' countervailing power, shielding them from surplus extraction by sellers.

The rest of the paper is organized as follows. Section 2 sets up the model, and Section 3 studies the equilibrium under some intuitive algorithms. In Section 4, we characterize the optimal algorithm and contrast it with a hypothetical welfare-maximizing algorithm. Section 5 extends the base model by allowing for monetary transfer. Section 6 concludes.

2 Model

2.1 Primitives

We model social media platform as a monopolistic two-sided online marketplace where users produce and consume digital content, such as articles, music, and videos. The platform employs a personalized recommendation algorithm to distribute content and advertisements to maximize its advertising revenue.

Creators and Viewers We consider two kinds of platform users: measure one of content creators and measure one of content viewers, which we call *creators* and *viewers* for short.

A creator is characterized by a two-dimensional type (θ, j), where $\theta \in \mathbb{R}_+$ represents his ability, and $j \in \mathcal{N} \equiv \{1, \ldots, N\}$ indicates the horizontal category of a creator's contents. The marginal distribution of type-j creators is denoted $\mu_j > 0$. The conditional distribution of θ given j is denoted by a continuous C.D.F. $F(\theta|j)$.

For simplicity, we only account for horizontal heterogeneity among viewers and assume that every viewer is interested in only one content category. So, a viewer is labeled by $k \in \mathcal{N}$, the horizontal category that she is interested in.[10] The proportion of type-k viewers is denoted $\nu_k > 0$.

Contents and Ads Each creator of type (θ, j) can put in costly *effort* $e_j^\theta \geqslant 0$ to produce one unit of *content*, which can be a post or a video, for example. We assume that effort affects the *quality* of the content generated, not quantity. With effort level e_j^θ, creator type (θ, j) generates content of quality $q_j^\theta \equiv \theta e_j^\theta$. Therefore, the ability and effort are supermodular: the higher ability θ, the more effective of effort.

Meanwhile, there is a competitive external market for *ads*. The market has an unlimited supply of ads that pays the platform a fixed piece rate whenever an ad is actually read by a viewer.

Algorithm We model the algorithm of the social media platform as a recommendation mechanism. Formally:

Definition 1 (Algorithm)
An algorithm $\mathscr{A} \equiv (\tilde{a}, \tilde{A})$ *consists of a* content recommendation $\tilde{a}_{j,k}^\theta(q) : \mathbb{R}_+ \times \mathcal{N}^2 \times \mathbb{R}_+ \to$

[10]Similar results arise if we allow a viewer to be interested in multiple categories.

$[0,1]$ *and an* ads recommendation $\tilde{A}_k : \mathcal{N} \to \mathbb{R}_+$.

The content recommendation $\tilde{a}^{\theta}_{j,k}(q)$ denotes the probability of recommending creator (θ, j)'s content to a type-k viewer, given the observed quality q. Remarkably, unlike ordinary commodities, digital content is *non-rival*: a creator's content can be simultaneously recommended to an unlimited number of viewers without diminishing its availability. The ads recommendation $\tilde{A}_k$ denotes the measure of ads from the external market that the platform recommends to a type-k viewer. With these two recommendation functions, the personalized *recommendation set* for a type-k viewer consists of $\tilde{a}^{\theta}_{j,k}(q)$ proportion of creator type (θ, j)'s contents, plus a measure $\tilde{A}_k$ of ads. Here, we allow $\tilde{a}^{\theta}_{j,k}(q)$ to be any real number in $[0,1]$, as type (θ, j) is interpreted as the continuum of creators within the neighborhood of that type.

Consumption Given the recommendation set, each viewer has the discretion over how much to read from it. We assume the viewer cannot cherry-pick from the recommendation set. That is, the algorithm blends the contents and ads in the recommendation set so that the viewer cannot distinguish between them before actually read them. As a result, even if a viewer has preference over what to read, she must consume all contents and ads in proportion to the recommendation set. The action of a type-k viewer is thus to choose a share $\alpha_k \in [0,1]$ to read from the recommendation set.

Recommended contents and ads are not necessarily consumed, as a viewer k may want to choose some $\alpha_k < 1$. Therefore, we use $a^{\theta}_{j,k}$ and A_k to denote type-k viewer's actual *attention* on contents created by type (θ, j) and on ads, respectively. Because a viewer type k can only randomly pick α_k proportion within the recommendation set to read, we immediately have $a^{\theta}_{j,k} = \alpha_k \tilde{a}^{\theta}_{j,k}(q^{\theta}_j)$ and $A_k = \alpha_k \tilde{A}_k$.

Payoffs Creators derives benefits while incurring costs. For any creator, an effort e costs $c(e)$. As usual, we assume $c(0) = c'(0) = 0$ and $c'(e), c''(e) > 0$ for all $e > 0$. Moreover, $c(\cdot)$ is assumed to be log-concave for normality. Moreover, they enjoy utility from the attention by viewers. In particular, a creator derives utility $u > 0$ from every unit of attention, regardless of its source. This specification is a reduced form to capture creators' desire to gain online reach as it is the foundation of their psychological satisfaction, earning potential, influence, and career growth. In section 5 we allow for additional monetary incentives for the creators. Given realized effort e^{θ}_j, a type-(θ, j) creator's payoff equals:

$$u \sum_k \nu_k a^{\theta}_{j,k} - c(e^{\theta}_j).$$

On the other side, viewers have benefit-cost trade-offs. Attention is costly in that every unit of attention costs a viewer $t > 0$, regardless of where the attention is spent. In return, viewers derive ex post utility from reading contents (from entertainment or information, for example). If a viewer type k reads some content from creator type (θ, j), then the utility from reading is $q_j^\theta \mathbb{1}\{j = k\}$. This utility depends on two factors: the quality q_j^θ and *relevance* (i.e., whether $j = k$ or not). Here we make the simplifying assumption that reading *irrelevant content* ($j \neq k$) yields zero utility, but this can be easily generalized. The benefit from watching ads is always zero. Given realized attention $a_{j,k}^\theta$ and A_k, a type-k viewer's payoff equals:

$$\mu_k \int a_{k,k}^\theta q_k^\theta \mathrm{d}F(\theta|k) - t \cdot \left[\sum_j \mu_j \int a_{j,k}^\theta \mathrm{d}F(\theta|j) + A_k \right].$$

The platform profits from showing ads. The external market for ads is such that each unit of attention spent on an ad yields the platform a profit of $z > 0$. Given realized attention A_k, the platform's total profit reads:

$$z \sum_k \nu_k A_k.$$

Timeline To summarize the model, we lay out the timeline of the mechanism as follows.

1. The platform commits to the algorithm $\mathscr{A}$, and then all creators and viewers simultaneously decide to join the platform or not. The outside option for all users is zero.
2. All creators (θ, j) who join the platform simultaneously put in effort e_j^θ to produce content of quality $q_j^\theta = \theta e_j^\theta$.
3. After seeing all q_j^θ, the algorithm uses $\tilde{a}_{j,k}^\theta(q_j^\theta)$ and $\tilde{A}_k$ to compile personalized recommendation sets for all viewer types.
4. Each viewer k receives the recommendation set and chooses a proportion $\alpha_k \in [0, 1]$ to read within it.
5. Actual attentions $a_{j,k}^\theta = \alpha_k \tilde{a}_{j,k}^\theta$ and $A_k = \alpha_k \tilde{A}_k$ realize, and payoffs realize.

2.2 Model Discussions

The mechanism involves several important assumptions, which we discuss here. First, the platform can see the types of users as well as the quality of contents. This is because our model aims to characterize the long-run equilibrium instead of the transitory learning stage. With big data and long term interactions, a user's type is easily learned by algorithm. Quality is also mostly visible because a platform can hire a small set of test viewers or even AI to judge the quality of posts (Ghosh and McAfee, 2011).

Second, the assumption that content creators derive utility from attention is documented by a large empirical literature, and is also adopted in theory studies (e.g., Filippas et al., 2023). It explains why in many platforms with user-generated contents people voluntarily contribute even without direct monetary reward. For the purpose of this paper, we remain agnostic about the exact source of utility, be it psychological satisfaction or exogenous pecuniary benefits proportional to the creator's popularity.

Third, we assume that reading is an experience good. That is, viewers cannot cherry-pick contents within the recommended set without incurring some attention costs. This assumption is mostly appropriate when it comes to static content such as short text or photos, because by the time viewers determines whether they like it or not, the attention is already spent. That is, there is little difference between evaluating and consuming it. Longer videos or texts, on the other hand, are different in that users can try to filter contents from the first few seconds of reading. However, such screening is far from perfect, as "click bait" on social media platform is all but rare. Oftentimes, viewers watch a video till the end, only to find that it is a scam or an embedded ad. As long as viewers cannot perfectly filter, the no-cherry-picking assumption is innocuous.[11] Moreover, trying to guess the quality by the sequence of recommended contents and ads is not effective, as the algorithm can always prevent it by randomizing the sequence.

Finally, in our model viewers can only read within their tailored recommendation. Notably, this does not forbid viewers from following creators of their choice; rather, the constraints is that viewers cannot read "followed content" exclusively. In reality, recommended content represents the lion's share of what people consume social medias such as Tiktok and Instagram, while "followed content" is in decline. These platforms even start to weaken the "follow" function so that recommended contents and ads sneak in under the tab of "followed content."[12] In sum, the platforms realize the profitability of strengthening the algorithm and taking control of the attention flow. As such, we omit the "follow" function to simplify analysis.

[11]Suppose the platform shows a viewer a mass $A > 0$ of ads if the viewer cannot cherry pick at all. Now suppose the user can detect and skip irrelevant content with probability $1/2$, then the platform can just raise the mass of ads to $2A$ to achieve the same outcome.

[12]Similarly, on Quora and Reddit, ads and recommended threads are mixed among the pertinent contents.

3 Some Special Algorithms

Before solving for the optimal algorithm, we can first study some simple algorithms and understand why or why not such an algorithm works towards profit maximization. For simplicity, we assume symmetry among categories in that $\mu_j = \nu_j = \frac{1}{N}$ and $F(\theta|j) = F(\theta)$ for all j.

3.1 Laissez Faire

The Laissez Faire algorithm is a trivial one, in which the platform does not put any restriction on the attention recommendation. One can imagine the bulletin board system many years ago as a Laissez Faire algorithm because everything that is produced is equally accessible to all viewers. In fact, any platform that does not actively direct attention is arguably using this algorithm.

Formally, the Laissez Faire algorithm is simply $\tilde{a}^{\theta}_{j,k}(q) = 1$ for all θ, j, k, q. The payoff for a creator (θ, j) is then $u\frac{\sum_k \alpha_k}{N} - c(e^{\theta}_j)$. Effort brings only cost but no benefit. Therefore, in equilibrium $e^{\theta}_j = 0$. On the other side, a viewer k's payoff is reduced to $-\alpha_k t(1 + \tilde{A}_k)$, which is maximized at $\alpha_k = 0$. Therefore, actual attentions $a^{\theta}_{j,k}$ and A_k are both zero, and the platform earns zero profit.

3.2 Quality Control

Laissez Faire algorithm fails because no creator puts in effort. A natural remedy is to set a minimum quality requirement for the content to be recommended. Many forums have a moderator, being a human or bot, that filters out low quality contents. Other social platforms use algorithms to dampen the visibility of low quality contents.

Formally, the algorithm requiring minimum quality is the one that for all θ, j, k, set $\tilde{a}^{\theta}_{j,k}(q) = 1$ if $q \geqslant \underline{q}$, and $\tilde{a}^{\theta}_{j,k}(q) = 0$ otherwise, where $\underline{q}$ is the lowest allowable quality to be recommended. The payoff for a creator (θ, j) is piecewise: $u\frac{\sum_k \alpha_k}{N} - c(e^{\theta}_j)$ if $e^{\theta}_j \geqslant \underline{q}/\theta$, and $-c(e^{\theta}_j)$ if $e^{\theta}_j < \underline{q}/\theta$. This leaves only two candidate effort levels: the minimum qualifying effort $\underline{q}/\theta$, and zero effort. A creator prefers the former if and only if $\theta \geqslant \underline{\theta} \equiv \frac{\underline{q}}{c^{-1}(u\frac{\sum_k \alpha_k}{N})}$. On the other side, a viewer k's payoff is reduced to $\alpha_k((1 - F(\underline{\theta}))(\underline{q}/N - t) - t\tilde{A}_k)$. In order for $\alpha_k > 0$ for any k, a necessary condition is $\underline{q} \geqslant Nt$. In fact, as long as $\underline{q} \geqslant Nt$, the platform is able to insert $\tilde{A}_k \geqslant 0$ while keeping $\alpha_k = 1$. A positive profit is thus possible, but this is achieved at the cost of a very high quality requirement, reducing the total amount of contents and consequently the amount of ads. This scenario is similar to those traditional mass media such as news paper and news report, where only highly qualified creators are allowed to publish.

3.3 Quality Control with Perfect Targeting

Now that a minimum quality is required to ensure effort, the last example still faces the problem of too little participation. This is due to the lack of targeting. In expectation, only $1/N$ of the contents are relevant for a viewer, and this significantly dilutes the utility a viewer obtains from reading. A platform may want to utilize the power of algorithm to perfectly target the audience so that upon receiving a recommendation set, a viewer is much more willing to read to begin with, leaving more room for inserting ads.

Formally, the algorithm is the one that for all θ, j, k, set $\tilde{a}^{\theta}_{j,k}(q) = 1$ if $q \geqslant \underline{q}$ and $j = k$, and $\tilde{a}^{\theta}_{j,k}(q) = 0$ otherwise, where $\underline{q}$ is the lowest allowable quality to be recommended to relevant viewers. The payoff for a creator (θ, j) is piecewise: $u\alpha_k - c(e^{\theta}_j)$ if $e^{\theta}_j \geqslant \underline{q}/\theta$, and $-c(e^{\theta}_j)$ if $e^{\theta}_j < \underline{q}/\theta$. A creator prefers the minimum qualifying effort if and only if $\theta \geqslant \underline{\theta}_k \equiv \frac{\underline{q}}{c^{-1}(u\alpha_k)}$. On the other side, a viewer k's payoff is reduced to $\alpha_k(m_c(1-F(\underline{\theta}_k))(\underline{q}/N - t/N) - t\tilde{A}_k)$. In order for $\alpha_k > 0$ for any k, a necessary condition is $\underline{q} \geqslant t$. In fact, as long as $\underline{q} \geqslant t$, the platform is able to insert $\tilde{A}_k \geqslant 0$ while keeping $\alpha_k = 1$. Because of perfect targeting, the quality requirement is much lower than the last example. Targeting thus expands participation at the cost of quality, but eliminates irrelevant attention so as to insert more ads. The problem of this simple algorithm is that creators, especially high ability ones, get away with too low effort. They should be motivated to produce better contents, and a useful incentive is the promise of *irrelevant* attention. This hurts the targeting accuracy, but benefits the creators and motivates higher effort. These trade-offs are explored and balanced in the optimal algorithm developed in the next section.

4 Analysis

In this section we solve the profit maximization problem for the platform, and then contrast it with a hypothetical welfare maximization problem for a benevolent platform.

4.1 Simplifying the Problem

First, Revelation principle (Myerson, 1986) can reduce the space of algorithms to *obedient algorithms* without loss of generality.

Definition 2 (Obedient Algorithm)
An obedient algorithm consists of an effort recommendation $e_j^\theta : \mathbb{R}_+ \times \mathcal{N} \to \mathbb{R}_+$, *attention for content* $a_{j,k}^\theta : \mathbb{R}_+ \times \mathcal{N}^2 \to [0,1]$ *and attention for ads* $A_k : \mathcal{N} \to \mathbb{R}_+$, *such that:*

$$u \sum_k \nu_k a_{j,k}^\theta - c(e_j^\theta) \geqslant 0, \ \forall\, \theta, j, \tag{1}$$

$$\mu_k \int \theta a_{k,k}^\theta e_k^\theta \mathrm{d}F(\theta|k) - t \cdot \left[\sum_j \mu_j \int a_{j,k}^\theta \mathrm{d}F(\theta|j) + A_k \right] \geqslant 0, \ \forall\, k. \tag{2}$$

The first set of constraints, (1), are the obedience constraints for creators. When following the recommendation, their individual payoff must be greater than their outside option of zero. These constraints coincide with individual rationality for creators because the platform observes creator's ability and content quality, and whenever the realized quality q does not equal the recommended level θe_j^θ, the platform can use the harshest punishment $\tilde{a}_{j,k}^\theta(q) = 0$ for all $k \in \mathcal{N}$, resulting in a non-positive payoff. The second set of constraints, (2), represents the obedience constraints for viewers. These constraints ensure that, for every viewer type, the (constant) marginal benefit of spending more attention on the recommendation set exceeds the (constant) marginal cost, making it optimal for them to fully consume the recommended contents and ads. These constraints also coincide with individual rationality for viewers.

Lemma 1 (Revelation Principle)
For any equilibrium outcome induced by some algorithm, there exists an obedient algorithm inducing the same outcome.

Among obedient algorithms, the optimization problem can be written as:

$$\max_{\substack{\{e_j^\theta\}_{\theta,j} \geqslant 0, \{A_k\}_k \geqslant 0 \\ \{a_{j,k}^\theta\}_{\theta,j,k} \in [0,1]}} z \sum_k \nu_k A_k \quad \text{s.t.} \quad (1), (2).$$

Next, we make two observations of the constraints, (1) and (2). First, the viewers' obedience constraints (2) must be binding at optimum. If not for some k, then the platform should simply increase A_k to improve profits. Second, it is without loss of generality to require the creators' obedience constraints (1) to be binding. If not for some creator type (θ, j), then the platform can require a higher e_j^θ and weakly relaxing (2). This is summarized in the following lemma.

Lemma 2 (Binding Constraints)
There exists an optimal algorithm where the obedience constraint (1) *and the obedience constraint* (2) *are both binding.*

With the binding constraints, we can substitute e_j^θ and A_k, omit the positive multiplier $\frac{z}{t}$, and rewrite the problem as:

$$\max_{\substack{\{e_j^\theta\}_{\theta,j}\geqslant 0 \\ \{a_{j,k}^\theta\}_{\theta,j,k}\in[0,1]}} \quad \sum_k \nu_k\mu_k \int \theta a_{k,k}^\theta e_k^\theta \mathrm{d}F(\theta|k) - t\cdot\left[\sum_k\sum_j \nu_k\mu_j \int a_{j,k}^\theta \mathrm{d}F(\theta|j)\right] \tag{3}$$

$$\text{s.t.} \quad e_j^\theta = c^{-1}\left(u\sum_k \nu_k a_{j,k}^\theta\right),\ \forall\,\theta, j, \tag{4}$$

$$\mu_k \int \theta a_{k,k}^\theta e_k^\theta \mathrm{d}F(\theta|k) - t\cdot\left[\sum_j \mu_j \int a_{j,k}^\theta \mathrm{d}F(\theta|j)\right] \geqslant 0,\ \forall\, k. \tag{5}$$

The objective (3) is the total ads inserted in the recommendation set of all viewers. If the total benefit from reading is higher than the total attention cost from reading contents, then there is room for the platform to sneak in more ads. The constraint (4) is a rewriting of the binding obedience constraint. The constraint (5) is the previous non-negativity constraint on A_k. It appears here because A_k is no longer an explicit variable in the optimization.

4.2 Optimal Algorithm

To understand the platform's trade off of attention allocation, we examine the impact of adjusting a_{jk}^θ on the platform's advertising revenue. To do so, we replace e_j^θ in (3) by constraint (4) and take first derivative with respect to a_{jk}^θ.

For some generic (θ, j, k), if $q_j^\theta = \theta e_j^\theta = 0$, then the derivative of the objective with respective to $a_{j,k}^\theta$ is

$$-t\nu_k\mu_j f(\theta|j) < 0,$$

regardless of whether $k = j$. This means if the quality of a content is zero, it is never optimal to allocate any attention, be it relevant or irrelevant. This is intuitive. If $q_j^\theta = 0$, it must be either because the creator's ability is $\theta = 0$ or because their effort level is $e_j^\theta = 0$.

In the first case, allocating any attention to the creator has no incentive value. In the second case, since the creator's optimal effort is zero, there is no reason to allocate them any attention.

If $\theta e_j^\theta > 0$, however, the derivative of the objective with respective to $a_{j,k}^\theta$ reads (omitting positive multipliers):

$$\begin{aligned} -t \;+\; \theta e_j^\theta + \;\frac{u}{c'(e_j^\theta)}\theta \nu_j a_{j,j}^\theta \qquad &\text{if } k = j, \\ -t \;+\; \frac{u}{c'(e_j^\theta)}\theta \nu_j a_{j,j}^\theta \qquad &\text{if } k \neq j. \end{aligned}$$

The intuition for the derivative is clear. If $k = j$, then recommending (θ, j)'s content to k has three effects on the quantity of ads. The first term, $-t$, is the attention cost on the content that crowds out attention on ads. The second term, θe_j^θ, is the quality of content, which is also the benefit from reading. This relaxes the obedience constraint and thus allows for more ads. The third term, $\frac{u}{c'(e_j^\theta)}\theta \nu_j a_{j,j}^\theta$, is the most interesting force. By assigning attention, the creator receives higher utility, which in turn allows the platform to extract marginally $\frac{u}{c'(e_j^\theta)}$ more effort. The increased effort then boosts the quality of content by $\frac{u}{c'(e_j^\theta)}\theta$, benefiting all relevant viewers of mass $\nu_j a_{j,j}^\theta$.

If $k \neq j$, then the second term is missing as the viewers do not benefit from reading irrelevant contents. Nevertheless, the third term is still there, meaning that recommending irrelevant contents is not a pure waste of time. By increasing the reach among irrelevant viewers, a creator gains from attention and is willing to work harder subject to the obedience constraint. The resulting higher quality content benefits relevant viewers and relaxes their obedience constraint, thereby admitting more ads. This explains why the algorithm may want to mismatch contents on purpose.

The comparison between the two cases implies that as long as a creator produces content of positive quality, it must first reach to all relevant viewers before starting to reach irrelevant ones. Formally, define

- $\overline{a}_j^\theta \equiv a_{j,j}^\theta \in [0, 1]$ as the *reach* among the relevant viewers, or the attention a type-(θ, j) creator receives from type-j viewers, and
- $\underline{a}_j^\theta \equiv \frac{1}{1-\nu_j}\sum_{k\neq j} \nu_k a_{j,k}^\theta \in [0, 1]$ as the *reach* among the irrelevant viewers, or the attention a type-(θ, j) creator receives from all viewers other than type j.

We conclude the following.

Lemma 3 (Priority)
(i) If $\overline{a}_j^\theta = 0$*, then* $\underline{a}_j^\theta = 0$*;*
(ii) If $\underline{a}_j^\theta > 0$*, then* $\overline{a}_j^\theta = 1$*.*

Furthermore, the objective can now be simplified to:

$$\sum_j \mu_j \int \Big(\nu_j \overline{a}_j^\theta \theta c^{-1}\left(u\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\right) - t\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\Big)\mathrm{d}F(\theta|j), \tag{6}$$

where $\overline{a}_j^\theta \in [0,1]$ and $\underline{a}_j^\theta \in [0,1]$ are the only choice variables, and

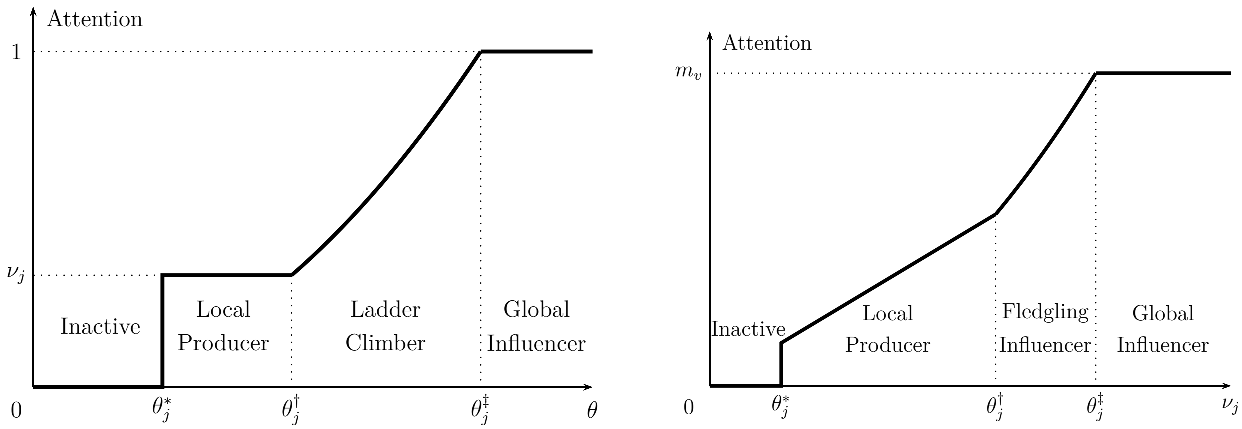
$$\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta$$

is the total attention creator (θ, j) receives. The reformulation reveals that what matters for the optimal attention allocation for each creator type-(θ, j) are the received attention from the relevant viewers and the one from irrelevant viewers. Exactly how to allocate attention over different irrelevant viewer types has no impact. This observation substantially reduces the dimensionality of our optimization analysis.

The objective is concave in $\underline{a}_j^\theta$. This is because in the second term, irrelevant attention comes with a linear cost, while in the first term, it generates a marginally decreasing effect in extracting the creators. In contrast, the objective is convex in $\overline{a}_j^\theta$. While the second term is still linear in $\overline{a}_j^\theta$, the first term is convex as there is complementarity between the attention from a viewer and the effort of a creator. The more relevant attention, the more profitable squeezing effort from creators; the higher effort, the more profitable it is to allocate relevant attention. Therefore, the multiplicative term $\overline{a}_j^\theta c^{-1}\left(u\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\right)$ is convex, given the log-concavity of c.

Therefore, the optimal algorithm must display a bang-bang solution for the reach of relevant contents but could feature interior reach of irrelevant contents. Indeed, this is confirmed in the following statement of the optimal algorithm. We focus on the case of low attention cost, where t is sufficiently small but positive. This guarantees that the constraint (5) does not bind at optimum.

Proposition 1 (Optimal Algorithm)
Suppose the attention cost t *is sufficiently low. For each category* j*, the optimal algorithm partitions abilities into four groups with cutoffs* $0 < \theta_j^* < \theta_j^\dagger < \theta_j^\ddagger$*:*
(1) If $\theta \leqslant \theta_j^*$*, the creator is* inactive, *with* $e_j^\theta = 0$*,* $\overline{a}_j^\theta = 0$ *and* $\underline{a}_j^\theta = 0$*;*
(2) If $\theta_j^* < \theta \leqslant \theta_j^\dagger$*, the creator is a* local producer, *with* $e_j^\theta = c^{-1}(u\nu_j)$*,* $\overline{a}_j^\theta = 1$ *and* $\underline{a}_j^\theta = 0$*;*
(3) If $\theta_j^\dagger < \theta < \theta_j^\ddagger$*, the creator is a* fledgling influencer, *with* $e_j^\theta = c'^{-1}(\theta u \nu_j / t)$*,* $\overline{a}_j^\theta = 1$ *and*



(a) Attention in θ

(b) Attention in ν_j

Figure 1: Effort and attention in the optimal mechanism. Parameter: $c(e) = e^2$. Panel (a): Total attention received as function of ability θ, fixing j. Panel (b): Total attention received as function of popularity ν_j, fixing θ.

$\underline{a}_j^\theta = \frac{c\left(c'^{-1}(\theta u\nu_j/t)\right) - u\nu_j}{u(1-\nu_j)} \in (0,1)$;
(4) If $\theta \geqslant \theta_j^\ddagger$*, the creator is a* global influencer*, with* $e_j^\theta = c^{-1}(u)$, $\overline{a}_j^\theta = 1$ *and* $\underline{a}_j^\theta = 1$.
The cutoffs are: $\theta_j^* = \frac{t}{c^{-1}(u\nu_j)}$, $\theta_j^\dagger = \frac{tc'(c^{-1}(u\nu_j))}{u\nu_j}$, $\theta_j^\ddagger = \frac{tc'(c^{-1}(u))}{u\nu_j}$.

For each category j, the optimal algorithm sorts creators into four segments according to their ability θ. The ones with lowest ability are *inactive creators*, excluded from production because their effort hardly generates any synergy with their ability, which does not justify any attention away from ads. The ones with slightly higher ability are called *local producers*, who puts in the same effort in exchange for attention from and only from relevant viewers. The ones with even higher ability are called *fledgeling influencers* as they not only cater to relevant viewers but also project their influence onto some of the irrelevant viewers. The total attention a creator receives increases in his ability θ, but the utility increase in completely offset by the higher effort level required by the platform. Finally, the ones with the highest ability are called *global influencers*. Their contents penetrate the entire market, relevant and irrelevant alike. Figure 1 plots the total attention $\nu_j\overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta$ a type-(θ, j) creator receives in the optimal algorithm, where $c(e) = e^2$. Panel (a) shows the total attention as an increasing function of ability θ, while Panel (b) plots the total attention as an increasing function of the popularity ν_j of his own category.

As is evident from the optimal algorithm, the platform thrives on irrelevant content recommendations. While irrelevant content and ads are both worthless for the viewers, they serve different roles in the maximization. Ad is the way to cash out the viewers' positive net utility, if any, and has a linear effect on the profit. On the other hand, while

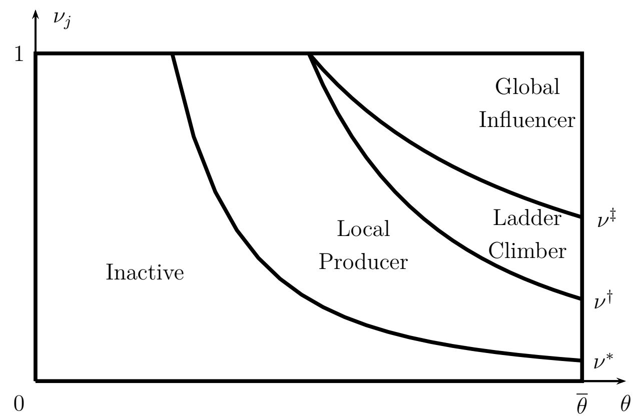


Figure 2: Four segments of creators on a $\nu_j - \theta$ panel, where $\overline{\theta}_j = \overline{\theta}$ for all j. ν^* is the critical mass for a category to exist. $\nu^\dagger$ (resp. $\nu^\ddagger$) is the critical mass for a category to hatch fledgling (resp. global) influencers.

irrelevant contents do not benefit the viewers, they boost the utility of the creators and allows the algorithm to extract more effort from them. This in turn relaxes the obedience constraint of the relevant viewers and makes room for more ads. In other words, the irrelevant content has a non-linear effect on profit after this feedback loop. This also explains why in the optimal algorithm, irrelevant attention increases in the ability θ or popularity ν_j. When assigning more irrelevant attention and extracting higher effort, it is the high-θ creators whose extra effort is most fruitful, and it is the creators in the popular category whose viewers benefit the most.

It is notable in Figure 1 that for fledgling influencers, the total attention grows faster than linearly. This implies that when we compare two creators in the same category, the one with higher ability will gain disproportionally larger attention; the same is true when we compare two creators of the same ability but born in categories of different popularity.

Corollary 1 (Skewness)
For fledgling influencer of type (θ, j), the total attention received is $c(c'^{-1}(\theta u \nu_j/t))$. Moreover, $\frac{c(c'^{-1}(\theta u \nu_j/t))}{\theta}$ increases in θ and $\frac{c(c'^{-1}(\theta u \nu_j/t))}{\nu_j}$ increases in ν_j.

Intuitively, a creator with higher ability or larger relevant audience are required to work harder, and due to the convex effort cost, the algorithm must allocate increasingly more attention to compensate the creators. This disproportional effect resonates well with the empirical finding that the attention distribution on the social platforms is skewed towards high-ability creators and popular categories.

Finally, we would like to discuss the sustainability of categories. Suppose the distribution of ability θ has a bounded support on $[0, \overline{\theta}_j]$ for category j, $j \in \mathcal{N}$. In order for any creator in category j to produce, we require $\overline{\theta}_j c^{-1}(u\nu_j) \geqslant t$, and therefore $\nu_j^* \equiv \frac{c(t/\overline{\theta}_j)}{u}$ is the critical mass for the category to remain active. The platform is viable only if $\nu_j \geqslant \nu_j^*$ for at least one category j. Similarly, in order for a category j to be popular enough to support any global influencers, we need $\frac{\overline{\theta}_j u \nu_j}{c'(c^{-1}(u))} \geqslant t$, and therefore $\nu_j^\ddagger \equiv \frac{tc'(c^{-1}(u))}{u\overline{\theta}_j}$ is the critical mass for this category to hatch a global influencer.

Corollary 2 (Critical Mass)
A category j is active on the platform only if $\nu_j \geqslant \nu_j^*$. A category j hatches fledgling (resp. global) influencers only if $\nu_j \geqslant \nu_j^\dagger$ (resp. $\nu_j \geqslant \nu_j^\ddagger$).

Figure 2 shows the four segments of creators across all possible θ and ν_j, where $\overline{\theta}_j = \overline{\theta}$ for all j. A category has to represent ν^* share of the viewer population in order to survive, and has to house $\nu^\dagger$ (resp. $\nu^\ddagger$) share to become a sufficiently popular category that hatches fledgling (resp. global) influencers.

4.3 Welfare Consequences

Having characterized the optimal algorithm that maximizes the ads income of the platform, we take a detour to contemplate on the welfare consequences of such modern algorithm that prevails the social media. Obviously, in the optimal algorithm, both sides of the users earn zero profit. The creators exert so much effort that they are on the verge of quitting. The viewers watch irrelevant contents and ads to the extent that they barely find the utility from reading worth their time.

Now suppose we consider a hypothetically benevolent platform, utilizing the algorithm to maximize users' welfare. To be specific, the users' welfare is a weight sum of creators' and viewers' net utility. As the weight varies, we obtain the Pareto frontier of what is achievable from an algorithm. Let $w_c, w_v > 0$ denote the Pareto weight on the creators and viewers, respectively.

We look for an obedient mechanism solving the following problem:

$$\max_{\substack{\{e_j^\theta\}_{\theta,j} \geqslant 0 \\ \{\overline{a}_j^\theta, \underline{a}_j^\theta\}_{\theta,j} \in [0,1]^2}} \quad w_c \sum_j \mu_j \int \Big(u\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right) - c\left(e_j^\theta\right) \Big) \mathrm{d}F(\theta|j)$$

$$+ w_v \sum_j \mu_j \int \Big(\nu_j \overline{a}_j^\theta \theta e_j^\theta - t\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right) \Big) \mathrm{d}F(\theta|j) \tag{7}$$

$$\text{s.t.} \quad c\left(e_j^\theta\right) \leqslant u\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right), \ \forall\, \theta, j. \tag{8}$$

Again, we consider the case where t is sufficiently small such that there exist $\{a^\theta_{j,k}\}_{\theta,j,k}$ consistent with $\overline{a}_j$ and $\underline{a}_j$ while (5) does not bind.

It appears that the welfare-maximizing algorithm crucially depends on the Pareto weights, in particular, the ratio of *payoff-adjusted* Pareto weights $\frac{tw_v}{w_c}$. We call it payoff-adjusted because t is a viewer's cost of reading while 1 is a creator's normalized utility from attention. The ratio $\frac{tw_v}{w_c}$ thus weighs the social cost of reading against the social gain from the same action. The next result characterizes welfare-maximizing algorithms under different ratios.

Proposition 2 (Welfare Maximization)
(i) When $\frac{tw_v}{w_c} \geqslant \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$ *for all* $j \in \mathcal{N}$*, the welfare-maximizing algorithm assigns the same* e^θ_j *and* $a^\theta_{j,k}$ *as in the main model, but sets* $A_k = 0$.
(ii) When $1 < \frac{tw_v}{w_c} < \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$ *for some* $j \in \mathcal{N}$*, the welfare-maximizing algorithm has a cutoff ability for local producers lower than that in the main model for category* j*, and some creators exert lower effort and enjoy positive utility.*
(iii) When $\frac{tw_v}{w_c} \leqslant 1$*, the welfare-maximizing algorithm has all* θ *producing, with* $\overline{a}^\theta_j = \underline{a}^\theta_j = 1$.

Part (i) claims that when the benevolent platform sufficiently values the viewer side, the profit-maximizing algorithm can be readily used for welfare maximization too, except that there are no ads inserted in the recommendation. Intuitively, when the creators' utility has a low weight, they will be required to exert effort up to the limit of the participation constraint, which is also the case in the profit maximization. Given zero utility of the creators, the remaining problem is to maximize the viewers' utility, and that coincides with profit maximization too. Indeed, maximum ads is achieved by maximizing viewers' utility before cashing it out by inserting ads. By comparing the two algorithms, the source of inefficiency in the profit maximization is clear: viewers waste time on the ads. Other than that, there is no distortion on the production or the attention allocation on contents.

Part (ii) proposes a different algorithm when creators' utility becomes more important. As the platform now wants to leave some creators a positive net utility, it must be the lowest-ability active creators who should relax. After all, due to their low abilities, reducing the effort level has a smaller impact on the content quality than the same change on a high-ability creator. Moreover, since the entry-level creators only exert a reduced level of effort, it is efficient to set the cutoff ability lower. Therefore, the source of inefficiency in the profit maximization is now two fold. First, the requirement on ability to start producing is too high. Second, the low ability active creators exert too much effort. In short, the profit-maximizing algorithm is not inclusive enough, and is too demanding on the low ability creators.

5 Monetary Incentives

So far we do not allow for monetary incentives for the creators: the only source of utility for creators has been the attention they receive. In this section we explicitly allow the platform to further incentivize the creators by two channels of pecuniary reward. One is the direct payment from the platform (e.g., YouTube), the other is to allow the creators to insert ads into their own contents and receive a payment from the advertisers (e.g., Instagram and Tiktok). For many creators, the monetary incentive has been the primary drive for effort.

To formally incorporate the two monetary channels, we assume that in the obedient algorithm, the platform recommends effort $e_j^\theta \geqslant 0$ and an amount $B_j^\theta \geqslant 0$ of creator-endorsed ads for a type-(θ, j) creator. If the recommendation is followed, it promises attention $a_{j,k}^\theta$ from type-k viewer, $k \in \mathcal{N}$, and a direct payment $\pi_j^\theta \geqslant 0$. The endorsed ads are integrated in their unit of content and hence cannot be skipped by the viewers. Therefore, now a type-k viewer receives a personalized recommendation set that includes organic contents (relevant and irrelevant) integrated with endorsed ads, and platform-inserted ads. Assume the competitive external ads market pays a piece rate $y > 0$ to a creator for each unit of endorsed ads. We allow y to differ from z, the piece rate that the platform receives, because the creator-endorsed ads and the platform-inserted ads can have different effectiveness.

With similar arguments to Lemma 2, the obedience constraints of both creators and viewers are binding. The platform's objective now reads:

$$\begin{aligned} &\frac{z}{t}\sum_j \nu_j\mu_j \int \overline{a}_j^\theta \theta c^{-1}\left[\pi_j^\theta + (u + B_j^\theta y)\left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\right] \mathrm{d}F(\theta|j) \\ &- z\sum_j \mu_j \int \left(\nu_j \overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)(1 + B_j^\theta)\mathrm{d}F(\theta|j) - \sum_j \mu_j \int \pi_j^\theta \mathrm{d}F(\theta|j), \\ \text{s.t.} \quad & \pi_j^\theta \geqslant 0,\ B_j^\theta \geqslant 0. \end{aligned} \tag{9}$$

The objective contains three terms. The first is proportional to the total utility of viewers derived from reading, where the effort is now boosted by the monetary incentives. The second is the attention cost of viewers, bloated by B_j^θ, which crowds out available attention on ads. The third is the total cost of the platform from direct payment.

It turns out that some form of monetary incentive is always optimal for high-ability creators, but generically, only one of the two incentives will be used, depending on the relative effectiveness of creator-sponsored ads. Moreover, the lucrativeness of the ads market indirectly affects the accuracy of targeting contents in the optimal algorithm.

5.1 Lucrative Ads Market

The ads market is considered "lucrative" if $\max\{y, z\} > u$. That is, either the platform-inserted ads or the creator-sponsored ads have a higher piece rate than the creator's ad hoc utility per unit of attention. The next result characterizes the optimal algorithm when both monetary incentives are available to use.

Proposition 3 (Monetary Incentives: Lucrative Ads)
Suppose $\max\{y, z\} > u$ and the attention cost t is sufficiently low. For each category $j \in \mathcal{N}$, there exists some $\hat{\theta}_j^*$ such that:
(i) For $\theta < \hat{\theta}_j^*$, the algorithm stipulates $e_j^\theta = 0$, $\overline{a}_j^\theta = 0$, $\underline{a}_j^\theta = 0$, $B_j^\theta = 0$ and $\pi_j^\theta = 0$;
(ii) For $\theta \geqslant \hat{\theta}_j^*$, the algorithm stipulates $e_j^\theta = c^{-1}(u\nu_j)$, $\overline{a}_j^\theta = 1$ and $\underline{a}_j^\theta = 0$. Moreover:

$$
\begin{aligned}
B_j^\theta &= \frac{1}{y\nu_j} \max\left\{c\left(c'^{-1}\left(\theta y \nu_j/t\right)\right) - u\nu_j, 0\right\}, \quad \pi_j^\theta = 0 \quad \textit{if } y > z, \\
\pi_j^\theta &= \max\left\{c\left(c'^{-1}\left(\theta u z \nu_j/t\right)\right) - u\nu_j, 0\right\}, \qquad B_j^\theta = 0 \quad \textit{if } y < z.
\end{aligned}
$$

According to the proposition, viewers *never* have to spend irrelevant attention on any contents when the ads market, either platform-inserted or creator-sponsored, pays well. The intuition is that if $y > u$ and the algorithm were to recommend some irrelevant contents from (θ, j) to k, then it can instead reduce them and compensate the corresponding creator by allowing them to increase the length of sponsored ads as a more efficient form of incentive. This frees up some attention on the viewer side and enables the platform to insert more ads. Alternatively, if $z > u$, then the platform can reduce irrelevant attention and directly pay creator (θ, j) to make up for the lost attention, and because attention space frees up for some viewer, the platform can insert more ads. In sum, with a lucrative ads market, the algorithm features accurate recommendation without irrelevant attention.

Moreover, due to the linear payoff specification, only the more efficient monetary incentive will be used, not both. The amount of sponsored ads B_j^θ in case of $y > z$, or the transfer π_j^θ in case of $y < z$, grows faster than linearly in the creator's ability θ and the popularity ν_j of the category whenever positive, similar to that of the irrelevant attention in the main model (Corollary 1). Indeed, money plays the role of a cheaper and unbounded substitute of irrelevant attention in incentivizing creators.

5.2 Meager Ads Market

The ads market is considered "meager" if $\max\{y, z\} < u$, that is, earnings from both form of ads are low compared to the direct utility from attention. The next result characterizes the optimal algorithm in this case, which features a fifth segment of creators: *paid* global influencers.

Proposition 4 (Monetary Incentives: Meager Ads)
Suppose $\max\{y, z\} < u$ *and the attention cost* t *is sufficiently low. The optimal algorithm is the same as in Proposition 1 except that among global influencers:*

$$B_j^\theta = \frac{1}{y} \max \left\{ c \left(c'^{-1} \left(\theta y \nu_j / t \right) \right) - u, 0 \right\}, \quad \pi_j^\theta = 0 \quad \textit{if } y > z,$$
$$\pi_j^\theta = \max \left\{ c \left(c'^{-1} \left(\theta u z \nu_j / t \right) \right) - u, 0 \right\}, \quad B_j^\theta = 0 \quad \textit{if } y < z.$$

The proposition claims that the platform should max out irrelevant attention before using either type of monetary incentives, if the ads market does not pay well. Indeed, holding off irrelevant attention in exchange for more ads is not profitable, and the algorithm should prioritize irrelevant attention as the more efficient form of incentive. As an immediate implication of a meager ads market, the algorithm deliberately uses *inaccurate targeting* as in the main model, and monetary incentives require a much higher threshold on creators to arrive.

Note that positive payment or sponsored ads is not guaranteed for all global influencers, due to $\max\{y, z\} < u$. In fact, there is a segment of unpaid global influencers because although they deserve full attention from irrelevant viewers, their content quality is not high enough to justify monetary incentives which has a discretely higher cost.

When $y < z < u$, this result aligns with the spirit of some ad revenue sharing programs in practice, where only high-end creators receives monetary from the platform. Youtube Partner Program sets a threshold on total views and number of subscribers, and pays only creators that meets the requirement. TikTok Pulse splits ad revenue only with creators whose videos rank in the top 4% of all TikTok contents. When $z < y < u$, the algorithm predicts another widely used form of monetization, where high-end creators integrate some sponsored ads within their organic contents (and acquiesced by the algorithm). Instagram and Tiktok thrive from this. Which form of monetary incentive prevails depends on the nature of the platform. On platforms with shorter contents, it is more effective to sponsor ads. Youtube used to feature longer contents and ads are easier to skip.

6 Conclusion

Social media has become an increasingly important part of many people's lives. The platforms of social media also fundamentally shape how information is created and distributed. In this paper we study a model with some crucial aspects of social media such as costly attention, directed attention by algorithm, vanity utility from attention, etc. We argue that since costly attention is a scarce resource to manage, the platform uses algorithms to meticulously allocate attention of the viewers and effort of the creators. When profit maximization is the goal of the designer, the optimal algorithm filters out low-ability creators, restricts medium-ability creators to niche audiences, and amplifies viral content from high-ability creators, creating a skewed distribution of attention. This naturally gives rise to a set of global influencers who is seen by all viewers, even if their specialization (horizontal location) does not necessarily fit all viewers. In contrast, when welfare maximization is the objective, the allocation of attention shifts away from prioritizing viral content and engagement-driven ad revenue. Instead, the platform broadens content exposure to better match viewers with creators who align with their preferences. This leads to a more inclusive ecosystem where low-ability creators are encouraged to participate. While global influencers still emerge, their dominance is reduced as the platform promotes a more diverse content landscape. Additionally, we explore how monetary transfers within the algorithm can mitigate some inefficiencies. Our results offer insights into the economics of content production, distribution, and consumption in digital markets, with direct implications for platform design, creator incentives, and regulatory interventions aimed at improving content allocation and market efficiency. Future research could explore how competition between platforms, alternative monetization models, or policy constraints influence these outcomes in digital markets.

Appendix: Proofs

Proof of Lemma 1. Given an arbitrary algorithm $(\tilde{a}, \tilde{A})$, suppose an equilibrium outcome features effort e_j^θ and reading proportion $\tilde{\alpha}_k$ from the recommendation set. Equilibrium requires:

$$
\begin{aligned}
&e_j^\theta \in \arg\max_{e\geqslant 0} u \sum_k \nu_k \tilde{a}^\theta_{j,k}(\theta e)\tilde{\alpha}_k - c(e), \forall\, \theta, j,\\
&u \sum_k \nu_k \tilde{a}^\theta_{j,k}(\theta e_j^\theta)\tilde{\alpha}_k - c(e_j^\theta) \geqslant 0, \forall\, \theta, j,\\
&\tilde{\alpha}_k \in \arg\max_{\alpha\in[0,1]} \alpha\mu_k \int \theta \tilde{a}^\theta_{k,k}(\theta e_k^\theta)e_k^\theta \mathrm{d}F(\theta|k) - \alpha t \cdot \left[\sum_j \mu_j \int \tilde{a}^\theta_{j,k}(\theta e_j^\theta)\mathrm{d}F(\theta|j) + \tilde{A}_k\right], \forall\, k,\\
&\tilde{\alpha}_k\mu_k \int \theta \tilde{a}^\theta_{k,k}(\theta e_k^\theta)e_k^\theta \mathrm{d}F(\theta|k) - \tilde{\alpha}_k t \cdot \left[\sum_j \mu_j \int \tilde{a}^\theta_{j,k}(\theta e_j^\theta)\mathrm{d}F(\theta|j) + \tilde{A}_k\right] \geqslant 0, \forall\, k.
\end{aligned}
$$

Now consider a new mechanism, in which the platform recommends effort e_j^θ and reading proportion $\alpha_k = 1$, and promises attention recommendation:

$$
a^\theta_{j,k}(q) = \begin{cases} \tilde{a}^\theta_{j,k}(\theta e_j^\theta)\tilde{\alpha}_k & \text{if} \quad q = \theta e_j^\theta, \\ 0 & \text{if} \quad q \neq \theta e_j^\theta. \end{cases},
$$

and ads recommendation $A_k = \tilde{A}_k \alpha_k$. Under the new mechanism, all constraints are satisfied:

$$
\begin{aligned}
&e_j^\theta \in \arg\max_{e\geqslant 0} u \sum_k \nu_k a^\theta_{j,k}(\theta e) - c(e), \forall\, \theta, j,\\
&u \sum_k \nu_k a^\theta_{j,k}(\theta e_j^\theta) - c(e_j^\theta) \geqslant 0, \forall\, \theta, j,\\
&1 \in \arg\max_{\alpha\in[0,1]} \alpha\mu_k \int \theta a^\theta_{k,k}(\theta e_k^\theta)e_k^\theta \mathrm{d}F(\theta|k) - \alpha t \cdot \left[\sum_j \mu_j \int a^\theta_{j,k}(\theta e_j^\theta)\mathrm{d}F(\theta|j) + A_k\right], \forall\, k,\\
&\mu_k \int \theta a^\theta_{k,k}(\theta e_k^\theta)e_k^\theta \mathrm{d}F(\theta|k) - t \cdot \left[\sum_j \mu_j \int a^\theta_{j,k}(\theta e_j^\theta)\mathrm{d}F(\theta|j) + A_k\right] \geqslant 0, \forall\, k.
\end{aligned}
$$

Finally, rewrite $a^\theta_{j,k} \equiv a^\theta_{j,k}(\theta e_j^\theta)$ to save notations. ■

Proof of Lemma 2. Suppose (2) is strict for some k. Then the platform should increase A_k to improve its profit. Now suppose (1) is strict for some (θ, j). Then the platform can increase e_j^θ, which weakly relaxes (2) and weakly improves its profit. Therefore, (2) being binding is necessary for optimization, while (1) being binding is without loss of generality. ■

Proof of Lemma 3. Plug (4) into (3), and then take the derivative w.r.t. $a^\theta_{j,k}$. If $\theta e_j^\theta = 0$, then the derivative reads $-t\nu_k\mu_j f(\theta|j) < 0$, and we must have $a^\theta_{j,k} = 0$ for all k.

If $\theta e_j^\theta > 0$, then the derivative reads $\nu_k\mu_j f(\theta|j)\left(-t + \theta e_j^\theta + \frac{u}{c'(e_j^\theta)}\theta\nu_j a^\theta_{j,j}\right)$ if $k = j$, and $\nu_k\mu_j f(\theta|j)\left(-t + \frac{u}{c'(e_j^\theta)}\theta\nu_j a^\theta_{j,j}\right)$ if $k \neq j$.

(i) Suppose $\overline{a}_j^\theta = a_{j,j}^\theta = 0$. If $\theta e_j^\theta = 0$, then $a_{j,k}^\theta = 0$ for all $k \neq j$. If $\theta e_j^\theta > 0$, then $-t + \frac{u}{c'(e_j^\theta)}\theta\nu_j a_{j,j}^\theta < -t + \theta e_j^\theta + \frac{u}{c'(e_j^\theta)}\theta\nu_j a_{j,j}^\theta \leqslant 0$, and again $a_{j,k}^\theta = 0$ for all $k \neq j$. As a result, $\underline{a}_j^\theta = \frac{1}{1-\nu_j}\sum_{k\neq j}\nu_k a_{j,k}^\theta = 0$.

(ii) Suppose $\underline{a}_j^\theta > 0$, then there exists some $k \neq j$ s.t. $a_{j,k}^\theta > 0$, and $\theta e_j^\theta > 0$. Then $0 \leqslant -t + \frac{u}{c'(e_j^\theta)}\theta\nu_j a_{j,j}^\theta < -t + \theta e_j^\theta + \frac{u}{c'(e_j^\theta)}\theta\nu_j a_{j,j}^\theta$. As a result, $\overline{a}_j^\theta = a_{j,j}^\theta = 1$. ■

Proof of Proposition 1. Differentiating (6) w.r.t. $\overline{a}_j^\theta$ and $\underline{a}_j^\theta$, we have respectively:

$$f(\theta|j)\mu_j\nu_j\left(-t + \theta c^{-1}\left(u\left(\nu_j\overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\right) + \frac{\theta u\nu_j\overline{a}_j^\theta}{c'\left(c^{-1}\left(u\left(\nu_j\overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\right)\right)}\right), \quad (10)$$

$$f(\theta|j)\mu_j(1-\nu_j)\left(-t + \frac{\theta u\nu_j\overline{a}_j^\theta}{c'\left(c^{-1}\left(u\left(\nu_j\overline{a}_j^\theta + (1-\nu_j)\underline{a}_j^\theta\right)\right)\right)}\right). \quad (11)$$

Notice that (11) is strictly decreasing in $\underline{a}_j^\theta$ because c' and c^{-1} are both increasing. However, (10) is strictly increasing in $\overline{a}_j^\theta$ because the second derivative reads:

$$\frac{\theta u\nu_j}{c'(e_j^\theta)^3}\left(2c'(e_j^\theta)^2 - c(e_j^\theta)c''(e_j^\theta) + u\underline{a}_j^\theta(1-\nu_j)c''(e_j^\theta)\right) > 0,$$

where the inequality follows from the log-concavity of c. Therefore, the optimizer must feature $\overline{a}_j^\theta \in \{0,1\}$. According to Lemma 3, $\underline{a}_j^\theta > 0$ implies $\overline{a}_j^\theta = 1$, and $\overline{a}_j^\theta = 0$ implies $\underline{a}_j^\theta = 0$.

Then we have potentially four cases. Case 3: $\overline{a}_j^\theta = 1$ and $\underline{a}_j^\theta \in (0,1)$. (11) implies that $\underline{a}_j^\theta = \frac{c\left(c'^{-1}(\theta u\nu_j/t)\right) - u\nu_j}{u(1-\nu_j)}$ and we need $\theta \in (\theta_j^\dagger, \theta_j^\ddagger)$ so that $\underline{a}_j^\theta \in (0,1)$. Moreover, (6) must be higher than when $\overline{a}_j^\theta = \underline{a}_j^\theta = 0$, which boils down to $\frac{\theta u\nu_j}{t}c'^{-1}\left(\frac{\theta u\nu_j}{t}\right) \geqslant c\left(c'^{-1}\left(\frac{\theta u\nu_j}{t}\right)\right)$. This is always true because $xc'^{-1}(x) - c(c'^{-1}(x)) \equiv \int_0^x c'^{-1}(x')\mathrm{d}x' > 0$ for all $x > 0$. Note that $\theta_j^\dagger < \theta_j^\ddagger$ because $\nu_j < 1$ and c' and c^{-1} are strictly increasing.

Case 4: $\overline{a}_j^\theta = 1$ and $\underline{a}_j^\theta = 1$. (11) implies $\theta \geqslant \theta_j^\ddagger$. Moreover, (6) must be higher than when $\overline{a}_j^\theta = \underline{a}_j^\theta = 0$, which boils down to $\theta\nu_j c^{-1}(u) \geqslant t$. This is always true for $\theta \geqslant \theta_j^\ddagger$ because we can set $x = c'(c^{-1}(u))$ and use the inequality $xc'^{-1}(x) - c(c'^{-1}(x)) > 0$ for all $x > 0$.

Case 2: $\overline{a}_j^\theta = 1$ and $\underline{a}_j^\theta = 0$. (11) implies $\theta \leqslant \theta_j^\dagger$. Moreover, (6) must be higher than when $\overline{a}_j^\theta = \underline{a}_j^\theta = 0$, which boils down to $\theta > \theta_j^*$. Note that $\theta_j^* < \theta_j^\dagger$ because we can set $x = c'(c^{-1}(u\nu_j))$ and use the inequality $xc'^{-1}(x) - c(c'^{-1}(x)) > 0$ for all $x > 0$.

Case 1: $\overline{a}_j^\theta = \underline{a}_j^\theta = 0$. We only require (6) to be higher than when $\overline{a}_j^\theta = 1$ and $\underline{a}_j^\theta$ is optimally chosen. If $\underline{a}_j^\theta > 0$, this is impossible from the analysis of Cases 3 and 4. Therefore, $\underline{a}_j^\theta = 0$, and the comparison boils down to $\theta \leqslant \theta_j^*$. Since $t > 0$, we have $\theta_j^* > 0$. ■

Proof of Corollary 1. The total attention for a fledgling influencer (θ, j) is $u(\overline{a}_j^\theta\nu_j + \underline{a}_j^\theta(1-$

$\nu_j)) = c\left(c'^{-1}(\theta u\nu_j/t)\right)$.

Note that:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\frac{c\left(c'^{-1}(\theta u\nu_j/t)\right)}{\theta} = \frac{c'(x)^2 - c(x)c''(x)}{\theta^2 c''(x)} > 0,$$

where $x = c'^{-1}(\theta u\nu_j/t)$. The inequality comes from the log-concavity of c. Considering the symmetry between θ and ν_j in $c\left(c'^{-1}(\theta u\nu_j/t)\right)$, the proof for $\frac{c\left(c'^{-1}(\theta u\nu_j/t)\right)}{\nu_j}$ is similar. ■

Proof of Corollary 2. If a category j supports active creators, we must have $\theta_j^* \leqslant \overline{\theta}_j$. By definition, this means $\overline{\theta}_j c^{-1}(u\nu_j) \geqslant t$, or equivalently $u\nu_j \geqslant c(t/\overline{\theta}_j)$. Therefore, $\nu_j \geqslant \nu_j^*$.

Similarly, if a category j hatches global influencers, we must have $\theta_j^\dagger \leqslant \overline{\theta}_j$. By definition, this means $\frac{\overline{\theta}_j u\nu_j}{c'(c^{-1}(u))} \geqslant t$, or equivalently $u\nu_j \geqslant \frac{tc'(c^{-1}(u))}{\overline{\theta}_j}$. Therefore, $\nu_j \geqslant \nu_j^\ddagger$. ■

Proof of Proposition 2. In welfare maximization, Lemma 3 still holds. Denote the Lagrangian multiplier for (8) as $w_c\lambda_j^\theta \geqslant 0$. The first order condition w.r.t. e_j^θ requires:

$$e_j^\theta = c'^{-1}\left(\frac{\theta u w_v \nu_j \overline{a}_j^\theta}{w_c(1+\lambda_j^\theta)}\right).$$

(i) We first examine conditions under which (8) always holds with equality. Note that when this is the case, the objective reduces to the one in the main model and the candidate solution is the same as in the main model. In particular, $e_j^\theta = c^{-1}(u(\overline{a}_j^\theta \nu_j + \underline{a}_j^\theta(1-\nu_j)))$. When $\theta < \theta_j^*$, we have $\overline{a}_j^\theta = \underline{a}_j^\theta = 0$, and the two expressions for e_j^θ trivially coincide. When $\theta \geqslant \theta_j^*$, we know $\overline{a}_j^\theta = 1$. Then, $\lambda_j^\theta \geqslant 0$ means:

$$c'^{-1}\left(\frac{\theta u w_v \nu_j \overline{a}_j^\theta}{w_p}\right) \leqslant c^{-1}(u(\overline{a}_j^\theta \nu_j + \underline{a}_j^\theta(1-\nu_j)))$$

for all $\theta \geqslant \theta_j^*$. Given the solution to the main model, the necessary and sufficient condition is $\frac{tw_v}{w_c} \geqslant \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$.

(ii) Suppose $1 < \frac{tw_v}{w_c} < \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$ for some j. Plugging in the effort e_j^θ from the first order condition, the derivative of (7) w.r.t. $\underline{a}_j^\theta$ yields $uw_c(1-\nu_j)(1-\frac{tw_v}{w_c}+\lambda_j^\theta)$. If $\theta > \theta_j^\dagger$ and $\lambda_j^\theta = 0$, then $\underline{a}_j^\theta = 0$, $\overline{a}_j^\theta = 1$ and the candidate e_j^θ violates (8). Therefore, $\lambda_j^\theta > 0$ and we end up with Cases 3 and 4 in Proposition 1. If $\frac{w_c}{w_v}\frac{c'(c^{-1}(u\nu_j))}{u\nu_j} < \theta < \theta_j^\dagger$ and $\lambda_j^\theta = 0$, then the same contradiction arises. Therefore, $\lambda_j^\theta > 0$ and we end up with Cases 2 in Proposition 1. If $\theta \leqslant \frac{w_c}{w_v}\frac{c'(c^{-1}(u\nu_j))}{u\nu_j}$, then regardless of λ_j^θ we must have $\underline{a}_j^\theta = 0$. If $\overline{a}_j^\theta = 1$ then (8) is not binding. This, compared to $\overline{a}_j^\theta = 0$, produces a higher objective if and only if $\theta \geqslant \tilde{\theta}_j^*$, where $\tilde{\theta}_j^* > 0$ uniquely solves

$$\int_0^{\tilde{\theta}_j^* w_v u\nu_j/w_c} c'^{-1}(x)\mathrm{d}x = u\nu_j\left(\frac{tw_v}{w_c} - 1\right).$$

Finally, we want to show $\tilde{\theta}_j^* < \theta_j^*$. We know $\tilde{\theta}_j^* > 0$ satisfies:

$$u\nu_j\left(\frac{tw_v}{w_c}-1\right)+c\left(c'^{-1}\left(\frac{\tilde{\theta}_j^* uw_v\nu_j}{w_c}\right)\right)-\frac{\tilde{\theta}_j^* uw_v\nu_j}{w_c}c'^{-1}\left(\frac{\tilde{\theta}_j^* uw_v\nu_j}{w_c}\right)=0.$$

Replacing $\tilde{\theta}_j^*$ with θ_j^* and taking the derivative of the left-hand side w.r.t. t, we have:

$$\frac{uw_v\nu_j}{c^{-1}(u\nu_j)}\left(c^{-1}(u\nu_j)-c'^{-1}\left(\frac{tw_v u\nu_j}{w_c c^{-1}(u\nu_j)}\right)\right)>0$$

for $\frac{tw_v}{w_c} < \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$, and the left-hand side becomes zero when $\frac{tw_v}{w_c} = \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$. Therefore, the left-hand side is negative for $\frac{tw_v}{w_c} < \frac{c'(c^{-1}(u\nu_j))c^{-1}(u\nu_j)}{u\nu_j}$. Since $xc'^{-1}(x) - c(c'^{-1}(x))$ increases in $x > 0$, we know that $\theta_j^* > \tilde{\theta}_j^*$.

(iii) Suppose $\frac{tw_v}{w_c} \leqslant 1$ for some j. Then according to the derivatives of (7) w.r.t. $\overline{a}_j^\theta$ and $\underline{a}_j^\theta$ yields $\overline{a}_j^\theta = \underline{a}_j^\theta = 1$ for all θ. In order for (8) to bind, we need $\theta > \frac{w_c}{w_v}\frac{c'(c^{-1}(u))}{u\nu_j}$. ■

Proof of Proposition 3. The derivatives of the objective w.r.t. $\underline{a}_j^\theta$, B_j^θ, and π_j^θ have the same sign as $\theta\overline{a}_j^\theta\frac{u+yB_j^\theta}{1+B_j^\theta}\nu_j - tc'(e_j^\theta)$, $\theta\overline{a}_j^\theta y\nu_j - tc'(e_j^\theta)$, and $\theta\overline{a}_j^\theta z\nu_j - tc'(e_j^\theta)$, respectively. When $u < \max\{y,z\}$, the derivative w.r.t. $\underline{a}_j^\theta$ is negative for all θ, since the other two derivatives are non-positive. Then $\underline{a}_j^\theta = 0$. Also, the objective is convex in $\overline{a}_j^\theta$ so that $\overline{a}_j^\theta \in \{0,1\}$.

If $\overline{a}_j^\theta = 0$, then B_j^θ is irrelevant, and the derivative w.r.t. π_j^θ requires $\pi_j^\theta = 0$. If $\overline{a}_j^\theta = 1$ and $y > z$, then $\pi_j^\theta = 0$ and $B_j^\theta = \frac{1}{y\nu_j}\max\left\{c\left(c'^{-1}\left(\theta y\nu_j/t\right)\right) - u\nu_j, 0\right\}$. Comparing $\overline{a}_j^\theta = 0$ versus $\overline{a}_j^\theta = 1$, the objective is greater if and only if

$$\theta \geqslant \frac{t(1+B_j^\theta)}{c^{-1}((u+yB_j^\theta)\nu_j)}. \tag{12}$$

If $\overline{a}_j^\theta = 1$ and $y < z$, then $B_j^\theta = 0$ and $\pi_j^\theta = \max\left\{c\left(c'^{-1}\left(\theta uz\nu_j/t\right)\right) - u\nu_j, 0\right\}$. Comparing $\overline{a}_j^\theta = 0$ versus $\overline{a}_j^\theta = 1$, the objective is greater if and only if

$$\theta \geqslant \frac{t(1+\frac{\pi_j^\theta}{z\nu_j})}{c^{-1}(\pi_j^\theta + u\nu_j)}. \tag{13}$$

When $\max\{y,z\} \geqslant \frac{c'(c'^{-1}(u\nu_j))c'^{-1}(u\nu_j)}{\nu_j}$, (12) and (13) must imply $B_j^\theta > 0$ when $y > z$ and $\pi_j^\theta > 0$ when $y < z$. If not, then (12) or (13) implies $\theta > \theta_j^*$. But then we must have $B_j^\theta > 0$ when $y > z$ or $\pi_j^\theta > 0$ when $y < z$, a contradiction. Therefore, $\overline{a}_j^\theta = 1$ whenever $\theta > \hat{\theta}_j^*$, where $\hat{\theta}_j^*$ uniquely solves

$$\int_0^{\hat{\theta}_j^* \max\{y,z\}\nu_j/t} c'^{-1}(x)\mathrm{d}x = \nu_j(\max\{y,z\} - u).$$

When $\max\{y,z\} < \frac{c'(c'^{-1}(u\nu_j))c'^{-1}(u\nu_j)}{\nu_j}$. For $\theta \geqslant \frac{tc'(c^{-1}(u\nu_j))}{y\nu_j}$, (12) and (13) must imply $B_j^\theta > 0$ when $y > z$ and $\pi_j^\theta > 0$ when $y < z$, for the same reason as above. For $\theta_j^* \leqslant \theta < \frac{tc'(c^{-1}(u\nu_j))}{y\nu_j}$, (12) and (13) must imply $B_j^\theta = \pi_j^\theta = 0$. If not, then $B_j^\theta < 0$ when $y > z$ and $\pi_j^\theta < 0$ when $y < z$, a contradiction. For $\theta < \theta_j^*$, (12) and (13) are violated. Then $\hat{\theta}_j^* = \theta_j^*$, coinciding with that in Proposition 1. ■

Proof of Proposition 4. The derivatives of the objective w.r.t. $\underline{a}_j^\theta$, B_j^θ, and π_j^θ have the same sign as $\theta\overline{a}_j^\theta \frac{u+yB_j^\theta}{1+B_j^\theta}\nu_j - tc'(e_j^\theta)$, $\theta\overline{a}_j^\theta y\nu_j - tc'(e_j^\theta)$, and $\theta\overline{a}_j^\theta z\nu_j - tc'(e_j^\theta)$, respectively. When $u > y > z$, the derivative w.r.t. π_j^θ is negative for all θ, since the derivatives w.r.t. B_j^θ is non-positive. Then, $\underline{a}_j^\theta$ must be maxed out at 1 before B_j^θ becomes positive. When $u > z > y$, the derivative w.r.t. B_j^θ is negative for all θ, since the derivatives w.r.t. π_j^θ is non-positive. Since $B_j^\theta = 0$, $\underline{a}_j^\theta$ must be maxed out at 1 before π_j^θ becomes positive. Also, the objective is convex in $\overline{a}_j^\theta$ so that $\overline{a}_j^\theta \in \{0,1\}$.

For $\theta \leqslant \frac{u}{\max\{y,z\}}\theta_j^\ddagger$, if $B_j^\theta > 0$ when $y > z$ or $\pi_j^\theta > 0$ when $y < z$, then $\underline{a}_j^\theta = \overline{a}_j^\theta = 1$. However,

$$
\begin{aligned}
B_j^\theta &= \frac{1}{y}\left(c\left(c'^{-1}\left(\frac{z}{u}\frac{\theta}{\theta_j^\ddagger}c'(c^{-1}(u))\right)\right) - u\right) \leqslant 0,\\
\pi_j^\theta &= c\left(c'^{-1}\left(\frac{y}{u}\frac{\theta}{\theta_j^\ddagger}c'(c^{-1}(u))\right)\right) - u \leqslant 0,
\end{aligned}
$$

a contradiction. Therefore, $B_j^\theta = \pi_j^\theta = 0$ for all $\theta \leqslant \frac{u}{\max\{y,z\}}\theta_j^\ddagger$, and the solution is the same as the main model for these θ.

For $\theta > \frac{u}{\max\{y,z\}}\theta_j^\ddagger$, we must have $B_j^\theta > 0$ when $y > z$ and $\pi_j^\theta > 0$ when $y < z$ from the above expressions. Therefore, $B_j^\theta > 0$ for all $\theta > \frac{u}{\max\{y,z\}}\theta_j^\ddagger$ when $y > z$, and $\pi_j^\theta > 0$ for all $\theta > \frac{u}{\max\{y,z\}}\theta_j^\ddagger$ when $y < z$. ■

References

Acemoglu, D., A. Ozdaglar, and J. Siderius (2024). A model of online misinformation. *Review of Economic Studies 91*(6), 3117–3150.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020). The welfare effects of social media. *American economic review 110*(3), 629–676.

Anderson, S. P. and S. Coate (2005). Market provision of broadcasting: A welfare analysis. *The review of Economic studies 72*(4), 947–972.

Aridor, G. and D. Gonçalves (2022). Recommenders' originals: The welfare effects of the dual role of platforms as producers and recommender systems. *International Journal of Industrial Organization 83*, 102845.

Bar-Isaac, H. and S. Shelegia (2022). *Monetizing steering*. Centre for Economic Policy Research.

Ben-Porat, O. and M. Tennenholtz (2018). A game-theoretic approach to recommendation systems with strategic content providers. *Advances in Neural Information Processing Systems 31*.

Bergemann, D. and A. Bonatti (2024). Data, competition, and digital platforms. *American Economic Review 114*(8), 2553–2595.

Berman, R. and Z. Katona (2020). Curation algorithms and filter bubbles in social networks. *Marketing Science 39*(2), 296–316.

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

Braghieri, L., R. Levy, and A. Makarin (2022). Social media and mental health. *American Economic Review 112*(11), 3660–3693.

Chen, N. and H.-T. Tsai (2024). Steering via algorithmic recommendations. *The RAND Journal of Economics 55*(4), 501–518.

Choi, J. P. and D.-S. Jeon (2023). Platform design biases in ad-funded two-sided markets. *The RAND Journal of Economics 54*(2), 240–267.

Collis, A. and F. Eggers (2022). Effects of restricting social media usage on wellbeing and performance: A randomized control trial among students. *PloS one 17*(8), e0272416.

Condorelli, D. and B. Szentes (2023). Buyer-optimal platform design.

De Corniere, A. and G. Taylor (2019). A model of biased intermediation. *The RAND Journal of Economics 50*(4), 854–882.

Fainmesser, I. P. and A. Galeotti (2021). The market for online influence. *American Economic Journal: Microeconomics 13*(4), 332–372.

Filippas, A., J. J. Horton, and E. Lipnowski (2023). The production and consumption of social media. Technical report, National Bureau of Economic Research.

Ghosh, A. and P. McAfee (2011). Incentivizing high-quality user-generated content. In *Proceedings of the 20th international conference on World wide web*, pp. 137–146.

Hagiu, A. and B. Jullien (2011). Why do intermediaries divert search? *The RAND Journal of Economics 42*(2), 337–362.

Ichihashi, S. and A. Smolin (2023). Buyer-optimal algorithmic consumption. *Available at SSRN 4635866*.

Iyer, G. and Z. Katona (2016). Competing for attention in social communication markets. *Management Science 62*(8), 2304–2320.

Janssen, M., T. Jungbauer, M. Preuss, and C. Williams (2023). Search platforms: Big data and sponsored positions. *CEPR Discussion Papers,,(18639)*.

Jullien, B., A. Pavan, and M. Rysman (2021). Two-sided markets, pricing, and network effects. In *Handbook of industrial organization*, Volume 4, pp. 485–592. Elsevier.

Krasnova, H., T. Widjaja, P. Buxmann, H. Wenninger, and I. Benbasat (2015). Research note—why following friends can hurt you: an exploratory investigation of the effects of envy on social networking sites among college-age users. *Information systems research 26*(3), 585–605.

Lindström, B., M. Bellander, D. T. Schultner, A. Chang, P. N. Tobler, and D. M. Amodio (2021). A computational reward learning account of social media engagement. *Nature communications 12*(1), 1311.

Mitchell, M. (2021). Free ad (vice): internet influencers and disclosure regulation. *The RAND Journal of Economics 52*(1), 3–21.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica: Journal of the Econometric Society*, 323–358.

Peitz, M. and A. Sobolev (2025). Inflated recommendations. *The RAND Journal of Economics forthcoming*.

Peitz, M. and T. M. Valletti (2008). Content and advertising in the media: Pay-tv versus free-to-air. *international Journal of industrial organization 26*(4), 949–965.

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

Ren, Q. (2024). Advertising and content creation on digital content platforms. *Marketing Science 43*(4), 734–750.

Teh, T.-H. and J. Wright (2022). Intermediation and steering: Competition in prices and commissions. *American Economic Journal: Microeconomics 14*(2), 281–321.

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