Rethinking Incentives in Recommender Systems: Are Monotone Rewards Always Beneficial?

Fan Yao [1] **Chuanhao Li** [1] **Karthik Abinav Sankararaman** [2] **Yiming Liao** [2] **Yan Zhu** [3] **Qifan Wang** [2] **Hongning Wang** [1] **Haifeng Xu** [4]

Abstract

The past decade has witnessed the flourishing of a new profession as media content creators, who rely on revenue streams from online content recommendation platforms. The rewarding mechanism employed by these platforms creates a competitive environment among creators which affect their production choices and, consequently, content distribution and system welfare. In this work, we uncover a fundamental limit about a class of widely adopted mechanisms, coined *Merit-based Monotone Mechanisms*, by showing that they inevitably lead to a constant fraction loss of the welfare. To circumvent this limitation, we introduce *Backward Rewarding Mechanisms* (BRMs) and show that the competition games resulting from BRM possess a potential game structure, which naturally induces the strategic creators' behavior dynamics to optimize any given welfare metric. In addition, the class of BRM can be parameterized so that it allows the platform to directly optimize welfare within the feasible mechanism space even when the welfare metric is not explicitly defined.

1. Introduction

Online recommendation platforms, such as Instagram and YouTube, have become an integral part of our daily life (Bobadilla et al., 2013). Their impact extends beyond merely aligning users with the most relevant content: they are also accountable for the online ecosystem it creates and the long-term welfare it promotes, considering the complex dynamics driven by the potential strategic behaviors of content creators (Qian & Jain, 2022). Typically, creators' utilities are directly tied to the visibility of their content or economic incentives they can gather from the platform, and they constantly pursue to maximize these benefits(Glotfelter, 2019; Hodgson, 2021). This fosters a competitive environment that may inadvertently undermine the *social welfare*, i.e., the total utilities of all users and content creators in the system (Fleder & Hosanagar, 2009).

To account for the potentially negative effects induced by strategic content creators, the platform can design reward signals that influence the creators' perceived utilities, thereby steering content distribution at equilibrium towards enhanced social welfare. In reality, many platforms share revenue with creators via various mechanisms (Meta, 2022; Savy, 2019; Youtube, 2023; TikTok, 2022). These incentives are typically proportional to user satisfaction measured by various metrics, such as click-through rate and engagement time. We model such competitive environment within a general framework termed *content creator competition* (C^3) game that generalizes and abstracts a few established models including (Yao et al., 2023; Ben-Porat & Tennenholtz, 2018; Jagadeesan et al., 2022; Hron et al., 2022), and frame a class of prevailing rewarding mechanisms as Merit-based Monotone Mechanisms ($\mathcal{M}^3$). The $\mathcal{M}^3$ are characterized by a few simple properties, intuitively meaning better content should be rewarded more (i.e., merit-based) and sum of creators' utilities increase whenever any creator increases her content relevance (i.e., monotone). However, we show that $\mathcal{M}^3$ necessarily incur a constant fraction of welfare loss in natural scenarios due to the fact that it undesirably encourages excessive concentration of creators around the majority user groups and leaves minority groups underserved. To resolve this issue, we introduce Backward Rewarding Mechanisms (BRM) which discards monotonicity in the sense that when creators' competition within some user group surpasses a limit that begins to harm welfare, their total reward is reduced. The strength of BRM lies in three aspects: 1. any C^3 game under BRM forms a potential game (Monderer & Shapley, 1996); 2. there exists a BRM instance such that the induced potential function is equivalent to the social welfare metric. As a result, the net effect of creators' competition aligns perfectly with maximizing the social welfare. 3. BRM contains a parameterized subspace that allows empirical optimization of social welfare. This flexibility becomes especially crucial in practice when the

[1]Department of Computer Science, University of Virginia, USA [2]Meta, USA [3]Google, USA [4]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 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

welfare is not explicitly defined. These merits of BRM are supported by our empirical studies, in which we developed simulated environments, demonstrating the welfare induced by BRM outperforms baseline mechanisms in $\mathcal{M}^3$.

2. Related Work

The theoretical studies of content creators' strategic behavior under the mediation of an RS date back to the seminal work from (Ben-Porat & Tennenholtz, 2017; 2018), where they proposed the Shapley mediator that guarantees the existence of pure Nash equilibrium (PNE) and several fairness-related requirements. In these studies, the RS was only empowered to design the matching probability, and it was observed that user welfare could be significantly compromised. In contrast, our work considers a more realistic and contemporary platform that can determine the reward for each content creator. We propose BRM and show that the Shapley mediator (Ben-Porat & Tennenholtz, 2018) is an instance of BRM. Several recent work (Hron et al., 2022; Jagadeesan et al., 2022; Yao et al., 2023) studied the properties of supply-side equilibrium under the C^3 game. In (Hron et al., 2022; Jagadeesan et al., 2022), creators are assumed to directly compete for user exposure without the mediation of an RS. These studies focus on characterizing the properties of Nash Equilibrium (NE). On the other hand, (Yao et al., 2023) demonstrated that the user welfare loss under a conventional RS using top-K ranking is upper-bounded by $O(\frac{1}{\log K})$. Our research reveals that any merit-based monotone mechanism, including those based on user exposure or engagement, unavoidably incurs a welfare loss of $\frac{1}{K}$. However, if the platform have the freedom to design arbitrary incentive signals, social welfare can be further optimized.

3. A General Model for Content Creator Competition

In this section, we formalize the Content Creator Competition (C^3) game under the platform's rewarding mechanisms. Each C^3 instance $\mathcal{G}$ is defined by a tuple $(\{\mathcal{S}_i\}_{i=1}^n, \{c_i\}_{i=1}^n, \mathcal{X}, \sigma, M, \{r_i\}_{i=1}^n)$:

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

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. For simplicity, we use $\sigma_{i,j}(\boldsymbol{s})$ to denote $\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j)$ given any joint strategy profile $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$. We may omit the argument $\boldsymbol{s}$ when the context is clear.

Rewarding mechanisms: given joint strategy $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$, the platform generates a reward $u_{i,j} \in [0, 1]$ for each user-creator pair $(\boldsymbol{s}_i, \boldsymbol{x}_j)$. We generally allow $u_{i,j}$ to depend on $\boldsymbol{s}_i$'s relevance $\sigma_{i,j}$ and also other creators' relevance $\sigma_{-i,j} = \{\sigma_{t,j} | 1 \leq t \leq n, t \neq i\}$. Thus, a rewarding mechanism M is a mapping from $(\sigma_{i,j}, \{\sigma_{-i,j}\})$ to $[0, 1]$, or formally,

$$u_{i,j} = M(\sigma_{i,j}, \sigma_{-i,j}). \tag{1}$$

Such rewarding mechanisms given by function $M(\cdot, \cdot)$ can be understood as the expected payoff for creator i under any user-content matching strategy and some post-matching rewarding scheme.

Creator utilities: creator-i's utility is defined as the sum of the reward gained from each individual user minus the cost for producing content $\boldsymbol{s}_i$, i.e.,

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} u_{i,j} - c_i(\boldsymbol{s}_i), \forall i \in [n], \tag{2}$$

where c_i is the cost function for creator-i. As an example, one may have $c_i(\boldsymbol{s}_i) = \lambda_i \|\boldsymbol{s}_i - \bar{\boldsymbol{s}}_i\|_2^2$ where $\bar{\boldsymbol{s}}_i$ represents the type of content that creator i is most comfortable or confident with.

User utility and the social welfare: Before formalizing the welfare objective, we first define each user-j's utility from consuming a list of ranked content. Since user's attention usually decreases in the rank positions, we introduce discounting weights $\{r_{k,j} \in [0, 1]\}_k$ for each user-j to represent his/her "attention" over the k-th *ranked* content. Naturally, we assume $r_{1,j} \geq \cdots \geq r_{n,j}$, i.e., higher ranked content receives more user attention. Consequently, user-j's utility from consuming a list of content $\{l_j(k)\}_{k=1}^n$, which is a permutation of $[n]$ ranked in a descending order of content relevance to user-j (i.e., $\sigma_{l_j(1),j} \geq \cdots \geq \sigma_{l_j(n),j}$), is defined by the following weighted sum

$$W_j(\boldsymbol{s}) = \sum_{k=1}^{n} r_{k,j} \sigma_{l_j(k),j}(\boldsymbol{s}). \tag{3}$$

In the special case where the platform always recommends the top-K ranked content, we shall have $r_i = 0, \forall i > K$. We refer to this important special case as the *top-K environment*, which is popular in almost all practical RS. A concrete example is when $r_{k,j} = \frac{1}{\log_2(k+1)}$, which reduces $W_j(\boldsymbol{s})$ to the well-known metric Discounted Cumulative Gain (DCG) (Järvelin & Kekäläinen, 2002) used in information retrieval. We remark that the user utility definition (3) is compatible with most natural ranking strategies in addition to top-K; we provide additional examples of more general ranking rules in Appendix A.1.

Finally, the social welfare is defined as the sum of total user utilities and total creator utilities, minus the platform's cost:

$$W(\boldsymbol{s};\{r_{k,j}\}) = \sum_{j=1}^{m} W_j(\boldsymbol{s}) + \sum_{i=1}^{n} u_i(\boldsymbol{s}) - \sum_{i=1}^{n}\sum_{j=1}^{m} u_{i,j} = \sum_{j=1}^{m} W_j(\boldsymbol{s}) - \sum_{i=1}^{n} c_i(\boldsymbol{s}_i). \quad (4)$$

To simplify our discussion in the following sections, we assume $\{r_{k,j}\}$ is independent of the user index j and simply use r_k in place of $r_{k,j}$. Our results can also be generalized to $r_{k,j}$ with dependency on j, up to more complex notations.

Our research questions: creator incentive design. Unlike previous works (Ben-Porat et al., 2019; Hron et al., 2022; Jagadeesan et al., 2022) that primarily focus on designing user-content matching mechanisms, we consider the design of a different "knob" to improve social welfare, i.e., creators' rewarding scheme. Each rewarding mechanism M establishes a competitive environment among content creators, encapsulated by a C^3 instance $\mathcal{G}(\{\mathcal{S}_i\},\{c_i\},\mathcal{X},\sigma,M,\{r_i\})$. Our objective is to design mechanisms M that: 1. guarantee the existence of PNE, thereby ensuring a stable outcome, and 2. maximize social welfare at the PNE.

4. The Fundamental Limit of Merit-based Monotone Mechanisms

In this section, we show an intrinsic limitation for a generic class of reward mechanisms commonly utilized in the practice. These reward schemes, employed by numerous platforms, exhibit common properties that can be summarized by the following definition.

Definition 1 (Merit-based Monotone Mechanism ($\mathcal{M}^3$)). We say M is a *merit-based monotone mechanism* if for any relevance scores $1 \geq \sigma_1 \geq \cdots \geq \sigma_n \geq 0$, M satisfies the following properties:

- Merit-based:
  - (Normality) $M(0,\sigma_{-i})=0$, $M(1,\{0,\cdots,0\})>0$,
  - (Fairness) $M(\sigma_i,\sigma_{-i}) \geq M(\sigma_j,\sigma_{-j}), \forall i>j$,
  - (Negative Externality) $\forall i$, if $\sigma_{-i} \preccurlyeq \sigma'_{-i}$ ($\sigma_j \leq \sigma'_j, \forall j \neq i$), then $M(\sigma_i,\sigma_{-i}) \geq M(\sigma_i,\sigma'_{-i})$.
- Monotonicity: the total rewards $\sum_{i=1}^{n} M(\sigma_i,\sigma_{-i}) : [0,1]^n \to \mathbb{R}_{\geq 0}$ is non-decreasing in σ_i for every $i \in [n]$.

In addition, we write $M \in \mathcal{M}^3$ if M is a merit-based monotone mechanism with parameter n.

The two properties underpinning $\mathcal{M}^3$ are quite intuitive. Firstly, the merit-based property consists of three natural sub-properties: 1. zero relevance content should receive zero reward, whereas the highest relevance content deserves a non-zero reward; 2. within the given pool of content with scores $\{\sigma_i\}_{i\in[n]}$, the higher relevance content should receive a higher reward; 3. any individual content's reward does not increase when other creators improve their content relevance. Secondly, monotonicity means if any content creator i improves her relevance σ_i, the total rewards to all creators increase. This property is naturally satisfied by many widely adopted rewarding mechanisms because platforms in today's industry typically reward creators proportionally to user engagement or satisfaction, the *total* of which is expected to increase as some creator's content becomes more relevant. Unsurprisingly, many popular rewarding mechanisms falls into the class of $\mathcal{M}^3$. For instances, two mechanisms that are widely adopted in current industry practices for rewarding creators (Meta, 2022; Savy, 2019; TikTok, 2022; Youtube, 2023) are exposure/engagement based, i.e., creators' utilities are set to the total content exposure/engagement. We show in Appendix A.2 that both of them are in $\mathcal{M}^3$. However, we show that *any* mechanism in $\mathcal{M}^3$ can result in a quite suboptimal welfare, even applied to some natural C^3 game environment. Specifically, we abstract out the following representative subset of C^3 instances, which capture the essence of many real-world situations. We term this subclass of situations as *Trend v.s. Niche* (TvN) games.

Definition 2 (TvN games). Denote $E=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\} \subset \mathbb{R}^n$ as the set of unit basis vectors in $\mathbb{R}^n$. Consider a user population $\mathcal{X}=\{\boldsymbol{x}_j\}_{j=1}^{2n}$ such that $\boldsymbol{x}_j=\boldsymbol{e}_1$, for $1 \leq j \leq n+1, \boldsymbol{x}_{n+2}=\boldsymbol{e}_2,\cdots,\boldsymbol{x}_{2n}=\boldsymbol{e}_n$. All creators have zero costs and share the same action set $\mathcal{S}_i=E$. The relevance is measured by the inner product, i.e., $\sigma(\boldsymbol{s},\boldsymbol{x})=\boldsymbol{s}^\top\boldsymbol{x}$. The attention discounting weights $\{r_i\}$ is induced by a top-K environment, i.e., $r_1 \geq \cdots \geq r_K \geq r_{K+1}=\cdots=r_n=0$. For any mechanism M, we call such a game instance $\mathcal{G}(\{\mathcal{S}_i\},\{c_i=0\},\mathcal{X},\sigma,M,\{r_i\})$ a TvN game.

The TvN game models a scenario where the user population comprises multiple interest groups, each with orthogonal preference representations. In this game, the largest group consists of nearly half the population. Each content creator has the option to cater to one—and only one—user group. While this game is simple and stylized, it captures the essence of real-world user populations and the dilemmas faced by creators. Our subsequent result shows that if the platform adopts $\mathcal{M}^3$ in the TvN game, this tension of content creation turns out to be a curse in the sense that a unique PNE is achieved when all players opt for the same strategy — catering to the largest user group — and we quantify the social welfare loss at this PNE in the following.

Theorem 1. *For any rewarding mechanism $M \in \mathcal{M}^3$ applied to any TvN instance, we have: 1. the resultant game admits a unique NE $\boldsymbol{s}^*$; 2. the welfare of this NE is at most*

$\frac{K}{K+1}$ *fraction of the optimal welfare for large n. Formally,*

$$\frac{W(\boldsymbol{s}^*)}{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})} \le \frac{K}{K+1} + O\left(\frac{1}{n}\right). \quad (5)$$

The proof is in Appendix A.3, where we explicitly characterize both $\boldsymbol{s}^*$ and the welfare maximizing strategy profile and calculate their difference in terms of welfare. Eq. (5) suggests that the welfare loss of $\mathcal{G}$ under $\mathcal{M}^3$ could be as significant as $1/2$ for users who primarily care about the top relevant content, which is shown to be realistic given the diminishing attention spans of Internet users (Carr, 2020).

5. Backward Rewarding Mechanisms

In observation of the negative result in Theorem 1, we introduce the class of Backward Rewarding Mechanisms (BRMs), in which the monotonicity is compromised. The name of BRM suggests its essential characteristic: the reward for a specific creator-i depends solely on their ranking and the relevance of content from creators ranked *lower* than i. The formal definition of BRM is provided below:

Definition 3 (BRM and BRCM)**.** We say M is a Backward Rewarding Mechanism (BRM) if for any relevance score sequence $1 \ge \sigma_1 \ge \cdots \ge \sigma_n \ge 0$, there exists a sequence of Riemann integrable functions $\{f_i(t) : [0,1] \to \mathbb{R}_{\ge 0}\}_{i=1}^n$, $f_1(t) \ge \cdots \ge f_n(t)$ such that

$$M(\sigma_i, \sigma_{-i}) = \sum_{k=i}^{n} \int_{\sigma_{k+1}}^{\sigma_k} f_k(t)dt, \quad (6)$$

where $\sigma_{n+1} = 0$ and $f_1(t) > 0, \forall t$. We use BRM to denote the set of all BRMs with parameter n, and use $M[f_1(t), \cdots, f_n(t)]$ to denote an element $M \in$ BRM when it is associated with $\{f_i(t)\}_{i=1}^n$.

In addition, we identify a subset BRCM$\subset$BRM which includes those M such that $\{f_i(t) \equiv f_i\}$ are a set of constant functions. Clearly, any $M \in$ BRCM can be parameterized by a n-dimensional point in the polytope $\mathcal{F} = \{(f_1, \cdots, f_n) | 1 \ge f_1 \ge \cdots \ge f_n \ge 0\}$.

To get an better intuition of how BRM works, let us consider a special case $M \in$ BRCM such that $f_1 = \cdots = f_K = 1$ and $f_k = 0, k \ge K+1$. By the definition, any score sequence $\sigma_1 \ge \cdots \ge \sigma_n$ will be mapped to a reward sequence of $(\sigma_1 - \sigma_{K+1}, \cdots, \sigma_K - \sigma_{K+1}, 0, \cdots, 0)$. Consequently, the top-$K$ ranked creators will experience a significant reduction in rewards if the $(K+1)$-th ranked creator increases its content relevance. This mechanism can deter an unnecessary concentration of creators on a specific strategy, as when the number of creators with high scores exceeds a certain threshold, even those ranked highly can receive a decreasing reward. This backward rewarding mechanism thus encourages diversity in content creation and mitigates the risk of oversaturation in any particular group of users. A more comprehensive understanding about the construction of BRM can be obtained through the lens of congestion games and we defer the discussion in Appendix A.4.

5.1. Properties of BRM

While the class of BRM might appear abstract at the first glance, one can confirm that it preserves all merit-based properties, making it a natural class of rewarding mechanisms. Nevertheless, in order to secure a better welfare guarantee, the monotonicity is dropped, as characterized in the following:

Proposition 1. *Any $M \in$ BRM is merit-based but not necessarily monotone.*

The detailed proof is provided in the Appendix A.5. Next we establish formal characterizations about the welfare guarantee of BRM. First, we show that any C^3 game under BRM possesses a PNE because it is a potential game (Monderer & Shapley, 1996). A strategic game is called a potential game if there exists a function $P : \prod_i \mathcal{S}_i \to \mathbb{R}$ such that for any strategy profile $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n)$, any player-$i$ and strategy $\boldsymbol{s}'_i \in \mathcal{S}_i$, whenever player-$i$ deviates from $\boldsymbol{s}_i$ to $\boldsymbol{s}'_i$, the change of his/her utility function is equal to the change of P, i.e., $P(\boldsymbol{s}'_i, \boldsymbol{s}_{-i}) - P(\boldsymbol{s}_i, \boldsymbol{s}_{-i}) = u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i}) - u_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i})$. This leads us to the main result of this section:

Theorem 2. *Any C^3 game $\mathcal{G}(\{\mathcal{S}_i\}, \{c_i\}, \mathcal{X}, \sigma, M, \{r_i\})$ induced by any $M \in$ BRM is a potential game and thus has a PNE. Moreover, if the mechanism $M = M[r_1, \cdots, r_n] \in$ BRCM$\in$BRM, the potential function is precisely the welfare function, i.e., $W(\boldsymbol{s}) = P(\boldsymbol{s}; M)$.*

The proof is in Appendix A.6, where we construct its potential function explicitly. According to (Monderer & Shapley, 1996), we also conclude: 1. the maximizers of P are the PNEs of $\mathcal{G}$, and 2. if the evolution of creators' strategic behavior follows a better response dynamics (i.e., in each iteration, an arbitrary creator deviates to a strategy that increases his/her utility), their joint strategy profile converges to a PNE. Theorem 2 suggests another appealing property of BRM: one can always select an M within BRM to align the potential function with the welfare metric, which can be simply achieved by setting each f_i identical to r_i. Consequently, any best response dynamic among creators not only converges to a PNE but also generates a strictly increasing sequence of W, thus ensuring at least a local maximizer of W. Denote the set of PNEs of $\mathcal{G}$ as $PNE(\mathcal{G})$. When $PNE(\mathcal{G})$ coincides with the global maximizers of its potential function, i.e., $PNE(\mathcal{G}) = \text{argmax}_{\boldsymbol{s}} P(\boldsymbol{s}; M)$, we conclude that any PNE of $\mathcal{G}$ also maximizes the welfare W. The following corollary indicates that such an optimistic situation occurs in TvN games, providing a stark contrast to the findings in Theorem 1. The proof is in Appendix A.7.

Corollary 1. *For any TvN instance $\mathcal{G}$, there exists $M \in$*

BRCM such that any PNE $\boldsymbol{s}^* \in PNE(\mathcal{G})$ attains the optimal W, i.e.,

$$\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s}) = W(\boldsymbol{s}^*). \quad (7)$$

5.2. Welfare Optimization within BRCM

Theorem 2 suggests that, provided the parameters $\{r_i\}$ are known, the platform can select a mechanism within BRCM with a better welfare guarantee. However, in many practical scenarios, the platform may not have access to the exact values of $\{r_i\}$ but can only evaluate the resulting welfare metric using certain aggregated statistics. This presents a challenge as it may not be analytically feasible to pinpoint the optimal M as suggested by Theorem 2. In these cases, we can formulate the following bi-level optimization:

$$\max_{M\in BRCM} \quad W(\boldsymbol{s}^*(M)) \quad (8)$$

$$\text{s.t.,} \quad \boldsymbol{s}^*(M) = \operatorname*{argmax}_{\boldsymbol{s}} P(\boldsymbol{s}; M) \quad (9)$$

In problem (8), the inner optimization (9) is executed by creators: for any given M, we have justified that the creators' strategies is very likely to settle at a PNE $\boldsymbol{s}^*(M)$ that corresponds to a maximizer of $P(\boldsymbol{s}; M)$. However, the exact solution to the inner problem is neither analytically solvable by the platform (owing to the combinatorial nature of P) nor observable from real-world feedback (due to creators' potentially long feedback cycles). Therefore, we propose to approximate its solution by simulating creators' strategic response sequences (See Appendix A.8, Algorithm 2), on top of which we solve (8). Algorithm 2 is a variant of better response dynamics, incorporating randomness and practical considerations to more accurately emulate creator behavior, and will be employed as a subroutine in Algorithm 1.

Another challenge of solving (8) lies in the presence of ranking operations in W, which makes it non-differentiable in $\boldsymbol{s}$ and renders first-order optimization techniques ineffective. Consequently, we resort to coordinate update and apply finite differences to estimate the ascending direction of W with respect to each M parameterized by $\boldsymbol{f} = (f_1, \cdots, f_n) \in \mathcal{F}$. Our proposed optimization algorithm for solving (8) is presented in Algorithm 1, which is structured into L_1 epochs. At the beginning of each epoch, the optimizer randomly perturbs the current M along a direction within the feasible polytope and simulates creators' responses for L_2 steps using Algorithm 2. Welfare is re-evaluated at the end of this epoch, and the perturbation on M is adopted if it results in a welfare increase.

6. Experiments

To validate our theoretical findings and demonstrate the efficacy of Algorithm 1, we simulate the strategic behavior of content creators and compare the evolution of social welfare under various mechanisms. These include Algorithm 1 and several baselines from both the $\mathcal{M}^3$ and BRCM classes.

Algorithm 1 Optimize W in BRCM

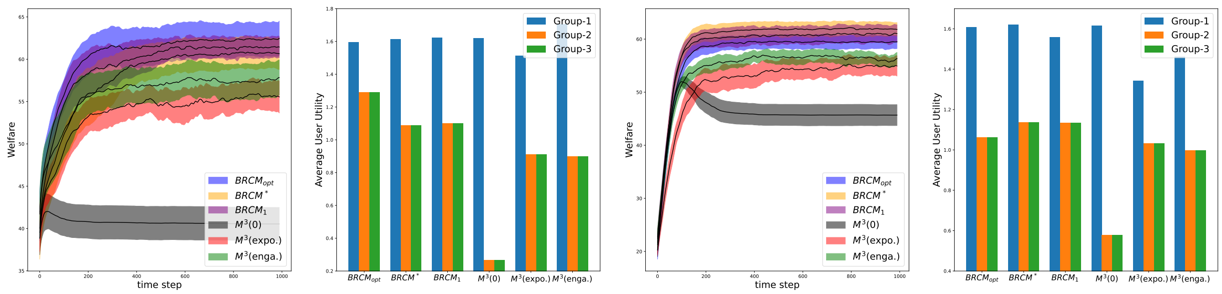
Input: Time horizon $T = L_1 L_2$, learning rate η_1, η_2, $(u_i(\boldsymbol{s}), \mathcal{S}_i)$ for each creator.
Initialization: Unit basis $\{\boldsymbol{e}_i\}_{i=1}^n$ in $\mathbb{R}^n$, initial strategy profile $\boldsymbol{s}^{(0)} = (\boldsymbol{s}_1^{(0)}, \cdots, \boldsymbol{s}_n^{(0)})$, initial parameter $\boldsymbol{f}^{(0)} = (f_1^{(0)}, \cdots, f_n^{(0)}) \in \mathcal{F}$ and mechanism $M[\boldsymbol{f}^{(0)}]$.
for $t = 0$ **to** $L_1 - 1$ **do**
  Generate $i \in [n]$ and $\boldsymbol{g}_i \in \{-\boldsymbol{e}_i, \boldsymbol{e}_i\}$ uniformly at random.
  Update $\boldsymbol{f}_i^{(t+\frac{1}{2})}$ as the projection of $\boldsymbol{f}_i^{(t)} + \eta_1 \boldsymbol{g}_i$ on $\mathcal{F}$.
  Simulate $\boldsymbol{s}^{(t+1)} =$ **simStra**$(\boldsymbol{s}^{(t)}; L_2, \eta_2, \{\boldsymbol{u}_i, \mathcal{S}_i\}_{i=1}^n, M[\boldsymbol{f}^{(t+\frac{1}{2})}])$.
  // Implemented by Algo. 2
  if $W(\boldsymbol{s}^{(t+1)}) > W(\boldsymbol{s}^{(t)})$ **then**
    $\boldsymbol{f}^{(t+1)} = \boldsymbol{f}^{(t+\frac{1}{2})}$.
  else
    $\boldsymbol{f}^{(t+1)} = \boldsymbol{f}^{(t)}$.

6.1. Specification of Environments

We conduct simulations on game instances $\mathcal{G}(\{\mathcal{S}_i\}, \{c_i\}, \mathcal{X}, \sigma, M, \{r_i\})$ constructed from synthetic data and MovieLens-1m dataset (Harper & Konstan, 2015). Result on MovieLens is deferred to Appendix A.9.

For the synthetic data, we first generate the user population $\mathcal{X}$ as follows: we fix the embedding dimension d and randomly sample Y cluster centers, denoted as $\mathbf{c}_1, \cdots, \mathbf{c}_Y$, on the unit sphere $\mathbb{S}^{d-1}$. For each center $\mathbf{c}_i$, we generate users belonging to cluster-i by first independently sampling from a Gaussian distribution $\tilde{\boldsymbol{x}} \sim \mathcal{N}(\mathbf{c}_i, v^2 I_d)$, and then normalize it to $\mathbb{S}^{d-1}$, i.e., $\boldsymbol{x} = \tilde{\boldsymbol{x}}/\|\tilde{\boldsymbol{x}}\|_2$. The sizes of the Y user clusters are denoted by a vector $\boldsymbol{z} = (z_1, \cdots, z_Y)$. In this manner, we generate a population $\mathcal{X} = \cup_{i=1}^Y \mathcal{X}_i$ with size $m = \sum_{i=1}^Y z_i$. The number of creators is set to $n = 10$, with action sets $\mathcal{S}_i = \mathbb{S}^{d-1}$. The relevance function $\sigma(\boldsymbol{x}, \boldsymbol{s}) = \frac{1}{2}(\boldsymbol{s}^\top \boldsymbol{x} + 1)$ is the shifted inner product such that its range is exactly $[0, 1]$. $\{r_i\}_{i=1}^n$ is set to $\{\frac{1}{\log_2(2)}, \frac{1}{\log_2(3)}, \frac{1}{\log_2(4)}, \frac{1}{\log_2(5)}, \frac{1}{\log_2(6)}, 0, \cdots, 0\}$.

We consider two types of game instances, denoted $\mathcal{G}_1$ and $\mathcal{G}_2$, distinguished by their cost functions. In $\mathcal{G}_1$, creators have zero cost and their initial strategies are set to the center of the largest user group. This environment models the situation where the social welfare is already trapped at suboptimal due to its unbalanced content distribution. In $\mathcal{G}_2$, creators have non-trivial cost functions $c_i = 0.5\|\boldsymbol{s}_i - \bar{\boldsymbol{s}}_i\|_2^2$, where the cost center $\bar{\boldsymbol{s}}_i$ is randomly sampled on $\mathbb{S}^{d-1}$. Their initial strategies are set to the corresponding cost centers, i.e., all creators start with strategies that minimize their costs. This environment models a "cold start" situation for creators: they do not have any preference nor knowledge about the user population and gradually learn about the environment under the platform's incentivizing mechanism. In our experiment, we set $(d, v, Y, m) = (10, 0.3, 8, 52)$ and



(a) Welfare on $\mathcal{G}_1$ (b) Avg. Group Util. on $\mathcal{G}_1$ (c) Welfare on $\mathcal{G}_2$ (d) Avg. Group Util. on $\mathcal{G}_2$

Figure 1. Social welfare curve and average user utilities per group. Error bars represent half standard deviation range (0.5σ), and are generated from simulations on 10 randomly sampled game instances.

the cluster sizes $\boldsymbol{z} = (20, 10, 8, 5, 3, 3, 2, 1)$. The 8 clusters are devided into 3 groups $((20), (10, 8), (5, 3, 3, 2, 1))$, namely group-1,2,3, corresponding to the majority, minority, and niche groups.

6.2. Algorithm and Baseline Mechanisms

We simulate the welfare curve produced by Algorithm 1 alongside five baseline mechanisms below. **BRCM**$_{opt}$: This refers to the dynamic mechanism realized by optimization Algorithm 1. The starting point is set to $\boldsymbol{f}^{(0)} = (1, 1, 1, 1, 1, 0, \cdots, 0)$. The parameters are set to $T = 1000, L_1 = 200, L_2 = 5, \eta_1 = \eta_2 = 0.1$. **BRCM***: This denotes the theoretically optimal mechanism within BRCM, as indicated by Theorem 2. The corresponding parameters of M are derived based on the knowledge of $\{r_i\}_{i=1}^n$. **BRCM**$_1$: $\text{BRCM}_1 = M[1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, 0, \cdots, 0] \in \text{BRCM}$. This baseline aims to assess the impact of deviation from the theoretically optimal mechanism on the result. $\mathbf{M}^3(0)$: This mechanism assigns each content creator a reward equal to the relevance score, i.e., $M(\sigma_i, \sigma_{-i}) = \sigma_i$. It is obvious that this mechanism belongs to the $\mathcal{M}^3$ class and is therefore denoted as $M^3(0)$. Under $M^3(0)$, each creator's strategy does not affect other creators' rewards at all, and thus every creator will be inclined to match the largest user group as much as their cost allows. This mechanism acts as a reference to indicate the worst possible scenario. $\mathbf{M}^3(expo.)$: The mechanism based on exposure, defined in Section 4 with $K = 5, \beta = 0.05$. $\mathbf{M}^3(enga.)$: The mechanism based on engagement, defined in Section 4 with $K = 5, \beta = 0.05$.

6.3. Results

We record the social welfare and average group utility distribution at the end of simulations with Algorithm 2. The results under two environments are shown in Figure 1. As illustrated in Figure 1(a), BRCM family consistently outperformed $\mathcal{M}^3$. As anticipated, $M^3(0)$ does little to enhance social welfare when creators have already primarily focused on the most populous user group. The $M^3(expo.)$ and $M^3(enga.)$ mechanisms demonstrate a notable improvement over $M^3(0)$ as they instigate a competitive environment for creators striving to reach the top-K positions. Nevertheless, they still do not perform as effectively as BRCM_1, even though BRCM_1's parameter deviates from the theoretically optimal one. Within the BRCMs, BRCM_{opt} exhibits remarkable performance and even surpasses the theoretically optimal instance BRCM^*. One possible explanation for the empirical sub-optimality of BRCM^* is the stochastic nature of creators' response dynamics, which prevent the convergence to PNE associated with the maximum welfare without sufficient optimization. This observation underscores the importance of Algorithm 1, as it empowers the platform to pinpoint an empirically optimal mechanism in more practical scenarios. As depicted in Figure 1(b), the primary source of advantage stems from the increased utility among minority and niche user groups: compared to $M^3(expo.)$ and $M^3(enga.)$, BRCM class results in higher average utility for groups 2 and 3 while preserving overall satisfaction for group-1. Similar observations can be made for $\mathcal{G}_2$. However, it is worth noting that BRCM_{opt} underperformed slightly in comparison to BRCM^* as shown in Figure 1(c). Despite this, the BRCM class of mechanisms continued to significantly surpass those in $\mathcal{M}^3$. Figure 1(d) highlights that BRCM mechanisms lead to a more equitable distribution of average user utility across different user groups. Nevertheless, the gap in comparison becomes less pronounced, which is probably because creators burdened with such costs are inherently inclined towards serving specific user groups, making them less susceptible to the influence of platform's incentives.

7. Conclusion

Our work reveals an intrinsic limitation of the monotone reward principle, widely used by contemporary online content recommendation platforms to incentivize content creators, in optimizing social welfare. As a rescue, we introduce BRM, which ensures a stable outcome, guides content creators' strategic responses towards optimizing social welfare, and offers a parameterized subspace that allows the platform to empirically optimize social welfare.

References

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.

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

Carr, N. *The shallows: What the Internet is doing to our brains*. WW Norton & Company, 2020.

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.

Glotfelter, A. Algorithmic circulation: how content creators navigate the effects of algorithms on their work. *Computers and composition*, 54:102521, 2019.

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.

Hodgson, T. Spotify and the democratisation of music. *Popular Music*, 40(1):1–17, 2021.

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.

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

Järvelin, K. and Kekäläinen, J. Cumulated gain-based evaluation of ir techniques. *ACM Transactions on Information Systems (TOIS)*, 20(4):422–446, 2002.

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.

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

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

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/.

Yao, F., Li, C., Nekipelov, D., Wang, H., and Xu, H. How bad is top-k recommendation under competing content creators? In *International Conference on Machine Learning*. PMLR, 2023.

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

A. Supplementary Material

A.1. Additional Examples of User Utility Function

For simplicity of notations we always assume $\sigma_1 \geq \cdots \geq \sigma_n$. As discussed in Section 3, if the platform presents the top-K ranked content in terms of their relevance quality, user j's utility function has the following form:

$$W_j(\boldsymbol{s}) = \sum_{k=1}^{n} r_k \sigma_k, \tag{10}$$

where $\sigma_k = \sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)$ is the relevance score between user j and content creator ranked at the k-th position, and $\{r_k\}_{k=1}^n$ are user-j's "attention" over the k-th ranked content such that $r_k = 0, k \geq K+1$.

It is worth noting that our user utility model is compatible with various matching strategies. In this regard, we provide additional examples that incorporate a modified version of the top-K approach, taking into account considerations of advertised content. For instance, let's consider a scenario where $K = 5$ and the platform intends to promote the content originally ranked at position 6 to position 2 with probability $p \in (0, 1)$. Consequently, the resulting utility function can be expressed as follows:

$$\begin{aligned}\tilde{W}_j(\boldsymbol{s}) &= p(r_1\sigma_1 + r_2\sigma_6 + r_3\sigma_2 + r_4\sigma_3 + r_5\sigma_4) + (1-p)(r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3 + r_4\sigma_4 + r_5\sigma_5) \\ &= r_1\sigma_1 + [pr_3 + (1-p)r_2]\sigma_2 + [pr_4 + (1-p)r_3]\sigma_3 + [pr_5 + (1-p)r_4]\sigma_4 + (1-p)r_5\sigma_5 + pr_2\sigma_6 \\ &\triangleq \sum_{k=1}^{n} \tilde{r}_k \sigma_k.\end{aligned}$$

This example shows that user utility function under any position-based perturbation of top-K ranking can be expressed in the form of (10), and in general the values of $r_k, k > K$ can be non-zero.

A.2. Examples of $\mathcal{M}^3$

In this section we formally justify that the two examples given in Section 4 belong to the class of $\mathcal{M}^3$.

1. When the creators' utilities are set to the total content exposure (Ben-Porat et al., 2019; Hron et al., 2022; Jagadeesan et al., 2022), we have $M(\sigma_i, \sigma_{-i}) = \mathbb{I}[i \leq K] \frac{\exp(\beta^{-1}\sigma_i)}{\sum_{j=1}^K \exp(\beta^{-1}\sigma_j)}$, with a temperature parameter $\beta > 0$ controlling the spread of rewards.

   The validity of three merit-based properties are straightforward. In terms of monotonicity, we have $\sum_{i=1}^n M(\sigma_i, \sigma_{-i}) = 1$ which is a constant and thus monotone.

2. When the creators' utilities are set to the total user engagement (Yao et al., 2023), we have $M(\sigma_i, \sigma_{-i}) = \mathbb{I}[i \leq K] \frac{\exp(\beta^{-1}\sigma_i)}{\sum_{j=1}^K \exp(\beta^{-1}\sigma_j)} \pi(\sigma_1, \cdots, \sigma_n)$, where $\pi(\sigma_1, \cdots, \sigma_n) = \beta \log\left(\sum_{j=1}^K \exp(\beta^{-1}\sigma_j)\right)$.

   The first two merit-based properties are obvious (Normality and Fairness). In terms of monotonicity, we have $\sum_{i=1}^n M(\sigma_i, \sigma_{-i}) = \beta \log\left(\sum_{j=1}^K \exp(\beta^{-1}\sigma_j)\right)$ which is monotone in each σ_j. To verify negative externality, it suffices to show the function $\frac{\log\left(\sum_{j=1}^K \exp(\beta^{-1}\sigma_j)\right)}{\sum_{j=1}^K \exp(\beta^{-1}\sigma_j)}$ is decreasing in $\sigma_j, \forall j$. Since $\exp(x)$ is increasing in x, and function $\frac{\log(t)}{t}$ is decreasing when $t > e$, we conclude that M satisfies negative externality when $n \geq 3$.

A.3. Proof of Theorem 1

Before showing the proof, we define the following notion of *local* maximizer:

Definition 4. We say $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n)$ is a *local* maximizer of $W(\boldsymbol{s})$ if for any $i \in [n]$ and any $\boldsymbol{s}_i' \in \mathcal{S}_i$,

$$W(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_i, \cdots, \boldsymbol{s}_n) \geq W(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_i', \cdots, \boldsymbol{s}_n).$$

The set of all the local maximizers of W is denoted by $Loc(W)$.

According to the definition, for any join strategy profile $\boldsymbol{s} \in Loc(W)$, no creator can unilaterally change his/her strategy to increase the value of function W. And clearly we have $\arg\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s}) \in Loc(W)$. Now we are ready to present the proof of Theorem 2.

Proof. We start by showing that any TvN game instance with $M \in \mathcal{M}^3$ possesses a unique NE at $\boldsymbol{s}^* = (\boldsymbol{e}_1, \cdots, \boldsymbol{e}_1)$. It suffices to show that:

1. For any joint strategy profile $(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n)$ in which there are $k < n$ creators occupy $\boldsymbol{e}_1$, there exists a creator who can receive a strict utility gain if she change her strategy to $\boldsymbol{e}_1$.

2. At $\boldsymbol{s}^* = (\boldsymbol{e}_1, \cdots, \boldsymbol{e}_1)$, any player would suffer a utility loss when changing her strategy.

For the first claim, suppose there are k players in $\boldsymbol{s}$ who play $\boldsymbol{e}_1$ and let i be any player who does not play $\boldsymbol{e}_1$. In addition, there are $t \le n-k$ players who play the same strategy as $\boldsymbol{s}_i$. By the definition of M_3, we have

$$\begin{aligned} u_i(\boldsymbol{s}_i;\boldsymbol{s}_{-i}) &= 1\cdot M(\underbrace{1,\cdots,1}_{t},\underbrace{0,\cdots,0}_{n-t}) + (n+1)\cdot M(\underbrace{0,\cdots,0}_{n-k},\underbrace{1,\cdots,1}_{k}) \\ &= 1\cdot\frac{1}{t}\cdot\pi(\underbrace{1,\cdots,1}_{t},\underbrace{0,\cdots,0}_{n-t}) + (n+1)\cdot 0 \\ &= \frac{1}{t}\cdot\pi(\underbrace{1,\cdots,1}_{t},\underbrace{0,\cdots,0}_{n-t}). \end{aligned} \tag{11}$$

If player-i changes her strategy from $\boldsymbol{s}_i$ to $\boldsymbol{s}_i' = \boldsymbol{e}_1$, the new utility would be

$$\begin{aligned} u_i(\boldsymbol{s}_i';\boldsymbol{s}_{-i}) &= (n+1)\cdot M(\underbrace{1,\cdots,1}_{k+1},\underbrace{0,\cdots,0}_{n-k-1}) + \sum_{j\neq i} 1\cdot M(0,\cdots) \\ &= (n+1)\cdot\frac{1}{k+1}\cdot\pi(\underbrace{1,\cdots,1}_{k+1},\underbrace{0,\cdots,0}_{n-k-1}) + 0 \\ &= \frac{n+1}{k+1}\cdot\pi(\underbrace{1,\cdots,1}_{k+1},\underbrace{0,\cdots,0}_{n-k-1}), \end{aligned} \tag{12}$$

From (11) and (12), $u_i(\boldsymbol{s}_i';\boldsymbol{s}_{-i}) > u_i(\boldsymbol{s}_i;\boldsymbol{s}_{-i})$ holds if and only if

$$\frac{1}{t}\cdot\pi(\underbrace{1,\cdots,1}_{t},\underbrace{0,\cdots,0}_{n-t}) < \frac{n+1}{k+1}\cdot\pi(\underbrace{1,\cdots,1}_{k+1},\underbrace{0,\cdots,0}_{n-k-1}). \tag{13}$$

And a sufficient condition for (13) to hold is

$$m = 2n > n-1+\max_{0\le k\le n-1}\left\{\frac{k+1}{t}\cdot\frac{\pi(\overbrace{1,\cdots,1}^{t},\overbrace{0,\cdots,0}^{n-t})}{\pi(\underbrace{1,\cdots,1}_{k+1},\underbrace{0,\cdots,0}_{n-k-1})}\right\}. \tag{14}$$

Denote $\tilde{\pi}_k = \pi(\underbrace{1,\cdots,1}_{k},\underbrace{0,\cdots,0}_{n-k})$. By the monotonicity of π, we have $\tilde{\pi}_n \ge \cdots \ge \tilde{\pi}_1 = M(1,0,\cdots,0) > 0$. Therefore, the RHS of (14) is a finite number. Moreover, when $t \le k+1$, we have

$$\frac{k+1}{t}\cdot\frac{\tilde{\pi}_t}{\tilde{\pi}_{k+1}} \le \frac{k+1}{t}\cdot\frac{\tilde{\pi}_{k+1}}{\tilde{\pi}_{k+1}} \le \frac{n-1+1}{1} = n,$$

and when $t > k+1$, based on the negative externality principle of merit-based rewarding mechanism we have

$$\frac{k+1}{t} \cdot \frac{\tilde{\pi}_t}{\tilde{\pi}_{k+1}} = \frac{M(\overbrace{1, \cdots, 1}^{t}, \overbrace{0, \cdots, 0}^{n-t})}{M(\underbrace{1, \cdots, 1}_{k+1}, \underbrace{0, \cdots, 0}_{n-k-1})} \leq 1.$$

Therefore, the RHS of Eq. (14) is strictly less than $2n-1$.

For the second claim, we have

$$u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}^*) = \frac{n+1}{n}\tilde{\pi}_n,$$

and if player-i changes her strategy from $\boldsymbol{s}_i^* = \boldsymbol{e}_1$ to any $\boldsymbol{s}_i' = \boldsymbol{e}_j, j \neq 1$, her new utility becomes

$$u_i(\boldsymbol{s}_i'; \boldsymbol{s}_{-i}^*) = \tilde{\pi}_1 \leq \tilde{\pi}_n < \frac{n+1}{n}\tilde{\pi}_n = u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}^*).$$

Therefore, we conclude that $\boldsymbol{s}^* = (\boldsymbol{e}_1, \cdots, \boldsymbol{e}_1)$ is the unique NE of $\mathcal{G}$.

Next we estimate the welfare loss of $\boldsymbol{s}^*$ under any sequence $\{r_i\}_{i=1}^K$. First of all, note that for any $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in Loc(W)$ and any $2 \leq k \leq n$, if there exists $i \neq j$ such that $\boldsymbol{s}_i = \boldsymbol{s}_j = \boldsymbol{e}_k$, then there must be $k' \in [n]$ such that $\boldsymbol{e}_{k'} \notin \boldsymbol{s}_j$. In this case, W strictly increases if $\boldsymbol{s}_j$ changes to $\boldsymbol{e}_{k'}$. Therefore, for any $2 \leq k \leq n$, the number of elements in $\boldsymbol{s}$ that equal to $\boldsymbol{e}_k$ is either 0 or 1. Let the number of elements in $\boldsymbol{s}$ that equal to $\boldsymbol{e}_1$ be q. By definition,

$$W(\boldsymbol{s}) = (n+1)\sum_{i=1}^{\min(K,q)} r_i + (n-q)r_1, \tag{15}$$
$$W(\boldsymbol{s}^*) = (n+1)\sum_{i=1}^{K} r_i.$$

Since q maximizes the RHS of (15), we have $1 \leq q \leq K$ and $(n+1)r_{q+1} \leq r_1 \leq (n+1)r_q$. Therefore,

$$\begin{aligned}
\frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{W(\boldsymbol{s}^*)} &\geq \frac{\min_{\boldsymbol{s}\in Loc(W)} W(\boldsymbol{s})}{W(\boldsymbol{s}^*)} \\
&\geq \frac{(n+1)\sum_{i=1}^{\min(K,q)} r_i + (n-q)r_1}{(n+1)\sum_{i=1}^{K} r_i} \\
&\geq \frac{(n+1)\sum_{i=1}^{q} r_i + (n-q)r_1}{(n+1)\sum_{i=1}^{q} r_i + (K-q)r_1} \\
&= 1 + \frac{(n-K)r_1}{(n+1)\sum_{i=1}^{q} r_i + (K-q)r_1} \\
&\geq 1 + \frac{(n-K)r_1}{[(n+1)q + (K-q)]r_1} \\
&\to 1 - \frac{1+1/q}{1+nq/K} + \frac{1}{q}, n \to \infty.
\end{aligned}$$

Since $1 \leq q \leq K$, we conclude that $\frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{W(\boldsymbol{s}^*)} > 1 - O(\frac{1}{n}) + \frac{1}{K}$ when n is sufficiently large. And therefore we conclude that

$$\frac{W(\boldsymbol{s}^*)}{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})} \leq \frac{K}{K+1} + O\left(\frac{1}{n}\right).$$

□

A.4. Additional Discussion of BRM

Another notable special case within BRCM∈BRM is $M^{SM} = M[1, \frac{1}{2}, \cdots, \frac{1}{n}]$, which coincides with the Shapley mediator proposed in (Ben-Porat & Tennenholtz, 2018). One key feature of M^{SM} is that for any sequence $1 \geq \sigma_1 \geq \cdots \geq \sigma_n \geq 0$, it holds that $\sum_{i=1}^{n} M^{SM}(\sigma_i, \sigma_{-i}) = \sigma_1 \leq 1$. This implies that the platform can avoid providing explicit incentives and merely implement these rewards as matching probabilities. However, to do so, it must accommodate the possibility of not matching a user with any creators, corresponding to a probability of $1 - \sigma_1$. Furthermore, it does not support the top-K ranking strategy.

As pointed out by (Monderer & Shapley, 1996), every finite potential game is isomorphic to a congestion game. Furthermore, the definition of M as outlined in (6) can be interpreted as the utility that creator i acquires from the following congestion game:

1. The set of congestible elements are given by the continuum $E = \mathcal{X} \times [0, 1]$, where each element $(\boldsymbol{x}, t) \triangleq \boldsymbol{e} \in E$ corresponds to a user $\boldsymbol{x}$ with satisfaction level t.
2. The n players are n content creators.
3. Each creator's pure action $\boldsymbol{s}_i \in \mathcal{S}_i$ can be mapped to a subset of E in the following way: the action $\boldsymbol{s}_i$ determines the relevance score $\sigma(\boldsymbol{s}_i, \boldsymbol{x})$ over each $\boldsymbol{x} \in \mathcal{X}$, and then $\boldsymbol{s}_i$ is mapped to a subset $\{(\boldsymbol{x}, t) | \boldsymbol{x} \in \mathcal{X}, t \in [0, \sigma(\boldsymbol{s}_i, \boldsymbol{x})]\} \triangleq S_i \subseteq E$.
4. For each element $\boldsymbol{e}$ and a vector of strategies $(S_1, \cdots, S_n)$, the load of element $\boldsymbol{e}$ is defined as $x_{\boldsymbol{e}} = \#\{i : \boldsymbol{e} \in S_i\}$, i.e., the number of players who occupy $\boldsymbol{e}$.
5. For each element $\boldsymbol{e}$, there is a payoff function $d_{\boldsymbol{e}} : \mathbb{N} \to \mathbb{R}_{\geq 0}$ that only depends on the load of $\boldsymbol{e}$.
6. For any joint strategy $(S_1, \cdots, S_n)$, the utility of player i is given by $\sum_{\boldsymbol{e} \in S_i} d_{\boldsymbol{e}}(x_{\boldsymbol{e}})$, i.e., the sum of reward he/she collects from all occupied elements. For each occupied element $\boldsymbol{e}$, the reward is determined by its "congestion" level $x_{\boldsymbol{e}}$, which is characterized by the payoff function $d_{\boldsymbol{e}}$.

To better understand the constructed congestion game and the utility definition given in (6), we can consider each element in E (i.e., a user with a particular satisfaction level) as an atomic "resource". Each production strategy adopted by an individual creator can be thought of as occupying a subset of these resources. Given a fixed strategy profile, the load of $\boldsymbol{e} = (\boldsymbol{x}, t)$ is determined by the number of creators who achieve a relevance score exceeding t for user $\boldsymbol{x}$, thereby linking the ranking of each creator in the relevance score sequence for $\boldsymbol{x}$. Consequently, we can reformulate the utility for a creator who is ranked in the i-th position for user $\boldsymbol{x}$ as

$$\begin{aligned} \sum_{\boldsymbol{e} \in S_i} d_{\boldsymbol{e}}(x_{\boldsymbol{e}}) &= \sum_{t \in [0, \sigma(\boldsymbol{s}_i, \boldsymbol{x})]} d_t(x_{\boldsymbol{e}}) \\ &= \sum_{k=i}^{n} \sum_{t \in [\sigma(\boldsymbol{s}_{k+1}, \boldsymbol{x}), \sigma(\boldsymbol{s}_k, \boldsymbol{x})]} d_t(x_{\boldsymbol{e}}) \\ &= \sum_{k=i}^{n} \sum_{t \in [\sigma(\boldsymbol{s}_{k+1}, \boldsymbol{x}), \sigma(\boldsymbol{s}_k, \boldsymbol{x})]} d_t(k) \\ &\triangleq \sum_{k=i}^{n} \int_{\sigma_{k+1}}^{\sigma_k} f_k(t) dt. \end{aligned} \tag{16}$$

(16) holds because for any resource $\boldsymbol{e} = (\boldsymbol{x}, t)$ such that $t \in [\sigma(\boldsymbol{s}_{k+1}, \boldsymbol{x}), \sigma(\boldsymbol{s}_k, \boldsymbol{x})]$, the load of $\boldsymbol{e}$ is exactly given by k. As a result, by letting $f_k(t) = d_t(k)$, we recover the utility function defined in (6), where the value of function $f_i(t)$ at $t = t_0$ indicates the atomic reward for each creator if his/her strategy covers "resource" $(\boldsymbol{x}, t_0)$, given that there are exactly i creators occupy $(\boldsymbol{x}, t_0)$. This relationship also rationalizes why it is natural to assume that $f_1 \geq \cdots \geq f_n$: as an increase in competition for the same resource from multiple creators should correspondingly reduce the return that can be accrued from that resource.

A.5. Proof of Proposition 1

Proof. To prove that any $M \in$ BRM is merit-based, we need to verify the following by definition:

1. $M(0, \sigma_{-i}) = \int_0^0 f_n(t)dt = 0$, $M(1, \{0, \cdots, 0\}) = \int_0^1 f_1(t)dt > 0$.
2. $M(\sigma_i, \sigma_{-i}) - M(\sigma_j, \sigma_{-j}) = \sum_{k=j}^{i-1} \int_{\sigma_{k+1}}^{\sigma_k} f_k(t)dt \geq 0$.
3. for any $\{\sigma_j\}_{j=1}^n, \{\sigma'_j\}_{j=1}^n$ such that $\sigma_{-i} \preccurlyeq \sigma'_{-i}$, we can transform $\{\sigma_j\}_{j=1}^n$ to $\{\sigma'_j\}_{j=1}^n$ by taking finite steps of the following operations: 1. increase a certain value of $\sigma_j, j \neq i$ to $\tilde{\sigma}_j$ and it does not change the order of the current sequence; 2. increase a certain value of $\sigma_j, j \neq i$ to $\tilde{\sigma}_j$, and σ_i's ranking position decreases after this change. We will show that after each operation the value of $M(\sigma_i, \cdot)$ under the perturbed sequence does not increase.

   Let the perturbed sequence be $\tilde{\sigma}$. For the first type of operation, if $j < i$, we have $M(\sigma_i, \tilde{\sigma}_{-i}) = M(\sigma_i, \sigma_{-i})$. If $j > i$, we have

$$\begin{aligned} M(\sigma_i, \tilde{\sigma}_{-i}) - M(\sigma_i, \sigma_{-i}) &= \int_{\tilde{\sigma}_j}^{\sigma_{j-1}} f_{j-1}(t)dt + \int_{\sigma_{j+1}}^{\tilde{\sigma}_j} f_j(t)dt - \int_{\sigma_j}^{\sigma_{j-1}} f_{j-1}(t)dt - \int_{\sigma_{j+1}}^{\sigma_j} f_j(t)dt \\ &= \int_{\sigma_j}^{\tilde{\sigma}_j} (f_j - f_{j-1})(t)dt \leq 0. \end{aligned}$$

   For the second type of operation, with out loss of generality let's assume σ_{i+1} has increased to $\tilde{\sigma}_{i+1}$ such that $\sigma_i \leq \tilde{\sigma}_{i+1} \leq \sigma_{i-1}$. In this case we have

$$\begin{aligned} M(\sigma_i, \tilde{\sigma}_{-i}) - M(\sigma_i, \sigma_{-i}) &= \int_{\sigma_{i+2}}^{\sigma_i} f_{i+1}(t)dt - \int_{\sigma_{i+1}}^{\sigma_i} f_i(t)dt - \int_{\sigma_{i+2}}^{\sigma_{i+1}} f_{i+1}(t)dt \\ &= \int_{\sigma_{i+1}}^{\sigma_i} (f_{i+1} - f_i)(t)dt \leq 0. \end{aligned}$$

Therefore, M is merit-based. On the other hand, there exist instances in BRM that are not monotone. For example, if we let $f_1(t) = 1$ and $f_k(t) = 0, \forall k \geq 2$. Then we have

$$\begin{aligned} M(1, 0, 0, \cdots, 0) &= \int_0^1 f_1(t)dt > 0, \\ M(1, 1, 0, \cdots, 0) &= \int_0^0 f_1(t)dt + \int_0^1 f_2(t)dt = 0. \end{aligned}$$

As a result, $\pi(1, 0, 0, \cdots, 0) > 0 = \pi(1, 1, 0, \cdots, 0)$, which violates monotonicity. □

A.6. Proof of Theorem 2

Proof. For the first claim, consider the potential function of the following form:

$$P(\boldsymbol{s}) = \sum_{j=1}^m \sum_{i=1}^n \int_0^{\sigma_{l_j(i),j}} f_{i,j}(t)dt - \sum_{i=1}^n c_i(\boldsymbol{s}_i),$$

where $\sigma_{i,j} = \sigma(\boldsymbol{s}_i, \boldsymbol{x}_j)$ and $\sigma_{l_j(1),j} \geq \sigma_{l_j(2),j} \geq \cdots \geq \sigma_{l_j(n),j}$.

By the definition of potential games, we need to verify that for any set of functions $\{f_{i,j}\}$ and a strategy pair $\boldsymbol{s}_i, \boldsymbol{s}'_i \in \mathcal{S}_i$ for player-i, it holds that

$$u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i}) - u_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i}) = P(\boldsymbol{s}'_i, \boldsymbol{s}_{-i}) - P(\boldsymbol{s}_i, \boldsymbol{s}_{-i}). \tag{17}$$

For any user $j \in [m]$, let $\sigma_{i,j} = \sigma(\boldsymbol{s}_i, \boldsymbol{x}_j), \sigma'_{i,j} = \sigma(\boldsymbol{s}'_i, \boldsymbol{x}_j), \forall i \in [n]$. It suffices to show that

$$M(\sigma_{i,j}, \sigma_{-i,j}) - M(\sigma'_{i,j}, \sigma_{-i,j}) = \sum_{i=1}^n \int_0^{\sigma_{l_j(i),j}} f_{i,j}(t)dt - \sum_{i=1}^n \int_0^{\sigma'_{l_j(i),j}} f_{i,j}(t)dt, \tag{18}$$

Since the summation of (18) over j gives (17). With out loss of generality, we omit subscript j in Eq. (18) and assume $\sigma_1 \geq \cdots \geq \sigma_{i'} \geq \cdots \geq \sigma_i \geq \cdots \geq \sigma_n$. After player-$i$ changes her strategy from $\boldsymbol{s}_i$ to $\boldsymbol{s}_i'$, the relevance ranking increases from i to i', i.e., $\sigma_1 \geq \cdots \geq \sigma_{i'-1} \geq \sigma_i' \geq \sigma_{i'} \geq \cdots \geq \sigma_n$.

Therefore, we have

$$\text{LHS of (18)} = \int_{\sigma_{i'}}^{\sigma_i'} f_{i',j}(t)dt + \sum_{k=i'+1}^{n} \int_{\sigma_{k-1}}^{\sigma_k} f_{k,j}(t)dt, \tag{19}$$

$$\begin{aligned}\text{RHS of (18)} &= \sum_{k=1}^{n} \int_0^{\sigma_k} f_{k,j}(t)dt - \left(\sum_{k=1}^{i'-1} \int_0^{\sigma_k} f_{k,j}(t)dt + \int_0^{\sigma_i'} f_{i',j}(t)dt + \sum_{k=i'+1}^{n} \int_0^{\sigma_{k-1}} f_{k,j}(t)dt \right) \\ &= \int_0^{\sigma_{i'}} f_{i',j}(t)dt + \sum_{k=i'+1}^{n} \int_0^{\sigma_k} f_{k,j}(t)dt - \int_0^{\sigma_i'} f_{i',j}(t)dt - \sum_{k=i'+1}^{n} \int_0^{\sigma_{k-1}} f_{k,j}(t)dt \\ &= \int_{\sigma_{i'}}^{\sigma_i'} f_{i',j}(t)dt + \sum_{k=i'+1}^{n} \int_{\sigma_{k-1}}^{\sigma_k} f_{k,j}(t)dt.\end{aligned}$$

Hence, (18) holds for any j which completes the proof.

For the second claim, we can verify that when $f_{i,j} = r_i, \forall i, j$,

$$\begin{aligned}P(\boldsymbol{s}) &= \sum_{j=1}^{m} \sum_{i=1}^{n} \int_0^{\sigma_{l_j(i),j}} f_{i,j}(t)dt - \sum_{i=1}^{n} c_i(\boldsymbol{s}_i) \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} r_i \sigma_{l_j(i),j} - \sum_{i=1}^{n} c_i(\boldsymbol{s}_i) \\ &= \sum_{j=1}^{m} W_j(\boldsymbol{s}) - \sum_{i=1}^{n} c_i(\boldsymbol{s}_i) \\ &= W(\boldsymbol{s}).\end{aligned}$$

□

A.7. Proof of Corollary 1

Proof. We show that any TvN game instance $\mathcal{G}$ with $M = M[r_1, \cdots, r_K, 0, \cdots, 0] \in$ BRCM possesses a unique NE $\boldsymbol{s}^*$ which maximizes $W(\boldsymbol{s})$. From Theorem 2 we know that under M, $\mathcal{G}$ is a potential game and its potential function P is identical to its welfare function W. Therefore, any PNE of $\mathcal{G}$ belongs to $Loc(W)$. Next we show that all elements in $Loc(W)$ yield the same value of W, thus any PNE of $\mathcal{G}$ maximizes social welfare W.

First of all, note that for any $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in Loc(W)$ and any $2 \leq k \leq n$, if there exists $i \neq j$ such that $\boldsymbol{s}_i = \boldsymbol{s}_j = \boldsymbol{e}_k$, then there must exist $k' \in [n]$ such that $\boldsymbol{e}_{k'} \notin \boldsymbol{s}_j$. In this case, W strictly increases if $\boldsymbol{s}_j$ changes strategy to $\boldsymbol{e}_{k'}$. Therefore, for any $2 \leq k \leq n$, the number of elements in $\boldsymbol{s}$ that equal to $\boldsymbol{e}_k$ is either 0 or 1. Let the number of elements in $\boldsymbol{s}$ that equal to $\boldsymbol{e}_1$ be q. By definition, the welfare function writes

$$W(\boldsymbol{s}) = (n+1) \sum_{i=1}^{\min(K,q)} r_i + (n-q) r_1. \tag{20}$$

It is clear that the q that maximizes (20) satisfies $1 \leq q \leq K$ and $(n+1) r_{q+1} \leq r_1 \leq (n+1) r_q$, and all such q yields the same objective value of W. Therefore, we conclude that any PNE of $\mathcal{G}$ attains the optimal social welfare W. □

A.8. Content Creator Response Simulator

Algorithm 2 functions as follows: at each step, a random creator i selects a random improvement direction $\boldsymbol{g}_i$. If creator i discovers that adjusting her strategy in this direction yields a higher utility, she updates her strategy along $\boldsymbol{g}_i$; otherwise, she

Algorithm 2 (simStra) Simulate content creators' strategy evolving dynamic

Input: Time horizon T, learning rate η, utility function strategy set $(u_i(\boldsymbol{s}), \mathcal{S}_i)$ for each player, current mechanism $M[\boldsymbol{f}]$ parameterized by $\boldsymbol{f}$.
Initialization: Initial strategy profile $\boldsymbol{s}^{(0)} = (\boldsymbol{s}_1^{(0)}, \cdots, \boldsymbol{s}_n^{(0)})$.
for $t = 0$ **to** $T - 1$ **do**
  Generate $i \in [n]$ and $\boldsymbol{g}_i \in \mathbb{S}^{d-1}$ uniformly at random.
  if $u_i(\boldsymbol{s}_i^{(t)} + \eta \boldsymbol{g}_i, \boldsymbol{s}_{-i}^{(t)}) \geq u_i(\boldsymbol{s}^{(t)})$ **then**
    $\boldsymbol{s}_i^{(t+\frac{1}{2})} = \boldsymbol{s}_i^{(t)} + \eta \boldsymbol{g}_i$.
    Find $\boldsymbol{s}_i^{(t+1)}$ as the projection of $\boldsymbol{s}_i^{(t+\frac{1}{2})}$ in $\mathcal{S}_i$.
  else
    $\boldsymbol{s}_i^{(t+1)} = \boldsymbol{s}_i^{(t)}$
Output: $\boldsymbol{s}^{(T)}$.

retains her current strategy. This approach is designed to more closely mimic real-world scenarios where content creators may not have full access to their utility functions, but instead have to perceive them as black boxes. While they may aim to optimize their responses to the current incentive mechanism, identifying a new strategy that definitively increases their utilities can be challenging. Therefore, we model their strategy evolution as a trial-and-exploration process. We should note that the specifics of the simulator are not critical to our proposed solution: the optimizer can select any equilibrium-finding dynamic to replace our Algorithm 2, as long as it is believed to better represent creators' responses in reality.

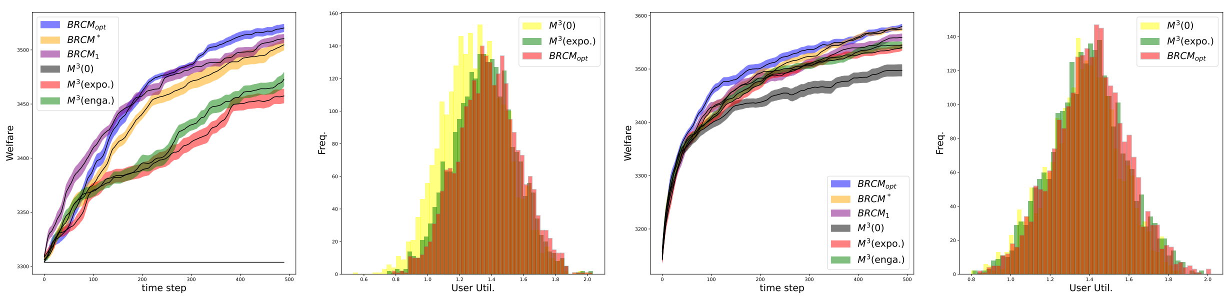
A.9. Additional Experiments

• **Experiments using MovieLens-1m** We use deep matrix factorization (Fan & Cheng, 2018) to train user and movie embeddings predicting movie ratings from 1 to 5 and use them to construct the user population $\mathcal{X}$ and creators' strategy set $\{\mathcal{S}_i\}$. The dataset contains 6040 users and 3883 movies in total, and the embedding dimension is set to $d = 32$. To validate the quality of the trained representation, we first performed a 5-fold cross-validation and obtain an averaged RMSE $= 0.739$ on the test sets, then train the user/item embeddings with the complete dataset.

To construct a more challenging environment for creators, we avoid using movies that are excessively popular and highly rated or users who are overly active and give high ratings to most movies. This ensures that the strategy of "producing popular content for the majority of active users" does not become a dominant strategy under any rewarding mechanism. Thus, we filtered out users and movies who have more than 500 predicted ratings higher than 4. After the filtering, we have $m = |\mathcal{X}| = 2550$ and $|\mathcal{S}_i| = 1783, \forall i \in [n]$. The remaining users are used as the user population $\mathcal{X}$, and remaining movies become the action set $\{\mathcal{S}_i\}$ for $n = 10$ creators. To normalize the relevance score to $[0, 1]$, we set $\sigma(\boldsymbol{s}, \boldsymbol{x}) = \text{clip}(\langle \boldsymbol{s}, \boldsymbol{x} \rangle / 2.5 - 1, 0, 1)$. $\{r_i\}_{i=1}^n$ is set to $\{\frac{1}{\log_2(2)}, \frac{1}{\log_2(3)}, \frac{1}{\log_2(4)}, \frac{1}{\log_2(5)}, \frac{1}{\log_2(6)}, 0, \cdots, 0\}$. We also consider two types of game instances, namely $\mathcal{G}_1$ and $\mathcal{G}_2$, as we elaborated on in Section 6.1. Specifically, in $\mathcal{G}_1$ creators' initial strategies are set to the most popular movie among all users (i.e., the movie that enjoys the highest average rating among $\mathcal{X}$) and the cost functions are set to be zero. In $\mathcal{G}_2$, we set creators' cost functions to $c_i = 10\|\boldsymbol{s}_i - \bar{\boldsymbol{s}}_i\|_2^2$ and let creator i start at the cost center $\bar{\boldsymbol{s}}_i$. $\{\bar{\boldsymbol{s}}_i\}_{i=1}^n$ are sampled at random from all the movies.

For each baseline in Section 6.2, we plot the welfare curve over $T = 500$ steps using Algorithm 2 and also the average user utility distribution at the end of simulations. The parameters of Algorithm 1 are set to $L_1 = 100, L_2 = 5, \eta_1 = 0.5, \eta_2 = 0.1, \boldsymbol{f}^{(0)} = (1, 1, 1, 1, 1, 0, \cdots, 0)$. The results are presented in Figure 2.

The new results obtained reinforce the findings presented in Section 6. In both the $\mathcal{G}_1$ and $\mathcal{G}_2$ environments, the BRCM family continues to outperform $\mathcal{M}^3$ overall. Specifically, BRCM$_{opt}$, BRCM$_1$, and BRCM* consistently demonstrate strong performance in social welfare, highlighting the robustness of BRCM across different environments. When creators initially adopt the most popular strategy in $\mathcal{G}_1$, $M^3(0)$ does not yield any improvement since no creator would change their strategy in such a situation under $M^3(0)$. In the case of $\mathcal{G}_2$, the advantage of BRCM over $\mathcal{M}^3$ diminishes slightly, which aligns with our observations from the synthetic dataset. The main reason is that the cost function discourages creators to deviate from their default strategies. Additionally, Figure 2(b) provides further evidence that the welfare gain achieved by BRCM arises from enhanced utility for a wider range of users.



(a) Welfare on $\mathcal{G}_1$ (b) User Util. Distr. on $\mathcal{G}_1$ (c) Welfare on $\mathcal{G}_2$ (d) User Util. Distr. on $\mathcal{G}_2$

Figure 2. Social welfare curve and average user utility distributions under two different environments. Error bars represent 0.2 standard deviation range, and they are generated from 10 independent runs. Game instances are generated from MovieLens-1m dataset.