How Bad is Top-K Recommendation under Competing Content Creators?

Fan Yao [1] **Chuanhao Li** [1] **Denis Nekipelov** [2] **Hongning Wang** [1] **Haifeng Xu** [3]

Abstract

This study explores the impact of content creators' competition on user welfare in recommendation platforms, as well as the long-term dynamics of relevance-driven recommendations. We establish a model of creator competition, under the setting where the platform uses a top-K recommendation policy, user decisions are guided by the Random Utility model, and creators, in absence of explicit utility functions, employ arbitrary no-regret learning algorithms for strategy updates. We study the user welfare guarantee through the lens of Price of Anarchy and show that the fraction of user welfare loss due to creator competition is always upper bounded by a small constant depending on K and randomness in user decisions; we also prove the tightness of this bound. Our result discloses an intrinsic merit of the relevance-driven recommendation policy, as long as users' decisions involve randomness and the platform provides reasonably many alternatives to its users.

1. Introduction

> *"(Producers) are led by an invisible hand to make nearly the same distribution of the necessaries of life... thus without intending it, without knowing it, advance the interest of the society."*
>
> — Adam Smith, *The Theory Of Moral Sentiments*, 1759.

Online recommendation platforms such as Instagram and YouTube have become prevalent in our daily life (Bobadilla et al., 2013). At the core of those platforms is a recommender system (RS) designed to match each user with the most relevant content based on predicted relevance. Such a practice, often referred to as the top-K recommendation, is believed to improve user satisfaction and has served as a rule-of-thumb principle in both academia and industry for decades (Konstan et al., 1997; Koren et al., 2009; He et al., 2017).

Until recently, the community came to realize that users' utilities cannot be maximized unilaterally due to the potential strategic behaviors of content creators (Qian & Jain, 2022). Because the content creators' utilities are directly tied to their content's exposure, they are motivated to adaptively maximize their own utilities. This leads to competition that may potentially harm the social welfare (defined as the total user satisfaction/engagement) (Fleder & Hosanagar, 2009). For example, consider a scenario where the user population contains a large group of sports fans and a small group of travel enthusiasts. Social welfare is maximized when the available content for recommendation covers both topics. However, one possible equilibrium of the competition is that all creators post homogeneous sports content when the gain from creating niche content cannot compensate for the utility loss caused by abandoning the exposure from the majority of users. It is thus urgent to understand in the long run how bad the social welfare loss could be under strategic content creators driven by a top-K RS.

In this work, we propose the *competing content creation game* to model the impact of the creators' competition on user engagement in a top-K RS. We measure the social welfare guarantee through the lens of Price of Anarchy (PoA) (Koutsoupias & Papadimitriou, 1999), which quantifies the inefficiency of selfish behavior by the ratio between the worst-case welfare value of the game's equilibrium and that of an optimal outcome. Some previous works touched upon this question under different competition models, and their answers are all pessimistic. For example, Ben-Porat & Tennenholtz (2018) noticed that the PoA of social welfare under the RS implemented by a Shapley mediator is unbounded. Ben-Porat et al. (2019) studied a competition model in 1-dimensional space and showed that the PoA under the top-1 matching principle could be as bad as a constant 2. These negative results are either based on a deterministic user choice model or assume creators compete for the shares of *content exposure*. We overturn these pessimistic conclusions by showing that the PoA induced by a top-K RS is at most $1 + O(\frac{1}{\log K})$ when (1) $K > 1$, (2) user choices have mild stochastic noises, and (3) creators are incentivized to

[1]Department of Computer Science, University of Virginia, USA [2]Department of Economics, University of Virginia, USA [3]Department of Computer Science, University of Chicago, USA. Correspondence to: Fan Yao <fy4bc@virginia.edu>, Hongning Wang <hw5x@virginia.edu>, Haifeng Xu <haifengxu@uchicago.edu>.

Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

compete for *user engagement* instead of content exposure. We also prove its tightness by analyzing a lower-bound instance. Thus an RS under these assumptions will approach the optimal efficiency (i.e., PoA ratio approaches 1) when K grows, though at a relatively slow rate of $1/\log K$. Notably, our PoA upper bound also holds in dynamic settings where creators gradually learn to improve their strategies in an online fashion. Extensive synthetic and real-world data based simulations also support these theoretical findings. Overall, our results robustly demonstrate that content creation competitions are efficient under properly set incentives. This echoes the famous insight of Adam Smith, as cited at the beginning of the section, about the market's "invisible hand" on driving socially efficient production of goods.

Our results rely on three key assumptions, all of which find their roots in recommendation literature and practice. First, on the platform side, we assume the top-K RS is based on a relevance function that best predicts user satisfaction if recommended content is consumed. To simplify our setting, we assume the true relevance function is known to the RS, since a tremendous amount of research has been spent on this aspect (Bobadilla et al., 2013; Konstan et al., 1997; Koren et al., 2009; He et al., 2017) and the goal of our study is not to improve its estimation. Second, on the user side, we employ the well-established *Random Utility (RU) model* (Manski, 1977) to specify the distributional structure of a user's choices and resulting utility when presented with a list of recommendations. The RU model has been widely adopted and found its success in marketing research to model consumer choices (Baltas & Doyle, 2001). Third, on the creator side, we assume that their utilities collected from matching their content with a user are proportional to the user's utility, as it is a common practice by platforms to set revenue sharing with content creators proportional to the user's satisfaction or engagement (Meta, 2022; Savy, 2019; Youtube, 2023; TikTok, 2022). When we move on to the dynamic setting where the creators do not have oracle access to their utility functions, we allow creators to adopt arbitrary no-regret learning algorithms, which cover a variety of rational learning behaviors.

2. Related Work

The theoretical studies of content creators' strategic behavior under the mediation of an RS date back to the seminal works from Ben-Porat & Tennenholtz (2018; 2017), where they extended the game setting in search and ranking systems (Ben Basat et al., 2015; Raifer et al., 2017) and proposed an RS based on Shapley value that leads to the unique PNE and several fairness-related requirements. However, they showed that the social welfare under the proposed Shapley mediator could be arbitrarily bad.

Another line of work studies the RS with strategic content creators under the Hotelling's spatial competition framework (Hotelling, 1929). First introduced by Hotelling (1929), Hotelling's model studied two restaurants trying to determine their locations to attract customers who are evenly distributed on the segment $[0, 1]$. The Nash equilibrium (NE) of the resulting game is that both restaurants locate at the center, known as the "principle of minimum differentiation". Recently, Shen & Wang (2016) proposed a variant of Hotelling's competition in which each player has its attraction region, and they showed that the PoA is 2 in the worse case. We show that their game settings are special cases of our proposed competing content creation game in Appendix H, and thus our main result directly implies their PoA bound. A more closely related work is from Ben-Porat et al. (2019), where they introduce the RS into the competition as a mediator who directs users to facilities. They studied mediators with different levels of intervention and proposed a limited intervention mediator with a good trade-off between social welfare and intervention cost. Interestingly, their game setting under a no-intervention mediator also turns out to be a special case of ours. We also note that the problem settings and theoretical discussions in both (Shen & Wang, 2016) and (Ben-Porat et al., 2019) are limited to pure strategy in 1-dimensional cases with a distance-induced user utility function, while our model and result apply to arbitrary dimensions and a generic form of user utility functions.

Two recent works (Hron et al., 2022; Jagadeesan et al., 2022) studied the supply-side competition where the creators' strategy space is high dimensional. Their models assume creators directly compete for user exposure without considering the role of an RS. They focused on the characterization of NE and the identification of conditions under which specialization among creators' strategies may occur. In contrast, we study the social welfare under the impact of a top-K RS without being limited to the existence of NE, and our result applies to general user utility functions.

3. A Model of Content Creator Competition

In this section, we formalize the competing content creation game. The game $\mathcal{G}$ is defined by a tuple $(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ with the following ingredients:

1. A finite set of users $\mathcal{X} = \{\boldsymbol{x}_j \in \mathbb{R}^d\}_{j=1}^m$, and a set of players (i.e., content creators[1]) denoted by $[n] = \{1, \cdots, n\}$. Each player i can take an action $\boldsymbol{s}_i$, often referred to as a *pure strategy* in game-theoretic literature, from an action set[2] $\mathcal{S}_i \subset \mathbb{R}^d$. $\boldsymbol{s}_i$ can be understood as the embedding for the *type* of content that creator i can

[1]We use these two terms interchangeably when there is no ambiguity.

[2]The action sets are not assumed to be finite and thus can be continuous.

produce. Let $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$ denote the space of joint strategies. As a convention, for any $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$, we use $\boldsymbol{s}_{-i}$ to denote the joint strategy $\boldsymbol{s}$ excluding $\boldsymbol{s}_i$. Moreover, we use $\boldsymbol{\alpha}_i \in \Delta(\mathcal{S}_i)$ to denote a mixed strategy of player i, which is a probability measure with support $\mathcal{S}_i$. Similarly, $\boldsymbol{\alpha} \in \Delta(\mathcal{S})$ is used to represent a (possibly correlated) joint strategy profile distribution over all players.

2. A relevance function $\sigma(\boldsymbol{s}, \boldsymbol{x}) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ which measures the *relevance* between a user $\boldsymbol{x} \in \mathcal{X}$ and content $\boldsymbol{s}$. Without loss of generality, we normalize σ to $[0, 1]$, where 1 suggests perfect matching. We focus on modeling the strategic behavior of creators and thus abstract away the estimation of σ.

3. Recommendation policy: given a joint strategy $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$ for all players, for each user $\boldsymbol{x}_j$, the RS first calculates the relevance scores $\{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j)\}_{i=1}^n$ over all available content and then generates $\mathcal{T}_j(\boldsymbol{s}; K)$, the subset of $\boldsymbol{s}$ containing the top-K recommendations for user j. Formally,

$$\mathcal{T}_j(\boldsymbol{s}; K) = \{\boldsymbol{s}_{l_i} | i = 1, \cdots, K\}, \tag{1}$$

where $(l_i)_{i=1}^n$ is a permutation of $[n]$ such that $\sigma(\boldsymbol{s}_{l_1}, \boldsymbol{x}_j) \geq \sigma(\boldsymbol{s}_{l_2}, \boldsymbol{x}_j) \geq \cdots \geq \sigma(\boldsymbol{s}_{l_n}, \boldsymbol{x}_j)$.[3]

4. User utility and choice model: we employ the widely adopted random utility (RU) model to capture users' utility and choices of recommendations. Formally, the RU model assumes that the utility for user $\boldsymbol{x}_j$ to consume content $\boldsymbol{s}_i$ is $\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i$, where ε_i is a noise term containing any additional uncertainty that cannot be captured by the RS's prediction $\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j)$ (e.g., user's mood at that moment). The RU model assumes that $\{\varepsilon_i\}$ are i.i.d. random, which are often assumed to follow the *Gumbel distribution* with cumulative distribution function $\text{Gumbel}(\mu, \beta) = e^{-e^{-\frac{x-\mu}{\beta}}}$.[4] We further assume ε_i is zero mean, thus implying $\mu = -\beta\gamma$ where $\gamma \approx 0.577$ is the Euler–Mascheroni constant. The variance of $\text{Gumbel}(-\beta\gamma, \beta)$ is $\frac{\pi\beta}{\sqrt{6}}$ and the parameter β measures the noise level.

   Upon receiving the recommended list $\mathcal{T}_j(\boldsymbol{s}; K)$, user j chooses $i_j^* \in \mathcal{T}_j(\boldsymbol{s}; K)$ that maximizes her utility:

$$i_j^* = \arg \max_{\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s}; K)} \{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i\}. \tag{2}$$

   Note that i_j^* is random, with randomness inherited from $\{\varepsilon_i\}$. Consequently, user j derives the following expected utility π_j from consuming the selected content

$$\pi_j(\boldsymbol{s}) \triangleq \mathbb{E}_{\{\varepsilon_i\}} \left[\max_{\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s}; K)} \{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i\} \right]. \tag{3}$$

5. Player utilities: following the convention, we assume that each player i's expected utility is the sum of the utilities from users that i served, i.e.,

$$u_i(\boldsymbol{s}) = \sum_{j=1}^m \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \cdot \mathbf{Pr}[\boldsymbol{x}_j \to \boldsymbol{s}_i], \tag{4}$$

   where "$\boldsymbol{x}_j \to \boldsymbol{s}_i$" denotes the event i_j^* defined in (2) equals i. Elegantly, $\mathbf{Pr}[\boldsymbol{x}_j \to \boldsymbol{s}_i] \propto e^{\beta\sigma(\boldsymbol{x}_j, \boldsymbol{s}_i)}$ for any $i \in \mathcal{T}_j(\boldsymbol{s}; K)$ (McFadden, 1974) and $\mathbf{Pr}[\boldsymbol{x}_j \to \boldsymbol{s}_i] = 0$ if $i \notin \mathcal{T}_j(\boldsymbol{s}; K)$.

6. Social welfare: the social welfare function is defined as the total utilities from all the users:

$$W(\boldsymbol{s}) = \sum_{j=1}^m \pi_j(\boldsymbol{s}). \tag{5}$$

   Note that under the player utility function (4), we have $W(\boldsymbol{s}) = \sum_{i=1}^n u_i(\boldsymbol{s})$. That is, the social welfare is also the total utility of players.

We remark that the player i's utility defined in (4) depends on not only the proportion of users matched with i, but also the user's engagement reflected in the term $\mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i]$. This differs crucially from the settings in Hotelling's competition (Hotelling, 1929) and its recent applications to recommender systems (Shen & Wang, 2016; Ben-Porat et al., 2019; Hron et al., 2022; Jagadeesan et al., 2022), where players' utilities are set to the total *user exposure*, i.e., total number or proportion of user visits (regardless of how satisfied the users are with the recommendations). Both metrics have been widely used in current industry practice to reward creators (Meta, 2022; Savy, 2019). In this paper, we primarily consider user engagement (i.e., the previously less studied case) as the creator's utility, and in Section 4.2 we will compare it with the *user exposure* metric to highlight their different impact.

Our research question and equilibrium concept. We are particularly interested in quantifying the average social welfare when creators learn to update their strategies adaptively. Specifically, we consider the repeated form of a competing

[3] When $(l_i)_{i=1}^n$ is not unique, $\mathcal{T}_j(\boldsymbol{s}; K)$ can be the top-K truncation of any such permutation with equal probability.

[4] There are many natural reasons to use the Gumbel noise model. This noise model is nearly indistinguishable from a Gaussian distribution empirically, but has slightly thicker tails, allowing for more aberrant user behavior. The RU model with Gumbel noise is also known as the multinomial logit model (McFadden, 1974). It deeply connects to the discrete choice model (McFadden, 1984), quantal response equilibrium to capture bounded rational behaviors (McKelvey & Palfrey, 1995), and entropy regularizer for optimizing randomized strategies (Ling et al., 2018).

content creation game played by n creators over a period of time T. At each time t, each creator chooses an action, observes the utility induced by all creators' strategies at that round, and uses the feedback to adjust their subsequent actions. Naturally, creators aim to optimize their accumulated utility over the course of interactions. However, in real-world online recommendation platforms, creators can only evaluate the utility of their chosen actions and have to gradually learn their optimal strategies through trial and error with such limited information (i.e., bandit feedback). A natural notion for capturing the "reasonable" learning behavior under such an environment is *no regret*. The (external) regret $R_i(T)$ for player i is defined as the difference between her optimal utility in hindsight and the realized accumulated utilities, i.e.,

$$R_i(T) = \max_{\boldsymbol{s}'_i} \sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}_{-i}\sim\boldsymbol{\alpha}^t_{-i}}[u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i})] - \sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[u_i(\boldsymbol{s})] \tag{6}$$

where $\boldsymbol{\alpha}^t = \prod_{i=1}^n \boldsymbol{\alpha}_i^t$ denotes the joint-strategy distribution at time t. Player i's learning has no regret if $R_i(T) = o(T)$, or equivalently, the average regret $R_i(T)/T \to 0$ as T goes to infinity. Note that such no-regret algorithms exist since any no-regret adversarial online learning algorithm (e.g., Exp3 in bandit literature (Auer et al., 2002)) guarantees no regret in such a multi-agent learning setup.

To characterize the outcome under no-regret learning players, we focus on an equilibrium concept termed coarse correlated equilibrium (CCE), as it is well known that the empirical action distribution of any no-regret playing sequence in a repeated game converges to its set of CCEs (Blum et al., 2008). The formal definition of CCE is as follows:

Definition 1. A *coarse correlated equilibrium* (CCE) is a distribution $\boldsymbol{\alpha}$ over the space of joint-strategy profile $\mathcal{S}$ such that for every player i and every action $\boldsymbol{s}'_i \in \mathcal{S}_i$,

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s})] \geq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i})]. \tag{7}$$

Thanks to the nice connection between no-regret dynamics and CCE, we first establish the welfare guarantee for CCE in Section 4.1 and then extend it to account for the accumulated welfare induced by repeated plays in Section 4.2.

We also note that the concept of CCE is particularly useful for two additional reasons. First, CCE always exists in any finite games (thus in our game), hence eliminating the necessity to address the existence of Nash equilibrium (NE), perhaps the most celebrated solution concept, as in previous research (Hron et al., 2022). In fact, when the action sets are continuous, the existence of NE (either pure or mixed) cannot be guaranteed in our game as the player utility function defined in (4) is not continuous. This is an inherent challenge of the problem, as any change in $\sigma(\boldsymbol{s}, \boldsymbol{x})$ may result in a different top-K recommendation list $\mathcal{T}_j(\boldsymbol{s}; K)$, leading to dramatically different player utilities. Similar challenges and the non-existence of mixed NE have also been observed by Hron et al. (2022), though their utility model and research questions differ from ours. Second, even in situations where NE exists, it is more realistic to assume that players eventually achieve some CCE rather than NE due to various criticisms about NE, including the computational concerns (Daskalakis et al., 2009) of NE.

4. The Price of Anarchy Analysis

We analyze the social welfare of any top-K RS under any possible CCE; or more specifically, *how bad can the welfare possibly be due to the competition among self-interested content creators* – compared to the *idealized* non-strategic situation in which the platform can "dictate" all creators' content choices and thus globally optimize the welfare function (5). This can be captured by the celebrated concept of the Price of Anarchy (PoA) (Koutsoupias & Papadimitriou, 1999). As its name indicates, PoA captures the welfare inefficacy due to players' selfish behavior. Our main result in this section is a comprehensive characterization of the PoA of competing content creation games.

Definition 2 (PoA under CCE)**.** Define the price of anarchy of a game $\mathcal{G}$ as

$$PoA(\mathcal{G}) = \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]}, \tag{8}$$

where CCE($\mathcal{G}$) is the set of CCEs of $\mathcal{G}$.

By definition, PoA($\mathcal{G}$) ≥ 1 always holds and larger values indicate worse welfare. Our choice of the CCE concept leads to the strongest possible welfare guarantee in the sense that any upper bound of PoA under CCE also trivially holds for the PoA under refined solution concepts such as correlated equilibrium (CE), PNE or mixed NE (if they exist), since these are all CCEs as well. Unless otherwise emphasized, any PoA in this paper refers to the PoA under CCE.

4.1. Matching PoA Upper and Lower Bounds

Our main theoretical findings are an upper and lower bound for the PoA, which match with each other and thus demonstrates the tightness of our analysis. We first present the upper bound as follows.

Theorem 1. *The PoA of any competing content creation game instance $\mathcal{G}$ with parameter $\beta \geq 0$ and $K \geq 1$ satisfies*

$$PoA(\mathcal{G}) < 1 + \frac{1}{c(\beta, K)}, \tag{9}$$

where $c(\beta, K)$ is defined as

$$c(\beta, K) = \frac{(b+1)\log(b+K)}{(b+K)(\log(b+K) - \log K)}, b = e^{\frac{1}{\beta}} - 1. \tag{10}$$

The proof of Theorem 1 is intricate and thus the detailed arguments are relegated to Appendix A. The primary challenge in the proof is to analyze various smoothness properties of the welfare and players' utility functions, especially how the welfare function changes after excluding any player i's participation. In Section 4.3, we highlight some of the noteworthy properties of the welfare function, including its submodularity, which we develop en route to proving Theorem 1 but is also of independent merit towards understanding the competing content creation game.

The format of $c(\beta, K)$ may not be intuitive enough for the readers to appreciate the derived PoA upper bound. We thus provide the following observations, which reveal various properties of $c(\beta, K)$ aiding the interpretation of (9):

1. For any $\beta > 0$ and $K \geq 1$, we have $c(\beta, K) \geq 1$ and thus $\text{PoA}(\mathcal{G}) < 2$ always holds.
2. $c(\beta, K) = 1$ if and only if $K = 1$ or $\beta \to 0$.
3. Fix any $\beta > 0$, $c(\beta, K)$ monotonically increases in K; similarly, fix any $K \geq 1$, $c(\beta, K)$ monotonically increases in β.
4. For sufficiently large β and K, $c(\beta, K) \approx (1+\beta)\log K$ asymptotically, and therefore

$$PoA(\mathcal{G}) < 1 + \frac{1}{(1+\beta)\log K}. \tag{11}$$

Based on these observations, Theorem 1 has multiple interesting and immediate implications. First, the welfare loss under any CCE is at most half in any situation, as the PoA is always upper bounded by 2. The second and third facts above show that such worst-case PoA occurs and only occurs when users' choices are made in a "hard" manner: either the RS dictates the user's choice by setting $K = 1$ or the randomness in users' choices is extremely low (i.e., $\beta \to 0$). Note that in the latter case, the user will only consume the most relevant content (i.e., the top-ranked content) due to small decision randomness.

Second, the welfare guarantee improves as either K increases (i.e., more items are recommended) or β increases (i.e., users' choices have more randomness). Welfare improvement in the latter situation is intuitive because when supplied with multiple items, the user can pick the content with large ε_i (i.e., the reward component that is not predictable by the RS) to gain utility. These together reveal an interesting operational insight that when the RS cannot perfectly predict user utility (i.e., $\beta > 0$), providing more items can help improve social welfare. This justifies top-K recommendation and the necessity of diversity in recommendation (Hurley & Zhang, 2011).

Our following second main result shows that this PoA upper bound is tight, up to negligible constants.

Theorem 2. *Given any $0 \leq \beta \leq 1$, $n > 2$ and any $1 \leq K \leq \min\{n-1, e^{\frac{1}{5\beta}}\}$, there exists a competing content creation game instance $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ such that*

$$PoA(\mathcal{G}) > \frac{n-1}{n} + \frac{1}{1+5\beta\log K}. \tag{12}$$

This theorem also implies that the argument we employed for Theorem 1, which is based on the *smoothness* proof developed by Roughgarden (2015), yields a tight PoA bound for our proposed game. The tightness of the smoothness argument is itself an intriguing research question. Only three classes of games are known to enjoy a tight PoA bound derived from the smoothness argument: congestion games with affine cost (Christodoulou & Koutsoupias, 2005), second price auctions (Christodoulou et al., 2008), and the valid utility game (Vetta, 2002), which are all fundamental classes of games. Theorem 2 suggests that our competing content creation game subscribes to this list. The proof of Theorem 2 is to explicitly construct a game instance which provably yields the stated PoA lower bound (see Appendix D).

4.2. Implications of the PoA Bounds

We have discussed some direct implications of Theorem 1. Now we develop new results which are either derived from or can be compared to Theorem 1 and 2. They will reveal additional insights from our main theoretical results.

Welfare implications to *learning* content creators. The PoA bounds presented in Theorem 1 and 2 are based on the assumption that creators are aware of the game parameters and play some CCEs of the game. While CCE is a reasonable equilibrium concept, one potential critique is that to find the CCE, it is assumed that each creator has knowledge about the system parameters (e.g., all other creators' strategies and the σ function), which can be unrealistic. Fortunately, in real-world scenarios where creators utilize no-regret algorithms to play a repeated competing content creation game with bandit feedback, we can still establish a slightly worse PoA upper bound leveraging the fact that the average strategy history of no-regret players converges to a CCE, as shown in the following Corollary 1.

Corollary 1. ***[Dynamic Version of Theorem 1]*** *Suppose each player in a repeated competing content creation game $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ independently executes some no-regret learning algorithm, with worst regret $R(T) = \max_i R_i(T)$ as defined in (6). Then we have*

$$\frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\frac{1}{T}\sum_{t=1}^T \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[W(\boldsymbol{s})]} < 1 + \Big(1 + \frac{n}{\beta\log K}\cdot\frac{R(T)}{T}\Big)\cdot\frac{1}{c(\beta,K)}, \tag{13}$$

where $\boldsymbol{\alpha}^t$ denotes the joint-strategy distribution at step t and $c(\beta, K)$ is the constant defined in (10).

In other words, the average welfare across all rounds $\frac{1}{T}\sum_{t=1}^T \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[W(\boldsymbol{s})]$ is close to the maximum possible

welfare $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$, up to a constant factor. The quantity in the LHS of (13) is also known as the "price of total anarchy" (Blum et al., 2008). It is a substitute for PoA when we want to characterize the welfare of an outcome from repeated play which does not necessarily fall into any equilibrium concept. The proof of Corollary 1 is presented in Appendix B. Because $R(T)/T \to 0$ as $T \to \infty$ for any no-regret algorithm (most no-regret algorithms have $R(T) = O(\sqrt{T})$), the RHS of (13) is still strictly less than 2 for any fixed constants (n, β, K).

Theorem 1 relies on two crucial platform features: 1. the player's utility in (4) is defined as the total user engagement that accounts for the user utility $\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_j$, as opposed to just "user exposure" (i.e., the expected total number of matches); 2. the platform uses the top-K recommendation policy. Next, we illustrate the insights revealed from Theorem 1 with respect to these two key features.

The importance of rewarding user engagement rather than solely exposure. A key reason for the nice PoA guarantee in our competing content creation game is each player i's utility is chosen as the user engagement $\sum_j \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i]\, \mathbf{Pr}(\boldsymbol{x}_j \to \boldsymbol{s}_i)$ in (4), while not the following user exposure metric:

$$\text{User exposure for player } i: \quad \sum_{j=1}^{m} \mathbf{Pr}(\boldsymbol{x}_j \to \boldsymbol{s}_i). \tag{14}$$

Our next result shows that incentivizing creators to maximize user exposure can lead to significantly worse welfare.

Proposition 1. *Let $\tilde{\mathcal{G}}$ denote the variant of the competing content creation game $\mathcal{G} = (\{\mathcal{S}_i\}_{i=1}^n, \{\boldsymbol{x}_j\}_{j=1}^m, \sigma, \beta, K)$ by substituting player utility function in (4) by the above user exposure in (14). Then for any $K \geq 1, 0 \leq \beta \leq \min\{0.14, \frac{1}{5\log K}\}$, there exist $\mathcal{G}$ and $\tilde{\mathcal{G}}$ such that*

$$PoA(\tilde{\mathcal{G}}) > 2 > PoA(\mathcal{G}). \tag{15}$$

Moreover, when $K = 1$ or β approaches 0, $PoA(\tilde{\mathcal{G}})$ can be arbitrarily large.

In stark contrast to Theorem 1 guaranteeing $\text{PoA}(\mathcal{G}) < 2$, Proposition 1 implies the deterioration of user welfare when content creators are incentivized to compete for the expected exposure of their content. However, we find in practice both metrics are used: for example, user engagement has been used more often as a reward metric for established creators, whereas user exposure is used more for new creators (Meta, 2022; Savy, 2019). Our result serves as a theoretical defense for rewarding creators by user engagement if the system aims to improve overall welfare of the users.

To prove Proposition 1, we construct a game instance in which the user welfare at NE is arbitrarily close to zero. Our construction also reveals interesting insights about situations where user welfare can be very bad. Hence, we briefly explain our construction here and leave our formal arguments in Appendix E. Our constructed game has two groups of users: one *dispersed* group that is fine with any content but is never very happy with it (i.e., a low relevant score for all content) and one *focused* group who looks for a specific type of high-quality content (a high relevance score on such content); but only a small group of specialized creators can produce such high-quality content. However, if players are incentivized to compete for exposure, even creators from the small group tend to produce low-quality content that appeals to the dispersed group rather than high-quality content that benefits the focused group. This, in the worse case, can lead to arbitrarily worse welfare for the platform.

The welfare efficiency of top-K recommendation policy. One may wonder whether the top-K recommendation is indeed a good policy for securing the platform's welfare, i.e., is it possible that other recommendation policies (e.g., a probabilistic policy based on Plackett-Luce model (Plackett, 1975; Luce, 1959)) may even lead to better equilibrium outcomes? Our following analysis, as a corollary of Theorem 1, shows that the answer is to some extent *no* since *any* recommendation policy cannot be better than the top-K rule by more than a tiny fraction of the theoretical optimality. We believe this finding also serves as a theoretical justification for the wide adoption of the top-K principle in practice.

Corollary 2. *Consider an arbitrary recommendation policy providing at most K recommendations, which induces a different competing content creation game $\mathcal{G}'$. Let $CCE(\mathcal{G}')$ denote the corresponding CCE set of $\mathcal{G}'$ and $W(\mathcal{G}') = \min_{\boldsymbol{\alpha}\in CCE(\mathcal{G}')} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$ be its worst-case CCE welfare. Then we have*

$$W(\mathcal{G}') \leq W(\mathcal{G}) + W_K^* / \Big(1 + \frac{K\log(K+b)}{K+b}\Big), \tag{16}$$

where W_K^* is the best possible social welfare achieved via any centralized recommendation policy with K slots.

As indicated by (16), the fraction of the loss of welfare is approximately $O(\frac{1}{\log K})$ as $\frac{K}{K+b} \sim O(1)$ when K is large. The proof is straightforward based on Theorem 1 and can be found in Appendix C.

4.3. Proof Highlights of Theorem 1

Our first step is to derive clean characterizations for the game primitives by utilizing properties of Gumbel distribution. The form of the user utility π_j and welfare W are corollaries of RU models (Baltas & Doyle, 2001), however, the closed-form of the creator utility u_i is a new property we derive. Detailed argument is deferred to Appendix A.1.

The main proof of Theorem 1 is based on a smoothness argument framework developed in the seminal work by Roughgarden (2015). For any strategy profile $\boldsymbol{s}$, $W(\boldsymbol{s}) =$

$\sum_i u_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i})$ is its total welfare function. A game is (λ, μ)-*smooth* if $\lambda W(\boldsymbol{s}') - \mu W(\boldsymbol{s}) \leq \sum_i u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i})$ for any $(\boldsymbol{s}, \boldsymbol{s}') \in \mathcal{S}$. Roughgarden (2015) observes that the PoA of any (λ, μ)-smooth game can be upper bounded by $\frac{1+\mu}{\lambda}$. After plugging in the expression of $W(\boldsymbol{s})$, the (λ, μ)-smoothness condition can be re-written as

$$\sum_i [\lambda u_i(\boldsymbol{s}'_i, \boldsymbol{s}'_{-i}) - u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i})] \leq \sum_i \mu u_i(\boldsymbol{s}).$$

Intuitively, the smoothness parameters bound how much *externality* other players' actions (i.e., $\boldsymbol{s}'_{-i}$ or $\boldsymbol{s}_{-i}$) impose on any player i's utility. Moreover, the tighter this bound is, the smoother the game is and the smaller the PoA is. To gain some intuition and also as a sanity check, consider the extreme situation in which each player's utility is not affected by other players' actions at all (i.e., the *no externality* situation), we have $\lambda = 1$ and $\mu = 0$ implying PoA=1. That is, if any player's utility is not affected by others, then self-interested utility-maximizing behaviors also maximize social welfare, which is a straightforward observation. Certainly, we cannot hope for such a nice property to hold in general, but fortunately, many well-known games have been shown to be smooth. For example, second-price auctions are $(1, 1)$-smooth as shown by Christodoulou et al. (2008), congestion games are $(\frac{5}{3}, \frac{1}{3})$-smooth as shown by Roughgarden (2015), and all-pay auctions are $(1/2, 1)$-smooth as shown by Roughgarden et al. (2017).

Hence, the key challenge in proving Theorem 1 is to pin down the tightest possible (λ, μ) parameters for our competing content creation game. This boils down to a fundamental question in top-K RS – i.e., *to what extent does the existence of other competing content creators affect a creator's utility*? To answer this question, we discover multiple interesting properties of the welfare and creator utility functions formulated as follows. Besides proving our main result in Theorem 1, we believe these properties are also of interest for us to understand recommender systems.

Our Lemma 1 demonstrates the *submodularity* of $W(\boldsymbol{s})$. That is, the marginal gain of welfare from adding a new player decreases as the total number of creators increases. Lemma 2 further relates this marginal welfare increase with the added player's own utility. It shows that the increased welfare after introducing a new player i with strategy $\boldsymbol{s}_i$ is at most i's utility under $\boldsymbol{s}_i$, multiplied by a shrinkage factor $c^{-1}(\beta, K) \in (0, 1]$. These two lemmas together allow us to prove that the competing content creation game is $(c^{-1}(\beta, K), c^{-1}(\beta, K))$-smooth, yielding Theorem 1. Detailed proofs are presented in Appendix A.2.

Lemma 1. *[Submodularity of Welfare] For any $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$, let $S = \{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$. The social welfare function defined in Eq (5) is submodular as a set function, i.e., for any $S, \boldsymbol{s}_x, \boldsymbol{s}_y$ it holds that*

$$W(S \cup \{\boldsymbol{s}_x\}) - W(S) \geq W(S \cup \{\boldsymbol{s}_x, \boldsymbol{s}_y\}) - W(S \cup \{\boldsymbol{s}_y\}).$$

Lemma 2. *[Smoothness of Welfare] For any $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$, $i \in [n]$ and $c(\beta, K)$ defined in Eq (10), player-i's utility function $u_i(\boldsymbol{s})$ defined in Eq (4) satisfies*

$$W(\boldsymbol{s}) - W(\boldsymbol{s}_{-i}) \leq c^{-1}(\beta, K) \cdot u_i(\boldsymbol{s}_i; \boldsymbol{s}_{-i}).$$

5. Experiments

To confirm our theoretical findings and also to empirically measure the social welfare induced by creators' competition, we conduct simulations on game instances $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ constructed from two synthetic datasets and the MovieLens-1m dataset (Harper & Konstan, 2015). Before presenting our results, we provide a detailed overview of the simulation environment, including the characteristics of the datasets utilized and the metrics employed for evaluation.

5.1. Specification of Datasets

Synthetic Dataset-1. Dataset-1 simulates the situation where content creators compete over an unbalanced user interest distribution. We construct n user clusters with the largest cluster containing half of the population, and let each strategy from a creator's action set generate content that only appeals to a specific user group.

Specifically, the user population is given by disjoint clusters $\mathcal{X} = \cup_{i=1}^n \mathcal{X}_i$ such that $|\mathcal{X}_1| = \frac{m}{2}$, and the sizes of smaller clusters $|\mathcal{X}_l|$ are sampled uniformly at random such that $\sum_{l=2}^n |\mathcal{X}_l| = \frac{m}{2}$. Players share the same action set $\mathcal{S}_i = \{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$, and the σ function satisfies that for any $i \in [n]$,

$$\sigma(\boldsymbol{s}_i, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in \mathcal{X}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Dataset-1 depends on the randomness of the partition $\cup_{i=1}^n \mathcal{X}_i$.

Synthetic Dataset-2. Dataset-2 simulates the situation where content creators can either "chase the trend" by generating mediocre content or cater to a specific user interest group with high-quality content. Similar to the construction of dataset-1, we let the user population comprise of n clusters and allow each player to take actions targeting at any specific user group. But, in addition, we also allow each player to take a "safe" action $\boldsymbol{s}_0$ by producing some popular content that can satisfy all users to a certain extent δ.

Specifically, the user population is also given by disjoint clusters $\mathcal{X} = \cup_{i=1}^n \mathcal{X}_i$, where the sizes of all clusters $|\mathcal{X}_l|$ are sampled uniformly at random such that $\sum_{l=i}^n |\mathcal{X}_l| = m$. Players share the same action set $\mathcal{S}_i = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$, and the σ function satisfies that for any $i \in [n]$,

$$\sigma(\boldsymbol{s}_i, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in \mathcal{X}_i, i \geq 1 \\ \delta, & \text{if } i = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

Table 1. PoA under $\beta = 0.1$. Results reflect the worst cases obtained from 10 independently sampled game instances.

K \ n	$*$	2	3	4	5
1	2.00	1.33	1.54	1.66	1.72
2	1.93	1.28	1.46	1.56	1.60
3	1.89		1.42	1.47	1.51
4	1.86			1.43	1.42
5	1.84				1.42

$*$ denotes the theoretical upper bound.

Table 2. PotA under $\beta = 0.1$. Results reflect the worst cases obtained from 10 independently sampled game instances.

K \ n	$*$	5	10	15	20	40
1	2.00	1.59	1.59	1.60	1.50	1.38
3	1.89	1.37	1.39	1.42	1.41	1.32
5	1.84	1.35	1.34	1.33	1.36	1.31
7	1.80		1.30	1.31	1.30	1.29

$*$ denotes the theoretical upper bound.

Dataset-2 depends on the randomness of the partition $\cup_{i=1}^{n} \mathcal{X}_i$ and the parameter $\delta \in [0, 1]$.

The Dataset Generated from MovieLens-1m. We use deep matrix factorization (Fan & Cheng, 2018) to train user and movie embeddings targeted at movie ratings from 1 to 5. The total number of users $m = 6040$, the number of movies $k = 3883$, and the embedding dimension is set to $d = 32$. To validate the quality of the trained representations, we first performed a 5-fold cross-validation and obtain an averaged RMSE =0.883 on the test sets, and then train the user/item embeddings with the complete dataset. The resulting user embeddings $\mathcal{X} = \{\boldsymbol{x}_j\}_{j\in[m]}$ are used as the user population. To construct each player-i's action set $\mathcal{S}_i$, we randomly sample 500 vectors from the trained movie embedding set $\mathcal{M}$ ($|\mathcal{M}| = 3883$) independently. To normalize the relevance score to $[0, 1]$, we let $\sigma(\boldsymbol{s}, \boldsymbol{x}) = 1$ when the predicted rating of movie $\boldsymbol{s}$ to user $\boldsymbol{x}$ is at least 4, i.e., $\sigma(\boldsymbol{s}, \boldsymbol{x}) = \mathbb{I}[\langle \boldsymbol{s}, \boldsymbol{x}\rangle \geq 4]$.

5.2. Evaluation Metrics

We use both PoA and PotA in our experiments. The evaluation of PoA requires solving two optimization problems, which are both intractable in general due to the non-concavity of $W(\cdot)$ and the undetermined structure of CCE($\mathcal{G}$). As a result, we use simulated annealing to approach $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$ when the exact computation is intractable. To compute $\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$, we compute its exact solution by solving a linear program with k^n variables and kn constraints (Papadimitriou & Roughgarden, 2005) for small n and a moderate size of action set k. To deal with larger problems, we let each player run Exp3 (Auer et al., 2002) over $T = 5000$ rounds and compute the price of total anarchy PotA($\mathcal{G}$) $= \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[W(\boldsymbol{s})]}$. More details are disclosed in Appendix I.

5.3. Results

Empirical PoA from simulations. We first demonstrate the empirical welfare under different game parameter (n, K, β) for dataset-1. We fix $\beta = 0.1$ and report PoA and PotA under varying n and K. Results are reported in Table 1 and 2. We observe that for fixed n, both PoA and PotA decrease w.r.t. K and β, as revealed in Theorem 1. Furthermore, under fixed (β, K), PoA approaches its theoretical upper bound as n increases. However, PotA follows this trend for values of n less than 15, but begins to decrease as n increases further. This discrepancy can be attributed to the fact that for larger values of n (i.e., in Table 2), the approximated optimal welfare becomes less accurate and as such, the PotA tends to be underestimated. We report the results under different β in Appendix J.1.

Comparison between user engagement/exposure metrics. Next we investigate the consequence of utilizing two different incentive metrics, namely *user engagement* vs., *user exposure*. However, Dataset-1 is no longer a good benchmark for this purpose, as the utility functions derived under a simple binary valued $\sigma(\cdot, \cdot)$ are almost indistinguishable under these two metrics. To this end, we use dataset-2, which has a more complex $\sigma(\cdot, \cdot)$ function that models the situation in which creators could focus on chasing the trends other than paying attention to the content quality.

We fix $(\beta, K) = (0.1, 2)$ and report PotA under different n and δ. The results, shown in Figure 1, demonstrate the advantage of using the user engagement metric, which consistently leads to a smaller PotA across different values of n and δ. For n larger than 10, PotA with user-exposure can exceed 2 as revealed by Proposition 1[5]. The performance gap between the two metrics is more distinct when δ gets smaller, which can be understood as when creators can produce popular content with lower effort, simply using exposure to reward creators can be catastrophic to the total user welfare. Additional results under different experimental conditions can be found in Appendix J.2.

Social welfare under different levels of rationality. In this experiment, we aim to investigate the competition outcomes when players utilize online-learning algorithms with varying levels of rationality. To better simulate what happens in practice, we employed the dataset generated from MovieLens-1m (Harper & Konstan, 2015). In our simulation, we model the scenario in which each player runs Exp3 under different exploration rates ϵ (i.e., with probability ϵ, each player will take a random action in each round). We

[5] Again, due to the approximation error in computing optimal W, the PotA could be underestimated as n gets larger.

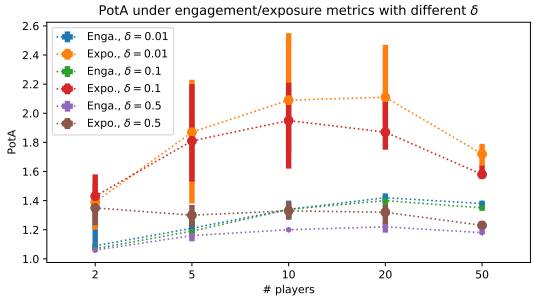


Figure 1. PotA under exposure/engagement metrics with $(\beta, K) = (0.1, 2)$. δ is the relevance score obtained from creators' "safe" action. The error bars indicate the largest/smallest values from 10 independent trials and the dots correspond to the mean values.

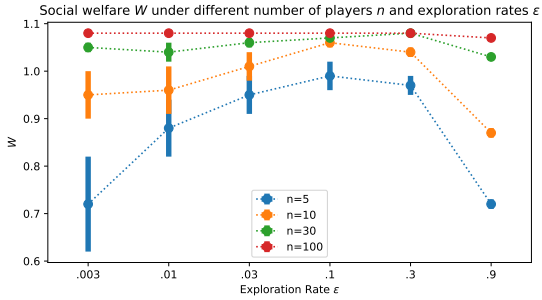


Figure 2. The averaged welfare $\bar{W}$ over $T = 1000$ rounds under different exploration rate ϵ and number of players n. Results are averaged over 10 independent runs under $(\beta, K) = (0.1, 5)$.

use this simulation setup to examine a practical situation, i.e., creators try to optimize their accumulated regret but with bounded rationality: since Exp3 is known to enjoy a sub-linear regret when $\epsilon \sim O(\sqrt{\frac{k \log k}{T}})$ (Auer et al., 2002), it would be less rational for creators to set ϵ to be too large or too small as it would incur a larger regret $R(T)$. We fix $(\beta, K, T) = (0.1, 5, 1000)$ and report the averaged social welfare over T rounds, i.e., $\bar{W} = \frac{1}{mT} \sum_{t=1}^{T} W(\boldsymbol{s}^{(t)})$ under different n and ϵ, as illustrated in Figure 2.

Our results indicate that the optimal exploration rate associated with the maximum welfare is around $\epsilon = 0.1$ across different values of n. When ϵ is set to be either too small or too large, the average welfare decreases, thereby confirming our claim in Corollary 1 that the welfare guarantee deteriorates as the accumulated regret of each player's learning algorithm increases. Additionally, we observed that the average welfare increases when more creators are involved, which is not unexpected given that users will have a higher chance of receiving a satisfactory recommendation when there is a larger pool of content on the platform. Furthermore, when the number of players is sufficiently large ($n = 100$), the welfare is fairly good even when players adopt nearly randomized strategies ($\epsilon = 0.9$). We also illustrate the PotA and content distribution induced by creators' sequence of play in Appendix J.3.

6. Conclusion and Future Work

We propose the competing content creation game, a game-theoretical framework for analyzing the strategic behaviors of creators on online content recommendation platforms. Our primary contribution is a comprehensive characterization of social welfare as the outcome of competition among creators, which suggests that the traditional top-K recommendation principle is effective when the platform utilizes user engagement as an incentive metric and offers a sufficient number of choices to users, resonating with the well-known "invisible hand" argument posited by Adam Smith.

Our positive result hinges on two crucial assumptions: 1. users on the platform employ the RU model with Gumbel noise when making decisions; 2. the incentive provided by the platform in creator utilities must align with the user utilities (i.e., user engagement) utlized in the social welfare metric. A fascinating and significant future direction may require the relaxation of either of these assumptions. Consequently, we propose two open questions:

Open Question 1: Does our main result still hold under different user choice models? For instance, what if users are permitted to make multiple choices, or users adhere to the RU model but with a different decision noise (e.g., Gaussian)? Although we have some preliminary empirical evidence suggesting a similar PoA upper bounded strictly less than 2, rigorous theoretical analysis can be challenging, as players' utility functions do not necessarily have closed forms and we might need novel techniques to derive a new smoothness constant.

Open Question 2: What can we say regarding the social welfare when user utility metrics employed in players' incentives and social welfare do not align? Although it is reasonable for the platform to design such aligned incentives, practical reasons might occasionally render this challenging to implement. Although our Proposition 1 demonstrates that the welfare can be arbitrarily bad if player incentives do not correspond with user engagement, it is still intriguing to examine the intermediate situations: what if player incentives and user engagement are correlated but not perfectly aligned? Any understanding of this extended setting can yield more profound impact on real-world systems.

7. Acknowledgement

This work is supported in part by the US National Science Foundation under grants IIS-2007492, IIS-2128019 and IIS-1838615. Haifeng Xu is supported in part by an ARO award W911NF-23-1-0030.

References

Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

Balog, M., Tripuraneni, N., Ghahramani, Z., and Weller, A. Lost relatives of the gumbel trick. In *International Conference on Machine Learning*, pp. 371–379. PMLR, 2017.

Baltas, G. and Doyle, P. Random utility models in marketing research: a survey. *Journal of Business Research*, 51(2): 115–125, 2001.

Ben Basat, R., Tennenholtz, M., and Kurland, O. The probability ranking principle is not optimal in adversarial retrieval settings. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval*, pp. 51–60, 2015.

Ben-Porat, O. and Tennenholtz, M. Shapley facility location games. In *International Conference on Web and Internet Economics*, pp. 58–73. Springer, 2017.

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

Ben-Porat, O., Goren, G., Rosenberg, I., and Tennenholtz, M. From recommendation systems to facility location games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 1772–1779, 2019.

Blum, A., Hajiaghayi, M., Ligett, K., and Roth, A. Regret minimization and the price of total anarchy. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pp. 373–382, 2008.

Bobadilla, J., Ortega, F., Hernando, A., and Gutiérrez, A. Recommender systems survey. *Knowledge-based systems*, 46:109–132, 2013.

Christodoulou, G. and Koutsoupias, E. The price of anarchy of finite congestion games. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pp. 67–73, 2005.

Christodoulou, G., Kovács, A., and Schapira, M. Bayesian combinatorial auctions. In *International Colloquium on Automata, Languages, and Programming*, pp. 820–832. Springer, 2008.

Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

Fan, J. and Cheng, J. Matrix completion by deep matrix factorization. *Neural Networks*, 98:34–41, 2018.

Fleder, D. and Hosanagar, K. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science*, 55(5):697–712, 2009.

Harper, F. M. and Konstan, J. A. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)*, 5(4):1–19, 2015.

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., and Chua, T.-S. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pp. 173–182, 2017.

Hotelling, H. (1929): Stability in competition. *Economic Journal*, 39(4):57, 1929.

Hron, J., Krauth, K., Jordan, M. I., Kilbertus, N., and Dean, S. Modeling content creator incentives on algorithm-curated platforms. *arXiv preprint arXiv:2206.13102*, 2022.

Hurley, N. and Zhang, M. Novelty and diversity in top-n recommendation–analysis and evaluation. *ACM Transactions on Internet Technology (TOIT)*, 10(4):1–30, 2011.

Jagadeesan, M., Garg, N., and Steinhardt, J. Supply-side equilibria in recommender systems. *arXiv preprint arXiv:2206.13489*, 2022.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., and Riedl, J. Grouplens: Applying collaborative filtering to usenet news. *Communications of the ACM*, 40(3):77–87, 1997.

Koren, Y., Bell, R., and Volinsky, C. Matrix factorization techniques for recommender systems. *Computer*, 42(8): 30–37, 2009.

Koutsoupias, E. and Papadimitriou, C. Worst-case equilibria. In *Annual symposium on theoretical aspects of computer science*, pp. 404–413. Springer, 1999.

Ling, C. K., Fang, F., and Kolter, J. Z. What game are we playing? end-to-end learning in normal and extensive form games. *arXiv preprint arXiv:1805.02777*, 2018.

Luce, R. D. Individual choice behavior: A theoretical analysis, new york, ny: John willey and sons, 1959.

Luce, R. D., Suppes, P., et al. Preference, utility, and subjective probability. 1965.

Manski, C. F. The structure of random utility models. *Theory and decision*, 8(3):229, 1977.

McFadden, D. The measurement of urban travel demand. *Journal of public economics*, 3(4):303–328, 1974.

McFadden, D. L. Econometric analysis of qualitative response models. *Handbook of econometrics*, 2:1395–1457, 1984.

McKelvey, R. D. and Palfrey, T. R. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

Meta. Meta is experimenting with new monetization options for creators, 2022. https://www.digitalinformationworld.com/2022/03/meta-is-experimenting-with-new.html.

Monderer, D. and Shapley, L. S. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

Papadimitriou, C. H. and Roughgarden, T. Computing equilibria in multi-player games. In *SODA*, volume 5, pp. 82–91. Citeseer, 2005.

Plackett, R. L. The analysis of permutations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 24(2):193–202, 1975.

Qian, K. and Jain, S. Digital content creation: An analysis of the impact of recommendation systems. *Available at SSRN 4311562*, 2022.

Raifer, N., Raiber, F., Tennenholtz, M., and Kurland, O. Information retrieval meets game theory: The ranking competition between documents' authors. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 465–474, 2017.

Roughgarden, T. Intrinsic robustness of the price of anarchy. *Journal of the ACM (JACM)*, 62(5):1–42, 2015.

Roughgarden, T., Syrgkanis, V., and Tardos, E. The price of anarchy in auctions. *Journal of Artificial Intelligence Research*, 59:59–101, 2017.

Savy. Will the new youtube algorithm impact your content?, 2019. https://savyagency.com/new-youtube-algorithm/.

Shen, W. and Wang, Z. Hotelling-downs model with limited attraction. *arXiv preprint arXiv:1611.05959*, 2016.

TikTok. What is the tiktok creator fund? here's how to join + start making money, 2022. https://www.backstage.com/magazine/article/tiktok-creator-fund-explained-how-to-join-75090/.

Vetta, A. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pp. 416–425. IEEE, 2002.

Youtube. Youtube partner program overview & eligibility, 2023. https://support.google.com/youtube/answer/72851.

Appendix to "How Bad is top-K Recommendation under Competing Content Creators?"

A. Proof of Theorem 1

First we derive the closed-forms of the utility and welfare functions of competing content creation game.

Lemma 3. *Given $\{\varepsilon_i\}$ are drawn i.i.d. from zero-mean Gumbel$(-\beta\gamma, \beta)$, the utility and welfare functions defined in (3), (4) and (5) have the following closed forms*

$$\pi_j(\boldsymbol{s}) = \beta \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big], \tag{19}$$

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \pi_j(\boldsymbol{s}) \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s};K)] \exp(\beta^{-1}\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}, \tag{20}$$

$$W(\boldsymbol{s}) = \beta \sum_{j=1}^{m} \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big]. \tag{21}$$

A.1. Proof of Lemma 3: Closed Forms of Utility and Welfare Functions

We start with a few known and useful properties of Gumbel distributions.

Lemma 4. *[e.g., (Balog et al., 2017)] Let $(v_1, \cdots, v_n) \in \mathbb{R}^n$ be any real-valued vector and $\varepsilon_1, \cdots, \varepsilon_n$ be independent samples from Gumbel(μ, β). Then*

$$\arg\max_i (v_i + \varepsilon_i) \sim \textit{Categorical}\Big(\frac{\exp(\beta^{-1}v_i)}{\sum_{j=1}^{n}\exp(\beta^{-1}v_j)}\Big), \tag{22}$$

and

$$\max_i (v_i + \varepsilon_i) \sim \textit{Gumbel}\Big(\mu + \beta \log\Big(\sum_{j=1}^{n}\exp(\beta^{-1}v_j)\Big), \beta\Big). \tag{23}$$

Derivation of user utility and welfare. These derivations follow easily from Lemma 4. Since we assumed that $\varepsilon_i \sim$Gumbel$(-\beta\gamma, \beta)$, leveraging properties in Lemma 4 we conclude that $\boldsymbol{x}_j$'s choice distribution over K alternatives $\{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K\} = \mathcal{T}_j(\boldsymbol{s};K)$ follows the soft-max rule

$$\mathbf{Pr}[\boldsymbol{x}_j \to \boldsymbol{s}_i] = \frac{\exp(\beta^{-1}\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}, \tag{24}$$

and the expected user utility after making choices has the following form

$$\pi_j(\boldsymbol{x}_j) = \mathbb{E}\left[\max_{i\in[K]}\{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i\}\right] = \beta \log\left[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))\right]. \tag{25}$$

Taking expectation over all users, we obtain the following welfare function

$$W(\boldsymbol{s}) = \sum_{j=1}^{m} \mathbb{E}\left[\max_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)}\{\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j) + \varepsilon_i\}\right] = \beta \sum_{j=1}^{m} \log\left[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\right]. \tag{26}$$

By setting $\tilde{W}(\boldsymbol{s}) = \beta W(\boldsymbol{s}), \tilde{\sigma}(\boldsymbol{s}, \boldsymbol{x}) = \beta^{-1}\sigma(\boldsymbol{s}, \boldsymbol{x})$, we have $\tilde{W}(\boldsymbol{s}) = \sum_{j=1}^{m} \log[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\tilde{\sigma}(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)]$. Therefore, under a rescaling of constant β it is with out loss of generality to consider a scoring function $\sigma \in [0, \frac{1}{\beta}]$, the user utility function and the social welfare function in the following form

$$\pi_j(\boldsymbol{s}) = \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big], \tag{27}$$

$$W(\boldsymbol{s}) = \sum_{j=1}^{m} \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big]. \tag{28}$$

Derivation of creator utility. This turns out to be a new result which requires non-trivial arguments. The players' utility is given by

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \cdot \mathbf{Pr}[\boldsymbol{x}_j \to \boldsymbol{s}_i] \tag{29}$$

$$= \sum_{j=1}^{m} \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \cdot \frac{\exp(\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}, \tag{30}$$

According to the definition in (30), what we need to show is that for i.i.d. random variables $\{\varepsilon_i\}_{i=1}^K$ sampled from Gumbel$(-\beta\gamma, \beta)$,

$$\mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] = \mathbb{E}[\max_{k\in[K]} \{\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j) + \varepsilon_i\}] = \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big], \tag{31}$$

i.e., for any $(v_1, \cdots, v_K) \in \mathbb{R}^K$ and i.i.d. random variables $\{\varepsilon_i\}_{i=1}^K$ sampled from Gumbel$(0, 1)$,

$$\mathbb{E}[v_i + \varepsilon_i | i = \arg\max_{k\in[K]} (v_k + \varepsilon_k)] = \gamma + \log\Big(\sum_{k=1}^{K} \exp(v_k)\Big). \tag{32}$$

Let $Y_i = \max_{k\in[K], k\neq i}(v_k + \varepsilon_k) \sim$ Gumbel$(\log(\sum_{k\neq i} \exp(v_k)), 1)$ and $X_i = v_i + \varepsilon_i \sim$ Gumbel$(v_i, 1)$. Then X_i has the probability density function

$$f_i(x) = \exp(-((x - v_i) + e^{-(x-v_i)})), \tag{33}$$

and Y has the cumulative distribution function

$$F_i(y) = \exp(-e^{-(y-\log(\sum_{k\neq i} \exp(v_k)))}). \tag{34}$$

Therefore we can explicitly compute the conditional expectation of X_i as follows:

$$
\begin{aligned}
&\mathbb{E}[v_i+\varepsilon_i | i = \arg\max_{k\in[K]}(v_k+\varepsilon_k)] \\
=&\mathbb{E}[v_i+\varepsilon_i | v_i+\varepsilon_i \geq \max_{k\in[K],k\neq i}(v_k+\varepsilon_k)] \\
=&\mathbb{E}[X|X\geq Y, X\sim \text{Gumbel}(v_i,1), Y\sim \text{Gumbel}(\log(\sum_{k\neq i}\exp(v_k)),1)] \qquad (35)\\
=&\frac{\int_{\mathbb{R}} x f_i(x)F_i(x)dx}{\int_{\mathbb{R}} f_i(x)F_i(x)dx} \\
=&\frac{\int_{\mathbb{R}} x\exp(-((x-v_i)+e^{-(x-v_i)}))\exp(-e^{-(x-\log(\sum_{k\neq i}\exp(v_k)))})dx}{\int_{\mathbb{R}} \exp(-((x-v_i)+e^{-(x-v_i)}))\exp(-e^{-(x-\log(\sum_{k\neq i}\exp(v_k)))})dx} \\
=&\frac{\int_{\mathbb{R}_{\geq 0}} -\ln t\cdot\exp(-t\sum_{k=1}^K\exp(v_k))dt}{\int_{\mathbb{R}_{\geq 0}}\exp(-t\sum_{k=1}^K\exp(v_k))dt} \qquad (36)\\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)+\frac{\int_{\mathbb{R}_{\geq 0}} -\ln s\cdot\exp(-s)ds}{\int_{\mathbb{R}_{\geq 0}}\exp(-s)ds} \qquad (37)\\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)-\frac{d}{d\alpha}\int_{\mathbb{R}_{\geq 0}} s^\alpha e^{-s}ds \\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)-\frac{d}{d\alpha}\Gamma(\alpha+1)\Big|_{\alpha=0} \\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)+\gamma. \qquad (38)
\end{aligned}
$$

where (35) holds because of Lemma 4, (36) and (37) hold by change of variables $t=e^{-x}$ and $s=t\sum_{k=1}^K\exp(v_k))$, and (38) is from the definition of Euler-Mascheroni constant. Therefore we show (32) and the players' utility function has the following form

$$
u_i(\boldsymbol{s}) = \sum_{j=1}^m\Big(\log\Big[\sum_{\boldsymbol{s}_k\in\mathcal{T}_j(\boldsymbol{s};K)}\exp\left(\sigma(\boldsymbol{s}_k,\boldsymbol{x}_j)\right)\Big]\Big)\frac{\mathbb{I}[\boldsymbol{s}_i\in\mathcal{T}_j(\boldsymbol{s};K)]\exp(\sigma(\boldsymbol{s}_i,\boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k\in\mathcal{T}_j(\boldsymbol{s};K)}\exp(\sigma(\boldsymbol{s}_k,\boldsymbol{x}_j))}. \qquad (39)
$$

A.2. Proof of Lemma 1 and 2: Properties of Utility and Welfare Functions

We consider the utility and welfare functions given in (27), (39) and (28) under the re-scaling of constant β with the new assumption that $\sigma(\boldsymbol{s},\boldsymbol{x})\in[0,\frac{1}{\beta}], \forall \boldsymbol{s}\in\cup_{i=1}^n\mathcal{S}_i, \boldsymbol{x}\in\mathcal{X}$. To simplify the subsequent analysis, we first specify some useful notations and conventions. For any joint strategy profile $\boldsymbol{s}=(\boldsymbol{s}_1,\cdots,\boldsymbol{s}_n)$, we use capital letter S to denote its set representation, i.e., $S=\{\boldsymbol{s}_1,\cdots,\boldsymbol{s}_n\}$. In this way we can view $\mathcal{T}_j(\boldsymbol{s};K),\pi_j(\boldsymbol{s}),u_i(\boldsymbol{s}),W(\boldsymbol{s})$ defined in (1), (3), (4), (5) as set functions $\mathcal{T}_j(S;K),\pi_j(S),u_i(S),W(S)$. From now on, we will use the set notation S and the vector notation $\boldsymbol{s}$ interchangeably, depending on the context. Similarly, we use S_{-i} to denote the set $\{\boldsymbol{s}_1,\cdots,\boldsymbol{s}_n\}$ excluding element $\boldsymbol{s}_i$. Moreover, we extend the definition of $\mathcal{T}_j(S;K)$ by allowing $|S|=K-1$ in the following sense: when $|S|=K-1$, we let $\mathcal{T}_j(S;K)=S\cup\{\bar{\boldsymbol{s}}\}$, where $\bar{\boldsymbol{s}}$ is a default external choice such that $\sigma(\bar{\boldsymbol{s}},\boldsymbol{x})=0$ for all $\boldsymbol{x}\in\mathcal{X}$. This extension captures the situation when the system does not have enough active content creators to allocate to the users. When such a situation happens, the system will put a default choice $\bar{\boldsymbol{s}}$ in the top-K list without any utility guarantee. We remark that this extended definition is introduced merely for the convenience of presentation and does not affect the implication of our main result.

Prior to the proofs for Lemma 1 and 2, we present two intermediate results in Proposition 2 and Lemma 5. Proposition 2 reveals a rather basic property of social welfare W which is useful in the proof of Theorem 1, and Lemma 5 is useful in the proof of Lemma 2.

Proposition 2. *Fix a joint strategy $S=\{\boldsymbol{s}_1,...,\boldsymbol{s}_n\}$ in any $n-$player competing content creation game $\mathcal{G}$. If we add an additional player indexed by $n+1$ with pure strategy $\boldsymbol{s}_{n+1}$ to the game and let $S'=\{\boldsymbol{s}_1,...,\boldsymbol{s}_n,\boldsymbol{s}_{n+1}\}$, the social welfare*

W will strictly increase, i.e.,

$$W(S') > W(S). \tag{40}$$

Proof. By definition,

$$W(S') = \sum_{j=1}^{m} \Big(\log \Big[\sum_{\boldsymbol{s} \in \mathcal{T}_j(S';K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \Big] \Big), \tag{41}$$

$$W(S) = \sum_{j=1}^{m} \Big(\log \Big[\sum_{\boldsymbol{s} \in \mathcal{T}_j(S;K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \Big] \Big), \tag{42}$$

It is obvious that for any fixed user j, the sum of exponential scores of top-K choices from S' is better than that from $\mathcal{S}$, i.e.,

$$\sum_{\boldsymbol{s} \in \mathcal{T}_j(S';K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \geq \sum_{\boldsymbol{s} \in \mathcal{T}_j(S';K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right).$$

Therefore, (40) holds immediately by the monotonicity of the logarithmic function. □

Proposition 2 reveals an important yet natural property of real-world content provider competitions: when there are more competitors in the market, users are facing more alternatives and thus their welfare will always increase.

Lemma 5. *The following function*

$$f(x, y) = \frac{(x+1)\log(x+y)}{(x+y)(\log(x+y) - \log y)}, (x, y) \in \mathbb{R}_+ \times \mathbb{N}_+, \tag{43}$$

is monotonically increasing in y for any $x \in \mathbb{R}_+$, and is monotonically decreasing in x for any integer $y \in \mathbb{N}_+$.

Proof. We first demonstrate the monotonicity of $f(\cdot, y)$ by directly calculating its partial derivatives. Note that $t \geq \log(1+t)$ holds for any $t \geq 0$, we have

$$\frac{1}{x+1} \frac{\partial f(x,y)}{\partial y} = \frac{\log(1+\frac{x}{y}) + \log(x+y)[\frac{x}{y} - \log(1+\frac{x}{y})]}{[(x+y)(\log(x+y) - \log y)]^2} > 0, \tag{44}$$

which implies that $f(x, y)$ is increasing in y. Now it remains to show the monotonicity w.r.t. x when fixing $y = K$, which is slightly more intricate. The derivative of $f(x, K)$ w.r.t. x now writes

$$f'(x, K) = \frac{(K-1)\log(x+K)\log(1+\frac{x}{K}) - (x+1)\log K}{[(x+K)(\log(x+K) - \log K)]^2} \triangleq \frac{-g(x, K)}{[(x+K)(\log(x+K) - \log K)]^2}, \tag{45}$$

and

$$\begin{aligned} g'(x, K) &= \frac{1}{x+K}\Big[(2K+x-1)\log K - 2(K-1)\log(x+K)\Big] && (46) \\ &= \frac{2(K-1)}{x+K}\Big[\frac{2K+x-1}{2(K-1)}\log K - \log(x+K)\Big] \\ &= \frac{2(K-1)}{x+K}\Big[\frac{x+1}{2(K-1)}\log K - \log(1+\frac{x}{K})\Big] && (47) \\ &\geq \frac{2(K-1)}{x+K}\Big[\frac{x+1}{2(K-1)}\log K - \frac{x}{K}\Big] \\ &\geq \frac{x}{x+K}\Big[\log K - \frac{2(K-1)}{K}\Big]. && (48) \end{aligned}$$

We claim $g'(x, K) \geq 0, \forall K \in \mathbb{N}^+$, and this is because

1. if $K = 1$, from (46) we have $g'(x, K) = 0$.

2. for $K \geq 5$, we can verify $\log K - \frac{2(K-1)}{K} > 0$. From (48) we have $g'(x, K) > 0$.
3. for $K \in \{2, 3, 4\}$, we can verify $\frac{x+1}{2(K-1)} \log K - \log(1 + \frac{x}{K}) > 0$ for any $x \geq 0$. Therefore (47) we have $g'(x, K) > 0$.

Now since $g'(x, K) \geq 0$, we conclude that $g(x, K) \geq g(0, K) = \log K \geq 0$, which implies $f'(x, K) \leq 0, \forall K \in \mathbb{N}^+$. Hence, $f(x, K)$ is decreasing in x. □

Now we are ready to prove Lemma 1 and 2.

Proof of Lemma 1. By the definition we only need to show the submodularity of $\pi_j(S)$ for any $j \in [m]$, i.e.,

$$\pi_j(\mathcal{T}_j(S \cup \{\boldsymbol{s}_x\}; K)) - \pi_j(\mathcal{T}_j(S; K)) \geq \pi_j(\mathcal{T}_j(S \cup \{\boldsymbol{s}_x, \boldsymbol{s}_y\}; K)) - \pi_j(\mathcal{T}_j(S \cup \{\boldsymbol{s}_y\}; K)). \tag{49}$$

With out loss of generality we assume $\sigma(\boldsymbol{s}_x, \boldsymbol{x}_j) \geq \sigma(\boldsymbol{s}_y, \boldsymbol{x}_j)$, and let

$$\{v_1, \cdots, v_K\} = \{\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)) | \boldsymbol{s} \in \mathcal{T}_j(S; K)\},$$

where $v_1 \leq \cdots \leq v_K$. Then depending on the values of $v_x = \exp(\sigma(\boldsymbol{s}_x, \boldsymbol{x}_j)), v_y = \exp(\sigma(\boldsymbol{s}_y, \boldsymbol{x}_j))$ and K, there are three situations :

1. $v_x \leq v_1$: (49) holds because its LHS and RHS are both equal to 0.
2. $v_x > v_1, K = 1$: The LHS of (49) is equal to $\log \frac{v_x}{v_1} > 0$, the RHS of (49) is equal to 0.
3. $v_x > v_1, K \geq 2$: The LHS of (49) is equal to $\log \frac{v_x+v_2+a}{v_1+v_2+a}$, the RHS of (49) is equal to $\log \frac{v_x+v_y+a}{v_y+v_2+a}$, where $a = \sum_{k=3}^{K} v_k$ if $K \geq 3$ and $a = 0$ if $K = 2$. We can verify

$$\begin{aligned} &(v_x + v_2 + a)(v_y + v_2 + a) - (v_1 + v_2 + a)(v_x + v_y + a) \\ =&(v_2 - v_1)(a + v_1 + v_2) + (v_x - v_1)(v_y - v_1) \geq 0. \end{aligned}$$

Therefore, (49) holds and Lemma 1 follows by summing (49) over all $j \in [m]$.

□

Proof of Lemma 2. By definition,

$$u_i(\boldsymbol{s}_i; \boldsymbol{s}_{-i}) = \sum_{j=1}^{m} \left(\log \left[\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp\left(\sigma(\boldsymbol{s}', \boldsymbol{x}_j)\right) \right] \right) \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(S; K)] \exp(\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp(\sigma(\boldsymbol{s}', \boldsymbol{x}_j))}, \tag{50}$$

and

$$W(S) = \sum_{j=1}^{m} \pi_j(\mathcal{T}_j(S; K)). \tag{51}$$

It is sufficient to prove that for any user j,

$$\left(\log \left[\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp\left(\sigma(\boldsymbol{s}', \boldsymbol{x}_j)\right) \right] \right) \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(S; K)] \exp(\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp(\sigma(\boldsymbol{s}', \boldsymbol{x}_j))} \geq c(\beta, K) \cdot \left[\pi_j(\mathcal{T}_j(S; K)) - \pi_j(\mathcal{T}_j(S_{-i}; K))\right]. \tag{52}$$

Note that when $\boldsymbol{s}_i \notin \mathcal{T}_j(S; K)$, (52) is trivial as its LHS=RHS=0. Now we suppose $\boldsymbol{s}_i \in \mathcal{T}_j(S; K)$ and thus $\mathcal{T}_j(S_{-i}; K)$ and $\mathcal{T}_j(; K)$ only differ in one element. Without loss of generality we let

$$\{\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)) | \boldsymbol{s} \in \mathcal{T}_j(S_{-i}; K)\} = \{v_1', v_2, \cdots, v_K\},$$

and

$$\{\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)) | \boldsymbol{s} \in \mathcal{T}_j(S; K)\} = \{v_1, v_2, \cdots, v_K\}, v_1 \geq v_1'.$$

Because of our extended definition of $\mathcal{T}_j(S;K)$, $\mathcal{T}_j(S_{-i},K)$ is well defined when $K=n$, under which case we have $v_1' = \exp(\sigma(\bar{\boldsymbol{s}}, \boldsymbol{x}_j)) = 1$. Now we let $z = v_2 + \cdots + v_K$, (52) is equivalent to

$$\frac{v_1}{v_1+z}\log(v_1+z) \geq c(\beta,K)\cdot\log\left[\frac{v_1+z}{v_1'+z}\right]. \tag{53}$$

Since $v_1' = \exp(\sigma(\cdot, \boldsymbol{x}_j)) \geq 1$, a sufficient condition for (53) to hold is

$$\frac{v_1}{v_1+z}\cdot\frac{\log(v_1+z)}{\log(v_1+z)-\log(1+z)} \geq c(\beta,K). \tag{54}$$

Note that $x = v_1 - 1 \in [0, e^{1/\beta}-1], y = z+1 \in [K, (K-1)e^{1/\beta}+1]$, the LHS of (54) becomes a function of (x,y) which has the following form

$$f(x,y) = \frac{(x+1)\log(x+y)}{(x+y)(\log(x+y)-\log y)}. \tag{55}$$

From Lemma 5 we know $f(x,y)$ is monotonically increasing in y for any $x>0$ and is monotonically decreasing in x any integer $K \geq 1$. Therefore, it holds that

$$f(x,y) \geq f(x,K) \geq f(e^{1/\beta}-1,K) = c(\beta,K). \tag{56}$$

Hence, (52) holds and we complete the proof.

□

A.3. Proof of Theorem 1

With the help of Proposition 2, Lemma 2 and 1, now we are ready to prove our claim in Theorem 1. We will demonstrate that any competing content creation game instance $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ is a smooth game with parameter $(\lambda,\mu) = (c(\beta,K), c(\beta,K))$ so that its PoA can be upper bounded by $\frac{1+\mu}{\lambda} = 1 + \frac{1}{c(\beta,K)}$.

Proof. Let $\boldsymbol{s} = (\boldsymbol{s}_1, ..., \boldsymbol{s}_n)$ and $\boldsymbol{s}^* = (\boldsymbol{s}_1^*, ..., \boldsymbol{s}_n^*)$ be two different strategy profiles. First, due to function W's sub-modular property disclosed in Lemma 1, for every $i \in [n]$ we have

$$W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i}) \geq W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}_i^*, \boldsymbol{s}]) - W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}]). \tag{57}$$

Summing over all player i we obtain

$$\begin{aligned}\sum_{i=1}^n (W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i})) &\geq \sum_{i=1}^n (W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}_i^*, \boldsymbol{s}]) - W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}])) \\ &= W([\boldsymbol{s}^*, \boldsymbol{s}]) - W(\boldsymbol{s}) \\ &> W(\boldsymbol{s}^*) - W(\boldsymbol{s}),\end{aligned} \tag{58}$$

where the last inequality holds because of Proposition 2. On the other hand, from Lemma 2 it also holds that

$$u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}) \geq c(\beta,K)\cdot\big[W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i})\big], \tag{59}$$

And therefore

$$\sum_{i=1}^n u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}) \geq c(\beta,K)\cdot\sum_{i=1}^n \big[W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i})\big] \tag{60}$$

$$> c(\beta,K)[W(\boldsymbol{s}^*) - W(\boldsymbol{s})]. \tag{61}$$

where inequality (60) holds by (59), and inequality (61) holds by (58).

Since (61) holds for any $\boldsymbol{s} \in \mathcal{S}$, for any $\boldsymbol{\alpha} \in CCE(\mathcal{G})$ we can take expectation over $\boldsymbol{s} \sim \boldsymbol{\alpha}$ and obtain

$$\sum_{i=1}^{n} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*;\boldsymbol{s}_{-i})] > c(\beta,K)[W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]]. \tag{62}$$

Therefore,

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})] &= \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[\sum_{i=1}^{n} u_i(\boldsymbol{s})] \\
&\geq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[\sum_{i=1}^{n} u_i(\boldsymbol{s}_i^*;\boldsymbol{s}_{-i})] && (63)\\
&\geq c(\beta,K)\cdot\sum_{i=1}^{n}\left[\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W([\boldsymbol{s}_i^*,\boldsymbol{s}_{-i}])] - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s}_{-i})]\right] \\
&> c(\beta,K)[W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]]. && (64)
\end{aligned}$$

where inequality (63) follows by the definition of CCE and inequality (64) holds by (61). Rearranging terms we obtain

$$PoA(\mathcal{G}) = \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]} < 1 + \frac{1}{c(\beta,K)}. \tag{65}$$

□

A.4. Proof of the Property of $c(\beta, K)$

Proof. The $c(\beta, K)$ function has the following form:

$$c(\beta,K) = \frac{(b+1)\log(b+K)}{(b+K)(\log(b+K)-\log K)}, b = e^{\frac{1}{\beta}} - 1. \tag{66}$$

We prove the following facts one by one.

1. Fix any $\beta > 0$, $c(\beta, K)$ is monotonically increasing in K; similarly, fix any $K \geq 1$, $c(\beta, K)$ is monotonically increasing in β.

   Note that $e^{\frac{1}{\beta}} - 1$ is decreasing in β, from Lemma 5 the claim holds.

2. $c(\beta, K) = 1$ if and only if $K = 1$ or $\beta \to 0$.

   When $K = 1$, $c(\beta, K) = 1$ directly holds. When $\beta \to 0$, $b \to +\infty$ and $c(\beta, K) \to 1$. The "only if" direction follows from the monotonicity property of $c(\beta, K)$.

3. For any $\beta > 0$ and $K \geq 1$, we have $c(\beta, K) \geq 1$ and thus $PoA(\mathcal{G}) < 2$ always holds.

   By the monotonicity of c, $c(\beta, K) \geq c(\beta, 1) = 1$. Hence, $PoA(\mathcal{G}) < 1 + \frac{1}{c(\beta,K)} \leq 2$.

4. For sufficiently large β and K, $c(\beta, K) \approx (1 + \beta)\log K$ asymptotically, and therefore

$$PoA(\mathcal{G}) < 1 + \frac{1}{(1+\beta)\log K}. \tag{67}$$

   When β is sufficiently large, $b = e^{\frac{1}{\beta}} - 1 \approx \frac{1}{\beta} \to 0$. Therefore,

$$\begin{aligned}
c(\beta,K) &= \frac{(b+1)\log(b+K)}{(b+K)\log(1+\frac{b}{K})} \\
&\approx \frac{(b+1)K\log(b+K)}{(b+K)b} && \text{since } \log(1+x)\approx x \text{ as } x\to 0 \\
&\approx \frac{(b+1)\log K}{b} && \text{since } K >> b \\
&\approx (1+\beta)\log K. && \text{since } b \approx \tfrac{1}{\beta}
\end{aligned}$$

□

B. Proof of Corollary 1

Proof. Let $\epsilon(T) = \frac{R(T)}{T}$, and $\boldsymbol{s}^* = (\boldsymbol{s}_1^*, ..., \boldsymbol{s}_n^*)$ be a global maximizer of $W(\boldsymbol{s})$. By definition,

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s})] \geq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i})] - \epsilon(T). \tag{68}$$

Summing over all player $i \in [n]$ we obtain

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})] = \sum_{i=1}^{n} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s})] \geq \sum_{i=1}^{n} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i})] - n\epsilon(T). \tag{69}$$

On the other hand, by (61) from the proof of Theorem 1, we have

$$\sum_{i=1}^{n} u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}) > c(\beta, K)[W(\boldsymbol{s}^*) - W(\boldsymbol{s})], \forall \boldsymbol{s} \in S. \tag{70}$$

Taking the expectation of $\boldsymbol{s}$ over distribution $\boldsymbol{\alpha}$ we obtain

$$\sum_{i=1}^{n} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i})] > c(\beta, K)\big(W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]\big). \tag{71}$$

(69) and (71) together imply that

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})] + n\epsilon(T) > c(\beta, K)\big(W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]\big). \tag{72}$$

Note that for any $\boldsymbol{s} \in \mathcal{S}$, we have $W(\boldsymbol{s}) = \sum_{j=1}^{m} \left(\log \left[\sum_{\boldsymbol{s}\in\mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \right] \right) \geq \beta \log K$ and therefore, $n\epsilon(T) \leq \frac{n\epsilon(T)}{\beta \log K} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$. Substituting it into (72), we obtain (13).

□

C. Proof of Corollary 2

Proof. Note that fix any players' strategy profile $\boldsymbol{s}$, the top-K matching mechanism maximizes the social welfare W. Therefore, $W(\mathcal{G}') \leq W_K^*$. On the other hand, from the PoA bound in Theorem 1 it holds that

$$\begin{aligned}
\frac{W_K^*}{W(\mathcal{G})} &< 1 + \frac{(b+K)(\log(1+b/K))}{(b+1)\log(b+K)} \\
&< 1 + \frac{(b+K)(b/K)}{(b+1)\log(b+K)} \\
&= 1 + \frac{b+K}{K\log(b+K)}.
\end{aligned}$$

Rearranging term yields $W(\mathcal{G}') - W(\mathcal{G}) \leq W_K^* - W(\mathcal{G}) \leq W_K^*/(1 + \frac{K\log(b+K)}{b+K})$. □

D. Proof of Theorem 2

Proof. Let $b = \exp(1/\beta) - 1$. Consider an n-player game where each player-i has the same action set $\mathcal{S}_i = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$. Let the user population $\mathcal{X}$ be a set with size $m = n + (n-1)a$, in which n users have profile $\boldsymbol{x}_1$ and a users have profile $\boldsymbol{x}_i$ for $i = 2, \cdots, n$. Here $a = \beta \log K + 1$ is a constant whose choice will become clear later. Let the scoring function σ be the indicator function defined as follow:

$$\sigma(\boldsymbol{s}, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{s} = \boldsymbol{x}, \\ 0, & \text{otherwise.} \end{cases} \tag{73}$$

First we lower bound the optimal welfare $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$. Consider the joint-strategy profile $\boldsymbol{s}^* = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)$, under which each user gets one player with σ score 1 and $K-1$ player with σ score 0. In this case, each user-j's utility $\pi_j(\boldsymbol{s}) = \beta\log(b+K)$ and the social welfare $W(\boldsymbol{s}) = m\beta\log(b+K)$. Therefore, the optimal social welfare

$$\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s}) \geq W(\boldsymbol{s}^*) = m\beta\log(b+K). \tag{74}$$

Next we show that $\boldsymbol{s} = (\boldsymbol{x}_1, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_1)$ is a pure NE of $\mathcal{G}$ and thus $\boldsymbol{s} \in CCE(\mathcal{G})$. Given players' joint-strategy $\boldsymbol{s}$, n users will be assigned with K players with σ score 1 and $(n-1)a$ users will be assigned with K players with σ score 0. Therefore, the utility for an arbitrary player-i is given by

$$\begin{aligned} u_i(\boldsymbol{s}) &= \Big[n\cdot(\beta\log K + 1) + a(n-1)\cdot\beta\log K\Big]/n \\ &= \beta\log K + 1 + \frac{a(n-1)\beta\log K}{n}. \end{aligned} \tag{75}$$

If player-i switches from strategy $\boldsymbol{s}_1$ to $\boldsymbol{s}_j$, n users still get K players with score 1, $(n-2)a$ users get players with score 0, and a users get K players with scores $(1, 0, \cdots, 0)$. Therefore, player-i's utility after the deviation is

$$\begin{aligned} u_i(\boldsymbol{s}_j, \boldsymbol{s}_{-i}) &= n\cdot 0 + a(n-2)\cdot\beta\log K\cdot\frac{1}{n} + a\cdot\beta\log(b+K)\cdot\frac{e^{\frac{1}{\beta}}}{e^{\frac{1}{\beta}} + K - 1} \\ &= \frac{a(n-2)\beta\log K}{n} + \frac{b+1}{b+K}\cdot a\beta\log(b+K). \end{aligned}$$

We can verify that $u_i(\boldsymbol{s}) \geq u_i(\boldsymbol{s}_j, \boldsymbol{s}_{-i})$ for any $2 \leq j \leq n$ if we take

$$a = \beta\log K + 1 \leq \frac{\beta\log K + 1}{\beta\Big(\frac{b+1}{b+K}\log(b+K) - \frac{1}{n}\log K\Big)}. \tag{76}$$

This is because: 1. The inequality in (76) always holds as $\frac{b+1}{b+K}\log(b+K) < \log(b+1) = \frac{1}{\beta}$ when $\beta \in [0,1]$ (this is due to the monotonicity of $\log x/x$); 2. $u_i(\boldsymbol{s}) = u_i(\boldsymbol{s}_j, \boldsymbol{s}_{-i})$ when $a = \frac{\beta\log K+1}{\beta\Big(\frac{b+1}{b+K}\log(b+K) - \frac{1}{n}\log K\Big)}$. Hence, $\boldsymbol{s}$ is an NE of $\mathcal{G}$.

Putting (74), (75), and (76) together, we have

$$\begin{aligned} PoA(\mathcal{G}) &= \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]} \geq \frac{W(\boldsymbol{s}^*)}{W(\boldsymbol{s})} \\ &= \frac{m\beta\log(b+K)}{nu_i(\boldsymbol{s})} \geq \frac{m}{nu_i(\boldsymbol{s})} \qquad (77) \\ &= \frac{n + (n-1)a}{n(\beta\log K + 1) + a(n-1)\beta\log K} \\ &= \frac{n-1}{n} + \frac{1 - t^2 a(a-1)}{a + ta(a-1)} \quad (t = \frac{n-1}{n}) \\ &> \frac{n-1}{n} + \frac{1}{5a-4} \qquad (78) \\ &= \frac{n-1}{n} + \frac{1}{1 + 5\beta\log K}. \end{aligned}$$

where inequality (77) holds because $\beta\log(b+K) \geq \beta\log(b+1) = 1$, and (78) holds because when $a = 1 + \beta\log K \in [1, 1.2]$ and $t = \frac{n-1}{n} \in [0.5, 1)$, it is easy to verify that $\frac{1-t^2a(a-1)}{a+ta(a-1)} > \frac{1}{5a-4}$. It is equivalent to $4t^2a + 4 > 5t^2a^2 + ta$, which is true because $t^2a(5a-4) + ta < a(5a-4) + a < 4$. □

E. Proof of Proposition 1

Proof. Consider a population of two users $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ and n players in which one player has two pure strategies and the other $n-1$ players only have access to a single strategy, i.e, $\mathcal{S}_i = \{\boldsymbol{s}_0\}, i = 2, \dots, n$ and $\mathcal{S}_1 = \{\boldsymbol{s}_1, \boldsymbol{s}_2\}$. Let the scoring function σ be

$$\sigma(\boldsymbol{s}, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{s} = \boldsymbol{s}_1, \boldsymbol{x} = \boldsymbol{x}_1, \\ \delta, & \text{if } \boldsymbol{s} = \boldsymbol{s}_2, \\ 0, & \text{otherwise.} \end{cases} \tag{79}$$

We will show that for any given $K \geq 1, 0 \leq \beta \leq \min\{0.14, \frac{1}{5\log K}\}$ there exists $\delta \in (0,1)$ such that the PoA of game $\tilde{\mathcal{G}}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ is always strictly greater than 2.

From the proof of Lemma 3, the user utility and welfare functions of $\tilde{\mathcal{G}}$ share the same form as in (19), (21), while the player utility functions of $\tilde{\mathcal{G}}$ have the following form

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s}; K)] \exp(\beta^{-1}\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s}; K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}. \tag{80}$$

Let $b = \exp(1/\beta) - 1$ and we choose any $\delta \in [\delta_0, 1)$ such that $\exp(\delta_0/\beta) + K - 1 = \frac{2}{\frac{1}{K} + \frac{1}{b+K}}$. Such $\delta_0 \in (0,1)$ must exist because function $f(\delta) = \exp(\delta/\beta) + K - 1$ is monotonically increasing in $[0,1]$ with range $[K, b+K] \supset \frac{2}{\frac{1}{K} + \frac{1}{b+K}}$. Given such choice of δ, we can verify that

$$\begin{aligned} 2u_1(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0) &= \frac{2\exp(\delta/\beta)}{\exp(\delta/\beta) + K - 1} \\ &\geq \frac{2\exp(\delta_0/\beta)}{\exp(\delta_0/\beta) + K - 1} \\ &= \frac{\exp(1/\beta)}{\exp(1/\beta) + K - 1} + \frac{1}{K} \\ &\geq \frac{\exp(1/\beta)}{\exp(1/\beta) + K - 1} + \frac{1}{n} = 2u_1(\boldsymbol{s}_1, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0), \end{aligned}$$

which indicates that $(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)$ is a PNE of $\tilde{\mathcal{G}}$. Therefore, by picking $\delta = \delta_0$ we have

$$\begin{aligned} PoA(\tilde{\mathcal{G}}) &= \frac{\max_{\boldsymbol{s} \in \mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha} \in CCE(\tilde{\mathcal{G}})} \mathbb{E}_{\boldsymbol{s} \sim \boldsymbol{\alpha}}[W(\boldsymbol{s})]} \geq \frac{W(\boldsymbol{s}_1, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)}{W(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)} \\ &= \frac{\log(\exp(1/\beta) + K - 1) + \log K}{2\log(\exp(\delta_0/\beta) + K - 1)} && (81) \\ &= \frac{\log(b+K) + \log K}{2\log[2K(b+K)] - 2\log(b+2K)} && \text{by the choice of } \delta_0 \quad (82) \\ &> 2, && (83) \end{aligned}$$

where (83) holds because (83) is equivalent to

$$(b+2K)^4 > 16K^3(b+K)^3. \tag{84}$$

And we show the correctness of (84) by verifying the following situations:

1. when $K \in \{2, 3\}$, (84) holds for all $\beta \in [0, 0.14]$, $b = \exp(1/\beta) - 1$.
2. when $K \geq 4$, from $\beta \leq \frac{1}{5\log K}$ we know $b + K = \exp(1/\beta) + K - 1 > K^5$ and thus $\frac{(b+2K)^4}{(b+K)^3} > b > K^5 \geq 16K^3$. Therefore, (84) holds.

Finally we show that when $K = 1$ or $\beta \to 0$, $PoA(\tilde{\mathcal{G}})$ can be arbitrarily large.

1. when $\beta \to 0$, we have $b \to \infty$. From (82) we have for any fixed K,

$$\lim_{\beta \to 0} PoA(\tilde{\mathcal{G}}) = \lim_{b \to +\infty} \left\{ \frac{\log(b+K) + \log K}{2\log[2K(b+K)] - 2\log(b+2K)} \right\} = \lim_{b \to +\infty} \log(b+K) \to +\infty.$$

2. when $K = 1$, the user's choice is deterministic and thus any $\delta \in (0,1)$ makes $(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)$ a PNE of $\tilde{\mathcal{G}}$. Let $\delta \to 0$ and from (81) we have for any fixed β,

$$\lim_{\delta \to 0} PoA(\tilde{\mathcal{G}}) = \lim_{\delta \to 0} \left\{ \frac{\log(\exp(1/\beta) + K - 1) + \log K}{2\log(\exp(\delta/\beta) + K - 1)} \right\} = \lim_{\delta \to 0} \frac{1}{2\delta} \to +\infty.$$

□

F. Additional Discission on Related Work

Our user decision model (see Section 3) stems from the RU model (Baltas & Doyle, 2001) in econometrics, which explains how an individual makes choices among a discrete set of alternatives. In the RU model, the utility that a decision maker could obtain from alternative j is decomposed into $U_j = V_j + \epsilon_j$, where V_j is the known parameterized part, and ϵ_j is the unknown stochastic part. The observed choice is then given by the alternative with the maximum utility. It is shown that if the unobserved stochastic utility follows the extreme value distribution (i.e., Gumbel distribution), then the choice probability is given by the logit formula, i.e., $P_j \propto \exp(V_j)$ (Luce et al., 1965). In our work, we apply the RU model to explain how a typical user allocates her attention across the recommended list.

To analyze the equilibrium efficiency of the competing content creation game, we employed the standard framework of the price of anarchy (PoA). This originates from the seminal work of (Koutsoupias & Papadimitriou, 1999) and has since led to an extensive literature on understanding the efficiency of numerous strategic games. Our discussion by no means can do justice to this rich literature; here, we only mention the few works that are closely related to ours. Since Nash equilibrium (NE) is not guaranteed to exist in our problem with non-continuous agent utilities (Hron et al., 2022), it is thus crucial for us to consider a solution concept that is weaker than NE and thus to prove a stronger PoA bound. Specifically, we consider coarse correlated equilibrium (CCE). The PoA for CCE is first studied by (Blum et al., 2008), who considered the efficiency of a dynamic setup with no-regret learners and coined the new notion of the price of total anarchy, which turns out to be equivalent to the PoA bound for CCE. This is precisely the question we want to address, but the structure of our new competing content creation game is significantly different from the games they studied, such as Hotelling's game on a graph and the valid utility game of (Vetta, 2002). Thus their techniques are not readily applicable to our problem. We instead employed a recent framework of (Roughgarden, 2015) using the smoothness argument. It is well-known that this framework can yield strong PoA bound applicable to CCE. However, the bounds obtained by this powerful framework are usually not tight; so far, it is only known that it yields tight PoA bounds for linear cost congestion games (Christodoulou & Koutsoupias, 2005), second price auctions (Christodoulou et al., 2008), and the valid utility games (Vetta, 2002). Interestingly, We show that the smoothness argument also yields a tight PoA bound for our competing content creation game and thus register an additional member to this important list of games.

G. Additional Discussion on Stability

Our characterization of social welfare does not require any stability property of the creators' competition. While previous works (Ben-Porat & Tennenholtz, 2018; 2017; Ben-Porat et al., 2019) strive to establish a unique pure Nash equilibrium (PNE) guarantee in similar game-theoretic settings, we do not consider such a stability requirement crucial for the system's design for two reasons. First, as demonstrated in previous works (Ben-Porat & Tennenholtz, 2018; 2017; Ben-Porat et al., 2019), the unique PNE is not guaranteed for top-K RS in general. However, our main result indicates the social welfare can be ensured under top-K RS regardless of the existence of a PNE, thus eliminating the need for tradeoffs between stability and the complexity of the recommendation algorithm. Second, even in cases where a PNE does exist, it is unclear how creators can achieve such a stable outcome in practice. For instance, Ben-Porat & Tennenholtz (2018; 2017) showed that creators following best-response dynamic (Monderer & Shapley, 1996) converge to the unique PNE in their game-theoretical setting, but it requires creators to have oracle access to their utility functions, which is unrealistic as creators can only evaluate the utility of their taken actions (i.e., bandit feedback) in practice. Therefore, we refrain from discussing stability and instead focus on characterizing the average social welfare under the evolving strategies of creators.

H. Connections to Existing Models

As an extended discussion to the related work, we show how our competing content creation games connect to the following three previously proposed competition models for content creators. All the following models do not consider the presence of an RS and match each user with all content creators (players), which corresponds to the case $K = n$ in our setting. Interestingly, we found that each of them turns out to be a special case of our competing content creation game.

Facility location games under the no intervention mediator (Ben-Porat et al., 2019)

Consider the following competing content creation game instance:

1. the user population $\mathcal{X} \subseteq [0, 1]$ is a finite set of size m,
2. each player $i \in [n]$ shares the same action set $\mathcal{S}_i = [0, 1]$,
3. the scoring function is given by $\sigma(s, x) = |s - x|$,
4. $(\beta, K) = (0, n)$,
5. utility function is the user exposure metric, i.e., $u_i(\boldsymbol{s}) = \sum_{x\in\mathcal{X}} \mathbf{Pr}(x \to s_i)$.

If we let $m \to \infty$ so that $\mathcal{X}$ becomes a continuum with density function g over the unit interval $[0, 1]$, the game instance $\tilde{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ [6] defined above is equivalent to the facility location game under the no intervention mediator proposed by (Ben-Porat et al., 2019).

Hotelling-Downs model with limited attraction under support utility functions (Shen & Wang, 2016)

Consider the following competing content creation game instance:

1. the user population $\mathcal{X} = \{x_1, \cdots, x_m\} \subseteq [0, 1]$ is a finite set of size m,
2. each player $i \in [n]$ shares the same action set $\mathcal{S}_i = [0, 1] \times [0, 1]$. For each action $\boldsymbol{s}_i = (s_i, w_i)$ taken by player-i, it is associated with an attraction region $R_i = [s_i - \frac{w_i}{2}, s_i + \frac{w_i}{2}] \cap [0, 1]$.
3. for each $i \in [n]$, the scoring function is given by $\sigma(\boldsymbol{s}_i, x) = \mathbb{I}[x \in R_i]$,
4. $(\beta, K) = (0, n)$,
5. the utility function is induced by the user engagement metric, i.e., $u_i(\boldsymbol{s}) = \sum_{j=1}^m \pi_j(\boldsymbol{s}) \mathbf{Pr}(x_j \to s_i)$.

In fact, given $\beta = 0$ and the above definition of σ, we can see the utility functions under both exposure and engagement metrics are identical, because it holds that $\pi_j(\boldsymbol{s}) \in \{0, 1\}$ and $\pi_j(\boldsymbol{s}) = 1$ if and only if $\mathbf{Pr}(x_j \to s_i) > 0$. We can verify that the game instance $G(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ defined above is equivalent to the Hotelling-Downs model with limited attraction under support utility functions proposed by (Shen & Wang, 2016).

Exposure games (Hron et al., 2022)

Consider the following competing content creation game instance:

1. the user population $\mathcal{X} \subseteq \mathbb{R}^d$ is a finite set of size m,
2. each player $i \in [n]$ is associated with an action set $\mathcal{S}_i$ on the unit sphere in $\mathbb{R}^d$, i.e., $\mathcal{S}_i \in \mathbb{S}^{d-1}$,
3. the scoring function is given by the inner product, i.e., $\sigma(\boldsymbol{s}, \boldsymbol{x}) = \langle \boldsymbol{s}, \boldsymbol{x} \rangle$,
4. $(\beta, K) = (\tau, n)$,
5. the utility function is induced by the user exposure metric, i.e., $u_i(\boldsymbol{s}) = \sum_{x\in\mathcal{X}} \mathbf{Pr}(x \to s_i)$.

Note that in exposure games the parameter β no longer represents the user decision noise but becomes a temperature parameter τ controlling the spread of exposure probabilities over items. The game instance $\tilde{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ defined above is equivalent to the exposure games proposed by (Hron et al., 2022).

[6] Note that we use $\tilde{G}$ to refer to the variant of G that utilizes the user exposure metric instead of the user engagement metric in player utility functions.

I. The Detailed Experimental Setup

I.1. The computation of globally optimal social welfare

We use two heuristic methods for the computation of $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$, when the exact solution is computationally infeasible: Simulated Annealing (SA) shown in Algorithm 1 and Best-response Search (BRS) shown in Algorithm 2. For SA, we set $T = 5000$ and the temperature schedule $\tau_t = 0.1/\sqrt{t}$; for BRS, we set $T = \max\{30, 2n\}$ and take the best output from 5 independent runs. Both methods perform as well as the brute-force search when the number of players n and the size of action set k are less than 6, thus yielding the exact global optimal under such situations. For larger problem scales, we run both methods and select the best output to approximate $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$.

Algorithm 1 Simulated Annealing for Computing the Globally Optimal Welfare

Input: Time horizon T, joint action space $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$, welfare function $W(\boldsymbol{s})$, temperature schedule $\{\tau_t\}_{t=1}^T$.
Initialization: A randomly selected joint action $\boldsymbol{s}^{(0)} = (\boldsymbol{s}_1^{(0)}, \cdots, \boldsymbol{s}_n^{(0)}) \in \mathcal{S}$.
for $t = 0$ **to** T **do**
  Randomly choose a player $i \in [n]$ and randomly perturb her action $\boldsymbol{s}_i^{(t)}$ in $\boldsymbol{s}^{(t)}$ to yield $\boldsymbol{s}'^{(t)}$.
  Compute $W(\boldsymbol{s}'^{(t)}), W(\boldsymbol{s}^{(t)})$.
  if $W(\boldsymbol{s}'^{(t)}) > W(\boldsymbol{s}^{(t)})$ **then**
    Set $\boldsymbol{s}^{(t+1)} = \boldsymbol{s}'^{(t)}$.
  else
    With probability $e^{(W(\boldsymbol{s}'^{(t)})-W(\boldsymbol{s}^{(t)}))/\tau_t}$, set $\boldsymbol{s}^{(t+1)} = \boldsymbol{s}'^{(t)}$; otherwise, $\boldsymbol{s}^{(t+1)} = \boldsymbol{s}^{(t)}$.
  end if
end for
Output: $\max_{t\in[T]} W(\boldsymbol{s}^{(t)})$.

Algorithm 2 Best-response Search for the Globally Optimal Welfare

Input: Time horizon T, joint action space $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$, welfare function $W(\boldsymbol{s})$.
Initialization: A randomly selected joint action $\boldsymbol{s}^{(0)} = (\boldsymbol{s}_1^{(0)}, \cdots, \boldsymbol{s}_n^{(0)}) \in \mathcal{S}$.
for $t = 0$ **to** T **do**
  Randomly choose a player $i \in [n]$ and search for her best-response that maximizes W, i.e.,

$$\boldsymbol{s}_i^{(t+1)} = \arg\max_{\boldsymbol{s}_i\in\mathcal{S}_i} W(\boldsymbol{s}_i, \boldsymbol{s}_{-i}^{(t)}).$$

  Set $\boldsymbol{s}^{(t+1)} = (\boldsymbol{s}_i^{(t+1)}, \boldsymbol{s}_{-i}^{(t)})$.
end for
Output: $W(\boldsymbol{s}^{(T)})$.

I.2. The computation of the worst case welfare under CCE

We can express the definition of CCE into a set of linear constraints. Let $\boldsymbol{\alpha} \in$CCE$(\mathcal{G})$ be a probability distribution over the joint action space $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$. Then $\boldsymbol{\alpha}$ satisfies

$$\forall i, \forall \boldsymbol{s}_i', \sum_{\boldsymbol{s}\in\mathcal{S}} \boldsymbol{\alpha}(\boldsymbol{s}) u_i(\boldsymbol{s}) \geq \sum_{\boldsymbol{s}\in\mathcal{S}} \boldsymbol{\alpha}(\boldsymbol{s}) u_i(\boldsymbol{s}_i', \boldsymbol{s}_{-i}). \tag{85}$$

And $\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$ is solving

$$\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \boldsymbol{\alpha}(\boldsymbol{s}) W(\boldsymbol{s}) \tag{86}$$

under linear constraints (85). Suppose each $\mathcal{S}_i$ shares the same size k, then we obtain a linear program with k^n variables and kn constraints.

I.3. The details of no-regret dynamic simulation and computation of PotA

For the computation of PotA, we need to simulate each player's sequence of play. We let all players run the following Exp-3 algorithm 3 simultaneously to update their strategies in a fixed time horizon $T = 5000$. According to (Auer et al.,

Table 3. PoA under $\beta = 0.5$. Results reflect the worst cases obtained from 10 independently sampled game instances.

K \ n	$*$	2	3	4	5
1	2.00	1.33	1.54	1.66	1.72
2	1.77	1.11	1.24	1.32	1.34
3	1.65		1.08	1.13	1.18
4	1.57			1.05	1.08
5	1.52				1.02

$*$ denotes the theoretical upper bound.

Table 4. PotA under $\beta = 0.5$. Results reflect the worst cases obtained from 10 independently sampled game instances.

K \ n	$*$	5	10	15	20	40
1	2.00	1.59	1.59	1.60	1.50	1.38
3	1.65	1.13	1.20	1.21	1.22	1.20
5	1.52	1.03	1.10	1.12	1.14	1.14
7	1.45		1.05	1.08	1.09	1.11

$*$ denotes the theoretical upper bound.

2002), Algorithm 3 enjoys sublinear regret if $\eta = \epsilon \sim O(\sqrt{\frac{k \log k}{T}})$. However, it is not realistic to assume content creators in practice are sophisticated enough to figure out the game parameters and the optimal learning/exploration rates. Hence, unless specified, we always use a fixed value $(\eta, \epsilon) = (0.1, 0.1)$ in our experiments.

Algorithm 3 Exp3 for player-i

Input: Time horizon T, number of actions k, exploration parameter ϵ, learning rate η.
Initialization: The score vector $\boldsymbol{y}_0 = (y_1(0), \cdots, y_k(0)) = (0, \cdots, 0)$.
for $t = 0$ **to** T **do**
Compute a mixed strategy from the accumulated scores:

$$p_j(t) = (1-\epsilon)\frac{\exp(y_j(t))}{\sum_{l\in[k]}\exp(y_l(t))} + \frac{\epsilon}{k}, \quad \forall j \in [k].$$

Draw action $\boldsymbol{s}_{i,t} \in [k]$ randomly accordingly to the distribution $\boldsymbol{p}_t = (p_1(t), p_2(t), \dots p_k(t))$.
Play action $\boldsymbol{s}_{i,t}$ and observe the utility $u_i(\boldsymbol{s}_{i,t}, \boldsymbol{s}_{-i,t})$.
Update the score

$$y_{\boldsymbol{s}_{i,t}}(t+1) = y_{\boldsymbol{s}_{i,t}}(t) + \eta \cdot \frac{u_i(\boldsymbol{s}_{i,t}, \boldsymbol{s}_{-i,t})}{p_{\boldsymbol{s}_{i,t}}(t)}.$$

end for

J. Additional Experiment Result

J.1. Empirical PoA from simulations

In Table 3 and 4 we summarize the PoA and PotA under $\beta = 0.5$ for game instances constructed from dataset-1. Keeping n fixed, the trend with respect to K is consistent with the results observed for $\beta = 0.1$. When compared to the results in Tables 1 and 2, it is observed that both PoA and PotA values are smaller for any fixed $K \geq 2$, thereby confirming the theoretical claim made in Theorem 1 that the PoA monotonically decreases with respect to β.

J.2. Comparison between user engagement/exposure metrics

The results presented in Figure 3 further support the superiority of utilizing the user engagement metric across a range of environments (β, K). It is also observed that the advantage of the engagement metric over the exposure metric diminishes as either β or K increases. This might suggest that the PotA under the exposure metric also decreases with respect to K and β, although this claim remains unproven. Nonetheless, it constitutes an interesting question for further investigation.

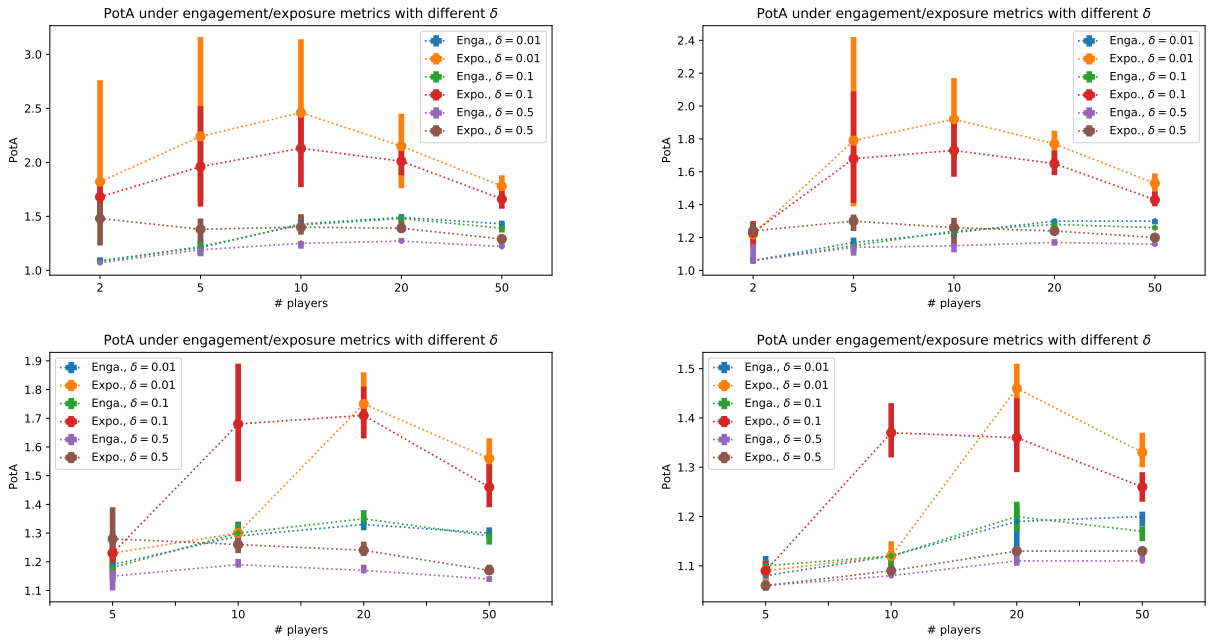


Figure 3. PotA under user exposure and user engagement metrics under different (β, K). δ is the relevance score obtained from creators' "safe" action. Upper-left: $(0.1, 1)$, Upper-right: $(0.3, 2)$, Lower-left: $(0.1, 5)$, Lower-right: $(0.3, 5)$. The error bars indicate the largest/smallest values from 10 independent trials and the dots correspond to the mean values.

Table 5. PotA under $(\beta, K) = (0.1, 5)$ for the MovieLens dataset. Results are averaged over 10 independent simulations under different exploration factor ϵ.

n \ ϵ	.003	.01	.03	.1	.3	.9
5	1.51	1.32	1.18	1.10	1.12	1.51
10	1.15	1.14	1.08	1.03	1.05	1.25
30	1.06	1.05	1.04	1.02	1.03	1.07
100	1.02	1.02	1.02	1.02	1.02	1.03

J.3. Additional Results on MovieLens-1m

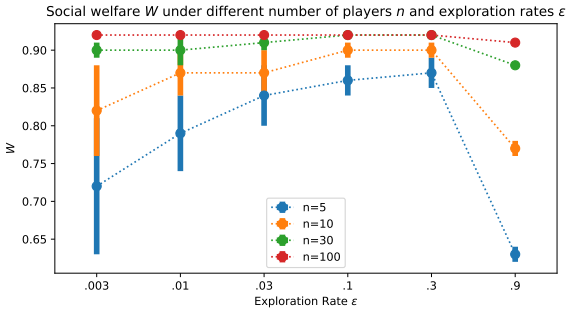
The social welfare under different levels of rationality

Together with Figure 2, plots in Figure 4 suggest a consistent optimal exploration rate $\epsilon = 0.1$ for $K = 5$. However, the optimal exploration rate for $K = 1$ is observed to be around $\epsilon = 0.3$, as demonstrated in the left panel of Figure 4. One potential explanation for this discrepancy is that when the RS provides a greater number of alternatives in the recommendation list, the randomness in users' decisions can better compensate for creators' exploration in finding more diversified strategies, leading to improved social welfare.

The PotA for MovieLens dataset

Under the same setting used in Figure 2 and 4, where we computed the average welfare for different numbers of players with various exploration rates, we also present the corresponding PotA in Table 5 and 6. As can be observed, when creators are able to utilize the optimal exploration rates that minimize regret, social welfare is unexpectedly high. Even with as few as five creators, the fraction of welfare loss is only around 10% (PotA=1.10). As the number of creators increases, the PotA rapidly approaches the ideal value of 1.0. Although the PotA may be underestimated due to approximation accuracy in determining the optimal welfare, such results still convey a very optimistic message.

The genre distribution resulting from the sequence of play



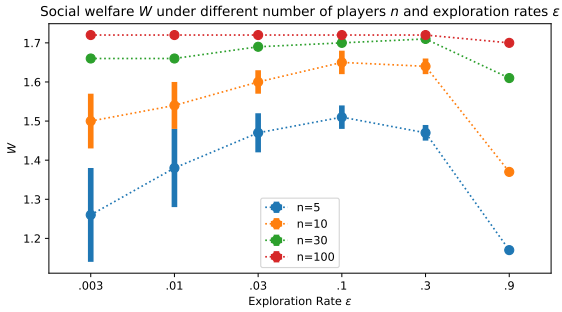


Figure 4. The averaged welfare $\bar{W}$ over $T = 1000$ rounds under different exploration rate ϵ and number of players n. Results are averaged over 10 independent simulations under $(\beta, K) \in \{(0.5, 1), (0.5, 5)\}$.

Table 6. PotA under $(\beta, K) = (0.5, 5)$ for the MovieLens dataset. Results are averaged over 10 independent simulations under different exploration factor ϵ.

n \ ϵ	.003	.01	.03	.1	.3	.9
5	1.37	1.25	1.18	1.15	1.18	1.48
10	1.15	1.12	1.08	1.05	1.06	1.26
30	1.04	1.04	1.02	1.02	1.01	1.07
100	1.01	1.01	1.01	1.01	1.01	1.02

In addition to evaluating the welfare metric, we also examined the content distribution induced by creators' sequence of play under different exploration rates, as shown in Figure 5. The content distribution is generated by counting the frequency of genres (in MovieLens-1m dataset, each movie is associated with multiple genre tags) throughout a simulation trajectory. Specifically, for each simulation path with a fixed time horizon $T = 1000$, we collect the complete set of content that all creators have produced and summarize their frequencies w.r.t. genres. We also compute the genre distributions for the optimal joint strategy that maximizes the social welfare under different n as references.

The leftmost panel in Figure 5 illustrates the genre distribution of the globally optimal content creation strategies under varying numbers of creators n. As n increases, the optimal genre distribution becomes "flatter", which is consistent with the expectation that new creators who want to maximize their utility should focus on niche interest groups to diversify the overall content distribution. The empirical genre distributions from simulations, shown in the three panels on the right side, are observed to be more polarized in comparison to the optimal distributions. For example, in all cases, creators generate content from the drama genre (the most popular genre) more than the ideal frequency. Such polarization effect is alleviated as the exploration factor on the creator side increases. Interestingly, if we look at the empirical frequency under different n for a fixed ϵ, we can find that the content distribution also becomes increasingly diversified as n gets larger, which is beneficial for social welfare. This explains our finding in Table 5 and 6 as well: that PotA gets consistently smaller under various ϵ as the number of creators increases. These observations provide insight into the positive impact of creators adopting regret minimization strategies: the resulting social welfare is fairly good, even if they are subject to bounded rationalities.

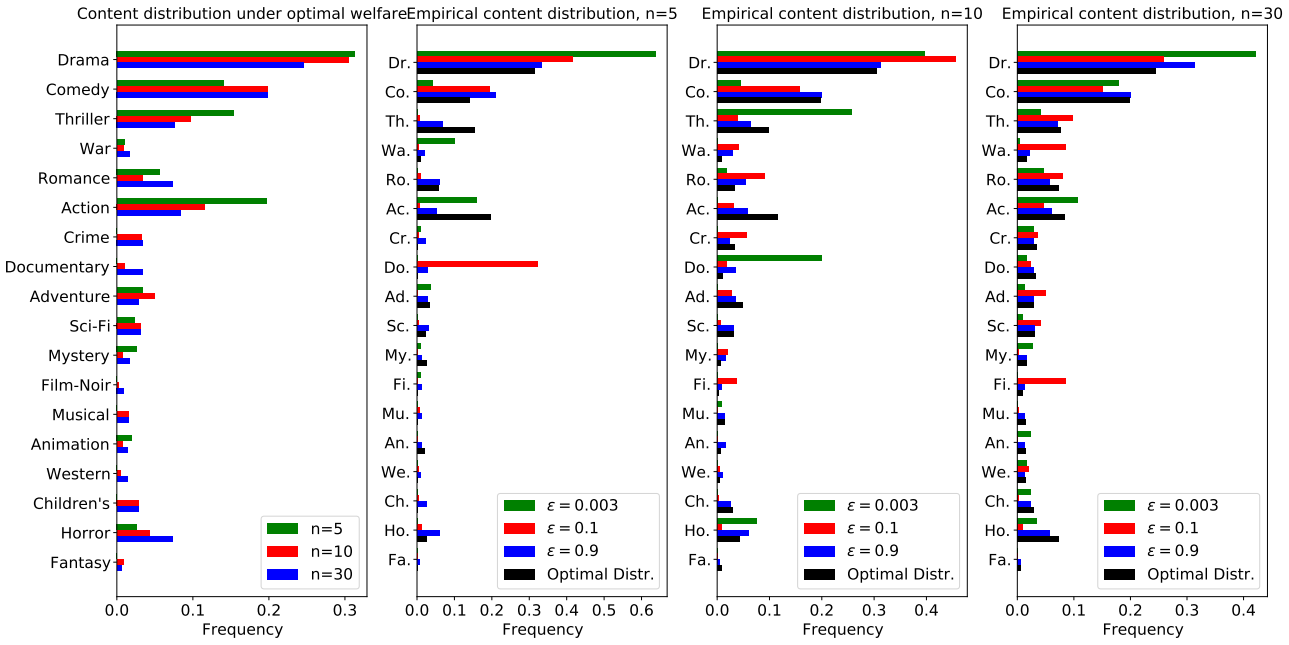


Figure 5. The genre distribution induced by creators' adaptive content creation strategies under different n. The left most panel shows the optimal genre distribution of the joint content creation strategy that maximizes social welfare under $n = 5, 10, 30$. The remaining 3 panels from left to right correspond to the genre distribution obtained from simulation under $n = 5, 10, 30$. The green/red/blue colors in the 3 panels on the right represent simulation results with different exploration rate $\epsilon = 0.003, 0.1, 0.9$ and the black bars denote the optimal distribution. All the frequency values are averaged over 5 independent runs.