Monetizing digital content with network effects: A mechanism-design approach*

Vincent Meisner[†] Pascal Pillath[‡]

Abstract

We design the profit-maximizing monetization scheme for a monopolistic digital content creator. She sells an excludable and non-rival good that is produced at a fixed cost. Users have heterogeneous private values that depend on how many users consume the content. Moreover, this audience size maps into additional creator profits, reflecting additional business opportunities. We show that the optimal allocation excludes low-type users, but may include otherwise unprofitable types that are solely included for their network value. We suggest an implementation in which users may voluntarily pay more than others or even subsidize them to make content provision to a larger audience more likely.

JEL-Classification: D82.
Keywords: Mechanism design, non-rival goods, network effects, digital content, creator economy.

*For useful comments and suggestions, we thank audiences at BSoE, UC3M, and the Lisbon GT Meeting, and, in particular, Helmut Bester, Tobit Gamp, Ángel Hernando-Veciana, Matthias Lang, Volker Nocke, Sebastian Schweighofer-Kodritsch, Johannes Schneider, and Roland Strausz. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged.

†Technical University Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany, vincent.meisner@tu-berlin.de.

‡Humboldt-Universität zu Berlin, pascal.pillath@hu-berlin.de.

1 Introduction

The "creator economy" is a system in which platforms such as Instagram, OnlyFans, Snapchat, Tiktok, Twitch or YouTube rely on content which is created by its users. The size of this market is estimated to be over $100 billion US dollars. While this industry has many examples of top earners, the vast majority of creators struggle to make a living from their content production.[1] Many companies have set up multi-billion dollar "creator funds" to attract and hold on to active content providers, and others have experimented with various features allowing creators to monetize their activity. In this paper, we contribute to the design of optimal monetization schemes for creators, who may only cater to niche audiences. Our model emphasizes the implications from three important features of this market: oligopsony, network effects, and the non-rivalry of digital content. We show that the optimal payment structure is fundamentally different from the simple pricing maximizing profit in large markets.

We isolate the content creator's profit maximization from many other economically relevant aspects of the digital creator economy. Specifically, we consider a monopolistic content creator as a mechanisms designer who faces a user population with heterogeneous values for her product. Digital content is produced at a one-time cost, and consumption is non-rival and excludable, i.e., once the content is produced it can be consumed by an unlimited number of users without extra cost, but access can be denied. Users draw a private value parameter, but their consumption value also depends on how many other users access content as well. In addition to this network effect on the user side, the creator's profit also features a network effect, e.g., by opening additional business opportunities such as sponsoring deals. We show that the profit-maximizing allocation can be implemented with a voluntary payment mechanism in which users can opt to pay more than others to make content provision more likely. A wide range of platforms such as OnlyFans, Substack, or Twitch indeed couple subscription fees with "tipping" or "donation" features. We also illustrate how positive network effects offer a self-interested rational for allowing "gifts" to other users, another only seemingly altruistic monetization element employed by, for instance, Twitch.

Besides modeling digital content as a club good, including network effects is integral both to capture the creator's objective and the consumers' value. Consumers value engagement with other consumers through comments, "likes," and chats such that the size of a creator's community enters their consumption value. Of course, we also allow for this network effect to be negative (congestion): a large audience may spark spam or come with a loss of a community feeling through reduced chances of directly engaging with the creator. The network effect on the creator's profit may reflect additional business opportunities that only emerge for popular creators with sufficiently large audiences. All creators in the "Top

[1]Some of the "Top Creators 2022" listed by Forbes (2022) such as Bhad Bhabie, FuckJerry, Jake Paul, or MrBeast made more than $30 million dollars that year, whereas the Economist (2021) suggests that more than 99% of content creators barely earn below minimum wage.

Creators 2022" list by Forbes (2022) make a significant fraction of their income through merchandise, advertisement deals or other partnerships that only arise through their fame. This feature of our model explains why, for instance, a blogger or podcaster may provide content for free in order to attract a large audience of which some then buy their book or their designed clothes.

The applied value of this paper lies in fitting the market for digital content creation into a mechanism-design framework to show that the commonly observed elements described above can indeed be part of an optimal monetisation scheme. Our technical contribution is to incorporate the network effect into a modified cost function, which allows us to characterize the optimal (dominant-strategy) implementable allocation with a straightforward algorithm. Although this allocation might exclude low-type users, it may include otherwise unprofitable types solely for their network value. Varying the direction and size of network effects allow us to nest several benchmark cases: the allocation of an indivisible private good (Myerson, 1981), which is essentially a non-rival good with very large negative network effects, and the allocation of a public good (Clarke, 1971; Groves, 1973), which is essentially a non-rival good with very large positive network effects such that it is only valued when it is consumed by all agents.

Cornelli (1996) considers our baseline model without the two network effects. Due to this relation, our indirect implementation through voluntary payment mechanisms also extends the payment scheme proposed in her paper. She considers a monopolistic mechanism designer who can produce a good at a fixed cost and zero marginal cost. By rewriting the creator's maximization problem, we essentially model the network effects as (possibly negative) costs. However, in contrast to Schmitz (1997), who extends the cost function of Cornelli (1996) to agent specific but constant costs, the "costs" in our setting depend on the audience size. The network effects on the user side feature in a type-dependent cost-benefit analysis. The similarities with these papers also connect our paper to the literature on crowdfunding (Belleflamme et al., 2015; Strausz, 2017; Ellman and Hurkens, 2019) and serial cost sharing (Moulin and Shenker, 1992; Moulin, 1994).

By modelling digital content as a club good (Buchanan, 1965), our work also relates to the mechanism-design literature on excludable public goods (Deb and Razzolini, 1999; Hellwig, 2003, 2005; Norman, 2004; Hellwig, 2007; Bierbrauer, 2011), where the goal is efficient provision rather than profit maximization. Birulin (2006) considers public goods with congestion, but, in contrast to us, models the congestion as a capacity constraint rather than directly in the agents' payoff function.

In contrast to other papers, we model the network externality by making agents' consumption value dependent on the number of other consumers. Mechanism design with allocation externalities was also studied by Jehiel et al. (1996), but they are concerned with the externality on agents who did not acquire the good rather than joint consumption. The externality of the network good in Csorba (2008) or Ostrizek and Sartori (2023) depends on the total production of the good in the economy. The consumption value in Segal (1999) depends on other agents' trades. Imas and Madarász (2021) provide evidence that consumers' valuations for the consumption of a good can be increasing in others' unmet desires. That is, all else equal the willingness-to-pay increases when other consumers are excluded from the market.

2 Model

Players and outcomes: A monopolistic content creator (mechanism designer) can produce a non-rival but excludable good (her digital content) at cost c. She faces N users (consumers) $i \in \mathcal{N}$, and she designs an arbitrary (finite) game that determines the outcomes, i.e., who has access to her content and the payments of all players. Formally, an outcome $o = (q_i, m_i)_{i \in \mathcal{N}}$ determines for each user i whether he can consume the good, $q_i \in \{0, 1\}$, and his payment $m_i \in \mathbb{R}$.

Types: User i privately learns his value type θ_i, an iid draw from a commonly known distribution with cdf F, continuous density f, and support $\Theta := [0, \overline{\theta}]$. The type profile is denoted by $\boldsymbol{\theta} := (\theta_i)_{i \in \mathcal{N}} \in \Theta^N$, and we sometimes use the notation $\boldsymbol{\theta} = (\theta_i, \boldsymbol{\theta}_{-i})$. By independence, the joint distribution of $\boldsymbol{\theta}$ is given by $G(\boldsymbol{\theta}) = \Pi_{j \in \mathcal{N}} F(\theta_j)$, and we define $G_{-i}(\boldsymbol{\theta}_{-i})$ analogously.

Payoffs: We call the subset of users that can consume the good the audience, $A = \{i : q_i = 1\} \subseteq \mathcal{N}$. Given an audience of size $|A| = k$, user i's valuation for the good is $v(\theta_i, k)$. We allow for positive, negative and also non-monotone user network effects, i.e., v can be increasing, decreasing or non-monotone in audience size k, but we assume the network effects go in the same direction for all types, i.e., $\text{sign}(v(\theta_i, k) - v(\theta_i, j)) = \text{sign}(v(\theta_i', k) - v(\theta_i', j))$ for all $j, k, \theta_i, \theta_i'$. Moreover, v is continuous and increasing in θ_i, ${}^{\partial v(\theta_i,k)}\!/\!_{\partial \theta_i} > 0$ for all θ_i and k, and we impose the following single-crossing assumption for all $x > y, k > k'$,

$$\text{sign}\{(v(x,k) - v(x,k')) - (v(y,k) - v(y,k'))\} = \text{sign}\{v(x,k) - v(x,k')\}. \quad \text{(SC)}$$

It says that the effect of a change in the audience size on the marginal utility in terms of types goes in the same direction as the network effect. That is, for positive network effects, larger types benefit more from an audience increase; for negative network effects, larger types lose more from an audience increase.

Given $k-1$ other users access the good, the utility of user i is given by

$$q_i v(\theta_i, k) - m_i.$$

The mechanism designer maximizes her profit, and she only incurs cost c when the content is provided, i.e., when she accepts at least one user, $k > 0$. For a given outcome, the creator's profit is

$$\sum_{i=1}^{N} m_i - c\mathbb{1}_{k>0} + \varphi(k), \tag{1}$$

where function $\varphi : \{0, 1, \ldots, N\} \to \mathbb{R}$ maps an audience size into creator network effects, an effect on creator profit that can be positive or negative depending on the application. Whereas v measures the network effect on users' valuations of the good, φ measures the network effect on the creator's profit.

Game: The content creator sets up an arbitrary (finite) game in which each agent has an action (plan) $\alpha_i \in \mathcal{A}_i$, and an outcome function maps all actions into a deterministic outcome $o \in \mathcal{O}$, $g : \mathcal{A} \to \mathcal{O}$ with $\mathcal{A} := (\mathcal{A}_i)_{i \in \mathcal{N}}$, where $g_i(\boldsymbol{\alpha}) \in \{0, 1\} \times \mathbb{R}$ is the final allocation decision and payment of user i given all players' actions $\boldsymbol{\alpha}$. Moreover, we allow users to walk away from the outcome such that each user must receive at least his outside option, which we normalize to zero. By the revelation principle,[2] we can restrict attention to incentive-compatible direct revelation mechanisms (DRM) in our quest to find the optimal allocation that can be implemented by any equilibrium in dominant strategies of any game designed by the creator. Because direct mechanisms are barely used in practice, we propose an indirect mechanism that implements this allocation, but in Bayesian Nash equilibrium instead of dominant strategies.

DRM: In a DRM, each user i reports his type, and functions $\langle q, m \rangle = (q_i, m_i)_{i \in \mathcal{N}}$ determine the outcome for each combination of types, $q_i : \Theta^N \to \{0, 1\}$ and $m_i : \Theta^N \to \mathbb{R}$. For any deterministic DRM, we can define the audience size as $k(\boldsymbol{\theta}) := \sum_{i \in \mathcal{N}} q_i(\boldsymbol{\theta})$.

Given the other users reported types $\widehat{\boldsymbol{\theta}}_{-i}$, the payoff of a user of type θ_i who reported $\widehat{\theta}_i$ to the DRM is

$$u_i(\widehat{\theta}_i, \widehat{\boldsymbol{\theta}}_{-i} | \theta_i) = q_i(\widehat{\theta}_i, \widehat{\boldsymbol{\theta}}_{-i}) v(\theta_i, k(\theta_i, \widehat{\boldsymbol{\theta}}_{-i})) - m_i(\widehat{\theta}_i, \widehat{\boldsymbol{\theta}}_{-i}).$$

Note that user i's utility depends on the final allocation through the audience size, which depends on the other users' reported types, not their true types. That is, we are still in a private-value setting. Moreover, for fixed $\widehat{\boldsymbol{\theta}}_{-i}$ and a given function q, user i's report fixes the allocation (and hence k) deterministically. The usual incentive and participation constraints when maximizing (1) can be expressed as

$$U_i(\theta_i, \widehat{\boldsymbol{\theta}}_{-i}) = u_i(\theta_i, \widehat{\boldsymbol{\theta}}_{-i} | \theta_i) \geq u_i(\widehat{\theta}_i, \widehat{\boldsymbol{\theta}}_{-i} | \theta_i) \qquad \forall i, \theta_i, \widehat{\theta}_i, \widehat{\boldsymbol{\theta}}_{-i} \qquad \text{(IC)}$$

$$u_i(\theta_i, \widehat{\boldsymbol{\theta}}_{-i} | \theta_i) \geq 0 \qquad \forall i, \theta_i, \widehat{\boldsymbol{\theta}}_{-i}. \qquad \text{(IR)}$$

In contrast to a rival-goods problem, we do not have a restriction $\sum_i q_i(\boldsymbol{\theta}) \leq 1$ because the good can be consumed by all users at the same time such that $q_i(\boldsymbol{\theta}) \in \{0, 1\}$ for all $\boldsymbol{\theta}$ and i is the only feasibility constraint of our deterministic mechanism.

Implementability: The constraint (IC) implies that q_i must be weakly increasing in type θ_i for all $\boldsymbol{\theta}_{-i}$ whereas the transfers are pinned down by the familiar integral form (4) below. Consequently, (IC) implies (IR) if the lowest type gets at least utility zero for all $\boldsymbol{\theta}_{-i}$. The following results are helpful to rewrite our problem.

[2]The revelation principle may not hold when restricting attention to deterministic mechanisms. However, this is not an issue in our setting with ex-post constraints. See Jarman and Meisner (2017).

Lemma 1. *A direct mechanism $\langle q, m\rangle$ is incentive compatible if and only if for every i and every $\boldsymbol{\theta}_{-i}$,*

(i) there is a type $\underline{x}(\boldsymbol{\theta}_{-i})$ such that for all $\theta_i > \underline{x}(\boldsymbol{\theta}_{-i}) > \theta_i'$:

$$q_i(\theta_i, \boldsymbol{\theta}_{-i}) = 1 \text{ and } q_i(\theta_i', \boldsymbol{\theta}_{-i}) = 0; \tag{2}$$

(ii) for all $\theta_i > \theta_i'' > \underline{x}(\boldsymbol{\theta}_{-i})$,

$$v(\theta_i, k(\theta_i, \boldsymbol{\theta}_{-i})) - v(\theta_i'', k(\theta_i, \boldsymbol{\theta}_{-i})) \geq v(\theta_i, k(\theta_i'', \boldsymbol{\theta}_{-i})) - v(\theta_i'', k(\theta_i'', \boldsymbol{\theta}_{-i})); \tag{3}$$

(iii) for all $\theta_i > \underline{x}(\boldsymbol{\theta}_{-i}) > \theta_i'$:

$$\begin{aligned} m_i(\theta_i', \boldsymbol{\theta}_{-i}) &= m(0, \boldsymbol{\theta}_{-i}), \\ m_i(\theta_i, \boldsymbol{\theta}_{-i}) &= m(0, \boldsymbol{\theta}_{-i}) + v(\theta_i, \boldsymbol{\theta}_{-i}) - \int_{\underline{x}(\boldsymbol{\theta}_{-i})}^{\theta_i} \frac{\partial v(t, k(t, \boldsymbol{\theta}_{-i}))}{\partial t} dt \end{aligned} \tag{4}$$

The above lemma gives familiar necessary and sufficient conditions (2) and (4) for incentive compatibility. The lemma below tells us that (3) and (SC) together imply that higher types who get the good must get a weakly "better" audience size.

Lemma 2. *Suppose* (SC) *holds. A direct mechanism $\langle q, m\rangle$ is incentive-compatible if and only if audience sizes are ordered $k_1, \ldots, k_n$ such that $v(\theta_i, k_j) \geq v(\theta_i, k_{j+1})$ for all θ_i, and given any $\boldsymbol{\theta}_{-i}$ the interval $(\underline{x}(\boldsymbol{\theta}_{-i}), \overline{\theta}]$ is partitioned by cutoffs $x_1(\boldsymbol{\theta}_{-i}) \geq \cdots \geq x_n(\boldsymbol{\theta}_{-i})$ such that $k(\theta_i, \boldsymbol{\theta}_{-i}) = j$ if $\theta_i \in (x_j(\boldsymbol{\theta}_{-i}), x_{j+1}(\boldsymbol{\theta}_{-i})]$.*

Rewriting the payoffs: Exploiting the integral form (4), we can rewrite the designer's objective as

$$\int_{\Theta^N} \left(\sum_{i=1}^{N} \psi(\theta_i, k(\boldsymbol{\theta})) q_i(\boldsymbol{\theta}) - c\mathbb{1}_{k(\boldsymbol{\theta})>0} + \varphi(k(\boldsymbol{\theta})) \right) dG(\boldsymbol{\theta}), \tag{5}$$

where the virtual value of type θ_i in an audience of size k is given by

$$\psi(\theta_i, k) = v(\theta_i, k) - \frac{1 - F(\theta_i)}{f(\theta_i)} \frac{\partial v(\theta_i, k)}{\partial \theta_i}, \tag{6}$$

where the latter part reflects the information rents needed to incentivize truthful type revelation. We assume it is strictly increasing in θ_i, and, in line with Myerson (1981), we call such environments regular. Because ${}^{\partial v(\theta_i,k)}/_{\partial \theta_i} \geq 0$ for all k and θ_i, the standard monotone hazard rate condition combined with ${}^{\partial^2 v(\theta_i,k)}/_{\partial \theta_i^2} \leq 0$ is sufficient for regularity.

3 Analysis

Road map: In the usual fashion, we approach our problem by first considering a relaxed problem, i.e., we maximize (5) without the constraints (2) and (3) implied by (IC). Next, we verify that our solution indeed satisfies these constraints and, hence, it also solves our original (more constrained) problem. Finally, we discuss Bayesian indirect implementations of our optimal allocation to rationalize commonly seen elements in real-life monetisation schemes.

3.1 The relaxed problem

Our first step towards the optimal allocation of the relaxed problem is a characterization of the type profiles under which the content can be provided profitably. The second step clarifies the structure of the audience in these cases for any audience size, while the third step provides an algorithm that finds the optimal audience size, establishing the optimal allocation for all type profiles. Finally, we find conditions under which the optimal allocation rule is simple.

1. Content provision: In the relaxed problem, we maximize

$$\sum_{i=1}^{N} \psi(\theta_i, k(\boldsymbol{\theta}))q_i(\boldsymbol{\theta}) - c\mathbb{1}_{k(\boldsymbol{\theta})>0} + \varphi(k(\boldsymbol{\theta})) \tag{7}$$

separately for all possible type profiles $\boldsymbol{\theta}$ and disregard the constraints (2) and (3) of Lemma 1. Because the problem is linear in q_i, it follows immediately that in the relaxed problem our restriction to deterministic mechanisms is without loss, $q_i(\boldsymbol{\theta}) \in \{0, 1\}$ for all $\boldsymbol{\theta}$.

Given a fixed type profile, the creator prefers to provide the good to an audience consisting of users in set J (left-hand side payoff) over not providing it at all (right-hand side payoff) if

$$\sum_{i \in J} \psi(\theta_i, |J|) - c + \varphi(|J|) \geq 0 + \varphi(0),$$

which we can rearrange to

$$\Psi(\boldsymbol{\theta}|J) \geq C(|J|)$$
$$\text{with } \Psi(\boldsymbol{\theta}|J) := \sum_{i \in J} \psi(\theta_i, |J|) \quad \text{and} \quad C(k) := c - (\varphi(k) - \varphi(0)). \tag{8}$$

This inequality simply expresses that providing the content is only profitable if the revenue extractable from audience J exceeds the network effect adjusted cost for an audience of size $|J|$. The extractable revenue is the sum of virtual values of admitted types, and $C(k)$ adjusts the total costs for creator network effects. For instance, a negative cost adjustment reflects positive creator network effects. By assumption, the cost adjustment $(\varphi(k) - \varphi(0))$ only depends on the size of the audience and not its composition.

Lemma 3. *Consider the relaxed problem and take any $\boldsymbol{\theta}$. In optimum, the good is provided if and only if a set $J \subseteq \mathcal{N}$ exists such that*

$$\Psi(\boldsymbol{\theta}|J) \geq C(|J|). \tag{8}$$

The result above gives a necessary and sufficient condition for content provision, but it does not delineate to whom the good shall be provided. The proof follows straightforwardly from the type-by-type creator profit function (7).

2. Audience structure and candidate sets: Given some type profile $\boldsymbol{\theta}$, the content creator prefers to entertain audience J (payoff on the left-hand side) instead of audience J' (payoff on the right-hand side) if

$$\sum_{j \in J} \psi(\theta_j, |J|) - c + \varphi(|J|) \geq \sum_{j \in J'} \psi(\theta_j, |J'|) - c + \varphi(|J'|),$$

which can be rearranged to

$$\Psi(\boldsymbol{\theta}|J) - \Psi(\boldsymbol{\theta}|J') \geq \varphi(|J'|) - \varphi(|J|). \tag{9}$$

Without loss of generality, let us relabel users in order of their (virtual) types, $\theta_i \geq \theta_{i+1}$, where regularity implies $\psi(\theta_i, k) \geq \psi(\theta_{i+1}, k)$ for all k. Since $(\varphi(k) - \varphi(k)) = 0$ for all $k \in \mathbb{N}$, we can immediately infer from (9) that out of all possible audiences of the same size k, the seller prefers $J_k = \{j : j \leq k\}$ the most. That is, in the relaxed problem, every optimal allocation that accepts k users must accept the k users with the highest (virtual) types. Hence, we can restrict attention to such allocation sets, and we only need to find the optimal audience size k^* for all type realizations $\boldsymbol{\theta}$. Let us call a set with this structure a candidate set. That is, J_k is a candidate set if and only if

$$J_k(\boldsymbol{\theta}) = \{j \in \mathcal{N} : j \leq k \text{ and } \theta_i \geq \theta_{i+1} \, \forall i\}. \tag{10}$$

Lemma 4. *Consider the relaxed problem and take any $\boldsymbol{\theta}$. If the good is produced, the optimal audience is a candidate set.*

This result follows from the objective (7) combined with the insights above. Having established the structure of the optimal audience for any size k, we can finalize the characterization of the optimal allocation by determining the optimal audience size for all type vectors $\boldsymbol{\theta}$.

3. Audience size and general algorithm: For any optimal audience J_k, it must be unprofitable a) to add the lowest-index users $i \notin J_k$ and b) to remove the highest-index users $j \in J$. If candidate set J_k is the optimal audience, (9) implies

$$\begin{aligned} &a)\Psi(\boldsymbol{\theta}|J_k) - \Psi(\boldsymbol{\theta}|J_{k+j}) &&\geq \varphi(k+j) - \varphi(k) && \forall j \in \{1, \dots, N-k\}, \text{ and} \\ &b)\Psi(\boldsymbol{\theta}|J_k) - \Psi(\boldsymbol{\theta}|J_{k-j}) &&\geq \varphi(k-j) - \varphi(k) && \forall j \in \{1, \dots, k\}. \end{aligned} \tag{11}$$

In general, these are many constraints to keep track of, and we essentially define an algorithm that compares the revenue of all candidate sets, in keeping with Lemma 4. Further assumptions on the network effects can reduce the complexity of our algorithm considerably.

Let $J \succeq_{\boldsymbol{\theta}} J'$ express that, for a given type vector $\boldsymbol{\theta}$, the creator weakly prefers audience J over audience J', i.e., it expresses that (9) holds. Next, we define thresholds on the additional revenue when adding the next j highest-value users $\{k+1, \ldots, k+j\}$ to audience J_k,

$$\gamma(k, k+j, \boldsymbol{\theta}_{\leq k}) := \overbrace{\varphi(k) - \varphi(k+j)}^{=C(k+j)-C(k)} + \sum_{i=1}^{k} [\psi(\theta_i, k) - \psi(\theta_i, k+j)] \tag{12}$$

with $\boldsymbol{\theta}_{\leq k} = (\theta_1, \ldots, \theta_k)$. This threshold reflects two changes in the creator's profit that the revenue extractable from the j additional users needs to compensate. First, the creator network effect changes by $(\varphi(k) - \varphi(k+j))$. This first part can be used to represent the variable part of the total adjusted total cost, $C(k) = c + \gamma(0, k, \cdot)$ for all k. Second, each user $i \leq k$ already admitted to the audience now garners value $v(\theta_i, k+j) \neq v(\theta_i, k)$. That is, due to the network effects on the users' side the creator can extract either more or less value from the users J_k already tentatively considered for her audience. While the sign of this effect only depends on whether we assume positive or negative network effects, its size also depends on $\boldsymbol{\theta}_{\leq k}$.

That is, for any candidate set J_k, we have

$$J_{k+j} \succeq_{\boldsymbol{\theta}} J_k \iff \sum_{i=1}^{j} \psi(\theta_{k+i}, k+j) \geq \gamma(k, k+j, \boldsymbol{\theta}_{\leq k}). \tag{13}$$

Without further assumptions on the form of v in k, it may be possible that $\psi(\theta_{k+1}, k+1) < \gamma(k, k+1, \boldsymbol{\theta}_{\leq k})$, but $\psi(\theta_{k+1}, k+2) + \psi(\theta_{k+2}, k+2) > \gamma(k, k+2, \boldsymbol{\theta}_{\leq k})$. That is, it can be unprofitable to add user $(k+1)$ to the audience alone, while it is profitable to add user $(k+1)$ and $(k+2)$ together. Similarly, it might be profitable to add a single user to the audience, but even more profitable to remove several users. Consequently, we cannot generally restrict attention to "local" (one-by-one) changes in our quest to find the optimal allocation. Nevertheless, once we have defined the general algorithm to find the optimal allocation, we can specify conditions under which a greedy algorithm that only considers such local changes succeeds as well.

General algorithm: To state the general algorithm that finds the optimal allocation in the relaxed problem, we define the operator

$$\overline{\max}_{\boldsymbol{\theta}}\{J, K\} = \begin{cases} J & \text{if } J \succeq_{\boldsymbol{\theta}} K, \\ K & \text{otherwise.} \end{cases}$$

Now, we, we define the direct mechanism $\langle \overline{q}, \overline{m} \rangle$ by the following algorithm. In each step k, this algorithm compares the step's candidate set J_k to the step's comparison set K_k. The preferred of the two sets becomes the comparison set in the next step. The final comparison set is the audience that generates the maximal revenue, and the algorithm's final step is to a) verify whether this maximal revenue covers the adjusted total cost and to b) use Myerson's revenue-equivalence formula (4) to determine the payments.

Definition 1. *$\langle \overline{q}, \overline{m} \rangle$ is defined by the following steps for each $\boldsymbol{\theta}$.*

step 0: *Set $J_0(\boldsymbol{\theta}) = K_1(\boldsymbol{\theta}) = \varnothing$.*

step 1: *Set $J_1(\boldsymbol{\theta}) = \{1\}$, and set $K_2(\boldsymbol{\theta}) = \overline{\max}_{\boldsymbol{\theta}}\{J_1(\boldsymbol{\theta}), K_1(\boldsymbol{\theta})\}$ ...*

step k: *Set $J_k(\boldsymbol{\theta}) = \{1, 2, \ldots, k\}$, and set $K_{k+1}(\boldsymbol{\theta}) = \overline{\max}_{\boldsymbol{\theta}}\{J_k(\boldsymbol{\theta}), K_k(\boldsymbol{\theta})\}$ ...*

Stop *The most profitable allocation set is $K_{N+1}(\boldsymbol{\theta})$ with size $\overline{k}(\boldsymbol{\theta}) = |K_{N+1}(\boldsymbol{\theta})|$. Set for all $i \in \mathcal{N}$*

$$\overline{q}_i(\boldsymbol{\theta}) = \mathbb{1}_{i \in \overline{K}_{N+1}(\boldsymbol{\theta}) \text{ and } \Psi(\boldsymbol{\theta}|K_{N+1}(\boldsymbol{\theta})) \geq C(\overline{k}(\boldsymbol{\theta}))}. \tag{14}$$

That is, the optimal audience only consists of users $K_{N+1}(\boldsymbol{\theta})$, and the good is provided if and only if the adjusted costs are covered. Set for all $i \in \mathcal{N}$

$$\overline{m}_i(\boldsymbol{\theta}) = v(\theta_i, \overline{k}(\boldsymbol{\theta}))\, \overline{q}_i(\boldsymbol{\theta}) - \int_0^{\theta_i} \overline{q}_i(x, \boldsymbol{\theta}_{-i}) \frac{\partial v(x, \overline{k}(\boldsymbol{\theta}))}{\partial x} dx. \tag{15}$$

This algorithm indeed solves our relaxed problem.

Lemma 5. *$\langle \overline{q}, \overline{m} \rangle$ as defined above is the solution to the relaxed problem.*

The algorithm finds the audience that allows to extract the maximal revenue by essentially comparing all candidate sets. Step N of the algorithm only guarantees that the revenue extractable from the concluding set K_{N+1} covers the variable part of the adjusted cost $C(\overline{k})$, $\Psi(\boldsymbol{\theta}|K_{N+1}) \geq \varphi(\overline{k}) - \varphi(0)$. Hence, an extra step is necessary to consider the fixed part c. The algorithm also makes sure that if the total adjusted costs are not covered for set K_{N+1}, they are not covered for any other audience such that the algorithm's provision decision is indeed final.

Simple allocation rule: Our algorithm $\overline{q}$ above solves the relaxed problem by comparing all candidate sets. However, the creator or a user may not find it straightforward to compute the allocation even for a given $\boldsymbol{\theta}$, but an understanding of the allocation mechanism (at least intuitively) is crucial for actual equilibrium play to ensue. We now consider the simpler allocation rule $\widehat{q}$, and give a necessary and sufficient condition such that $\widehat{q}$ and $\overline{q}$ are equivalent. In other words, under this condition the considerably less complex rule $\widehat{q}$ solves the relaxed problem. The transfer rule $\widehat{m}$ again just follows from the integral form (4).

Definition 2. *Allocation rule $\widehat{q}$ is defined as follows for all $\boldsymbol{\theta}$:*

$$\begin{gathered} \widehat{q}_i(\theta_i) = \mathbb{1}_{i \leq \widehat{k}(\boldsymbol{\theta}) \text{ and } \Psi(\boldsymbol{\theta}|J_{\widehat{k}(\boldsymbol{\theta})} \geq C(\widehat{k}(\boldsymbol{\theta}))} \\ \text{with } \widehat{k}(\boldsymbol{\theta}) = \arg\max_k \psi(\theta_k, k) \geq \gamma(k-1, k, \boldsymbol{\theta}_{\leq k-1}). \end{gathered} \tag{16}$$

For a given $\boldsymbol{\theta}$, allocation rule $\widehat{q}$ simply identifies the highest-index $\widehat{k}$ such that the virtual type exceeds the local cutoff for inclusion $\gamma(\widehat{k}-1, \widehat{k}, \boldsymbol{\theta}_{\leq k-1})$, and then it accepts all users $k \leq \widehat{k}$, while rejecting all others. That is, conditional on the provision of the content, the k-th highest type is admitted to the audience if and only if

$$\psi(\theta_k, k) \geq \gamma(k-1, k, \boldsymbol{\theta}_{\leq k-1}) \quad \iff \theta_k \geq x_k(\boldsymbol{\theta}_{-k}) := \psi^{-1}(\gamma(k-1, k, \boldsymbol{\theta}_{\leq k-1}), k), \tag{17}$$

where $\psi^{-1}(\cdot, k)$ is the inverse of function $\psi(\cdot, k)$ with respect to the first argument. This inverse exists because all virtual values for each k are strictly increasing by our regularity assumption.

This insight makes the allocation rule easy to understand for users. User i knows his own valuation parameter θ_i and knows that the other users types $\boldsymbol{\theta}_{-k}$ fix a sequence of cutoffs $(x_k)_{k\in\{1,\ldots,N\}}$. Any user i then understands that if he was the k-th highest type, he is conditional on content provision part of the audience if and only if $\theta_i \geq x_k(\boldsymbol{\theta}_{-i})$.

Before we give the condition under which $\overline{q}$ and $\widehat{q}$ are equivalent, we observe that it is always the case that any optimal mechanism rejects all users $i > \widehat{k}(\boldsymbol{\theta})$. That is, the algorithm of Definition 1 can essentially stop at step $\widehat{k}(\boldsymbol{\theta})$ without loss of generality for all $\boldsymbol{\theta}$. By definition of $\widehat{k}$, no type θ_i with $i > \widehat{k}$ exceeds its local cutoff, which implies that adding any number of users $i > \widehat{k}$ only reduces profit.

Lemma 6. *For any $\boldsymbol{\theta}$, $\overline{q}_i(\boldsymbol{\theta}) = 0$ for all $i > \widehat{k}(\boldsymbol{\theta})$.*

Given the good is provided, allocation rules $\overline{q}$ and $\widehat{q}$ are only equivalent when, in addition to the insight above, it is also not profitable to drop users from audience $J_{\widehat{k}(\boldsymbol{\theta})}$. The result below gives a necessary and sufficient condition for this case.

Lemma 7. *Allocation rule $\widehat{q}$ solves the relaxed problem if and only if*

$$j\,\gamma(\widehat{k}(\boldsymbol{\theta})-1, \widehat{k}(\boldsymbol{\theta}), \boldsymbol{\theta}_{\leq \widehat{k}(\boldsymbol{\theta})-1}) \geq \gamma(\widehat{k}(\boldsymbol{\theta})-j, \widehat{k}(\boldsymbol{\theta}), \boldsymbol{\theta}_{\leq \widehat{k}(\boldsymbol{\theta})-j}) \quad \forall j, \boldsymbol{\theta}. \tag{18}$$

Sufficient condition for an optimal simple rule: Condition (18) nests the special case of no network effects at all, when $\varphi(k) = 0$ and $\psi(\theta_i, k) = \widetilde{\psi}(\theta_i)$ for all k and θ_i. Here, $\gamma(k, j, \boldsymbol{\theta}_{\leq k}) = 0$ and $C(k) = c$ for all $k, j, \boldsymbol{\theta}$ such that the simple allocation rule of Cornelli (1996) is optimal: accept all types with non-negative virtual values and provide the good if the sum of these virtual values exceeds c.

Intuitively, (18) says that the inclusion conditions for lower (higher-index) types at a later stages of algorithm $\overline{q}$ can only become "harder." Hence, when adding user $\widehat{k}$ to audience $J_{\widehat{k}-1}$ is profitable it must also be profitable to add user $\widehat{k}$ together with users with higher (lower-index) types to a smaller candidate set. A sufficient condition for (18) is that both φ and ψ are concave in audience size. That is, as long as these functions have an inverted u-shape, they can also be non-monotone.

Taking stock, we have solved the relaxed problem in general, and we have provided a condition on the network effects such that the solution to the relaxed problem is simple. Now, we verify that our solution indeed satisfies the constraints which the relaxed problem ignores, and, as a result, the solutions to the relaxed problem and the constrained problem coincide.

3.2 The full constrained problem

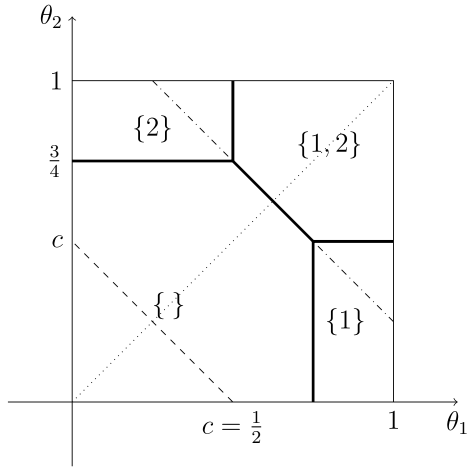
Given our regularity assumption on virtual valuations (6) and our single-crossing assumption (SC), the optimal allocation in the relaxed problem is indeed incentive compatible and individually rational. Hence, it also solves the more constrained problem. First, the algorithm behind $\overline{q}$ admits users to the audience in order of their virtual types, which under regularity coincides with the order of types. Therefore, any admitted user remains in the audience if his type is increased. Second, (SC) ensures that whenever two types obtain the good with different audience sizes, the larger type gets the audience size he prefers.

Proposition 1. *In regular environments given* (SC), $\langle \overline{q}, \overline{m} \rangle$ *of Definition 1 is the solution to the full constrained problem.*

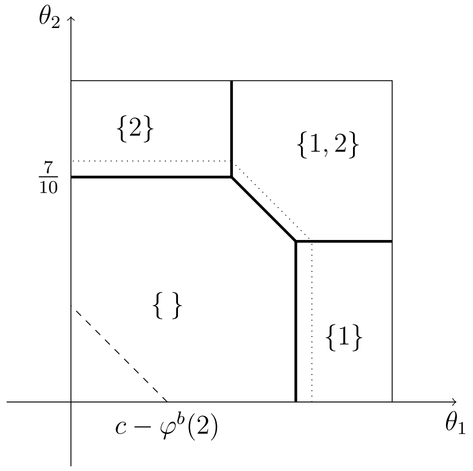
Discussion of non-regular environments: The optimality of $\overline{q}$ in the full constrained problem hinges on regularity in a fashion similar to the classical result by Myerson (1981). With non-monotone virtual values, $\overline{q}$ would violate the monotonicity constraint. As expected, the solution to this problem is to apply the classical ironing techniques applied separately for every audience size k and to then run the algorithm with with ironed virtual values. Ironing (or, alternatively, called bunching) implies that the types in the ironed region get the same contract $(\widetilde{q}, \widetilde{m})$. While this insight implies that the optimal allocation rule is stochastic in Myerson's model with a single rival good, this is not necessarily true in our model. For instance, when all network effects are weakly positive, the creator always wants to add all types in the bunching region whenever she wants to add one of them. That is, restricting attention to deterministic allocation rules is without loss. However, this is clearly not true when negative network effects are allowed. The most obvious example is Myerson's setting with a single scarce good, which is nested by our model when $v(\theta_i, 1) = \theta_i$ and $v(\theta_i, k) < 0$ for all θ_i and all $k > 1$.

Discussion of the optimal allocation against benchmarks: Figure 1 juxtaposes the optimal allocations in four exemplary settings with $N = 2$ users, uniformly distributed types on $[0, 1]$, and cost $c = 1/2$. Step by step and starting from a setting without any network effects (Cornelli, 1996), we add a component of our model in each panel. In Figure 2, we focus on user network effects, and we discuss the intermediate cases between the benchmark of Myerson (1981) (strong negative user network effects) and public goods (strong positive user network effects).

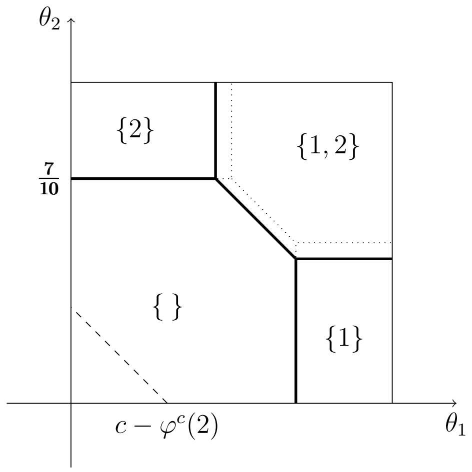
Panel 1a is copied from Cornelli (1996, Figure 1), who essentially solves our model without network effects, i.e., with $v^a(\theta_i, k) = \theta_i$ and $\varphi(k) = 0$ for all θ_i and k. Here, the creator's first-best solution is to provide the good to both users for all $\boldsymbol{\theta}$ north-east of the dashed line defined by cost c and not to provide it otherwise. Because incentive compatibility prevents the creator from extracting full surplus, the first-best allocation does not maximize profits when information



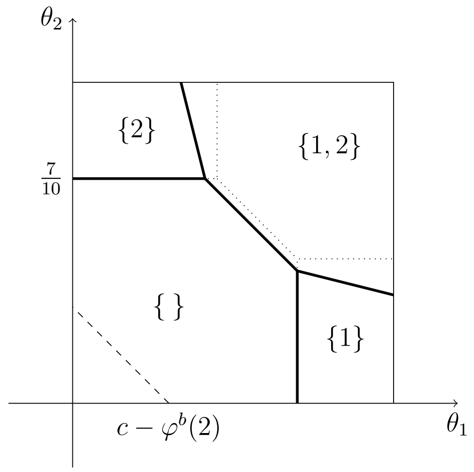
(a) No network effects.



(b) No user network effects, constant creator network effects, $\varphi^b(1) = \varphi^b(2)$.



(c) No user network effects, variable creator network effects, $\varphi^c(1) < \varphi^c(2)$.



(d) User and creator network effects.

Figure 1: Optimal allocations when adding components of our model step by step. The creator network effects essentially lead to the model of Cornelli (1996) with a different cost function. Lines are straight due to the assumed linearity of ψ (linear v and uniformly distributed types).

rents are accounted for. For type combinations north-east of the dash-dotted line in Panel 1a, the sum of virtual values covers the cost. Excluding types $\theta_i < 1/2$ with negative virtual values increases revenue, $\gamma^a(k-1, k, \boldsymbol{\theta}) = 0$ for all $k, \boldsymbol{\theta}$. Hence, the good is provided if and only if the sum of non-negative virtual values exceeds the cost, $C^a(k) = c$ for all k, and, in contrast to the first-best allocation, sometimes only a single user may access the content. In contrast to the public-good case, the possibility to provide the good while excluding low types allows to maintain higher prices. In contrast to the private-good case, the non-rivalry allows the seller to accept all users that increase revenue. Thus, the non-rivalry with a fixed cost creates a positive externality among users even without network effects. All types $\theta_i > 1/2$ have a positive externality because they help to cover the creator's cost c.

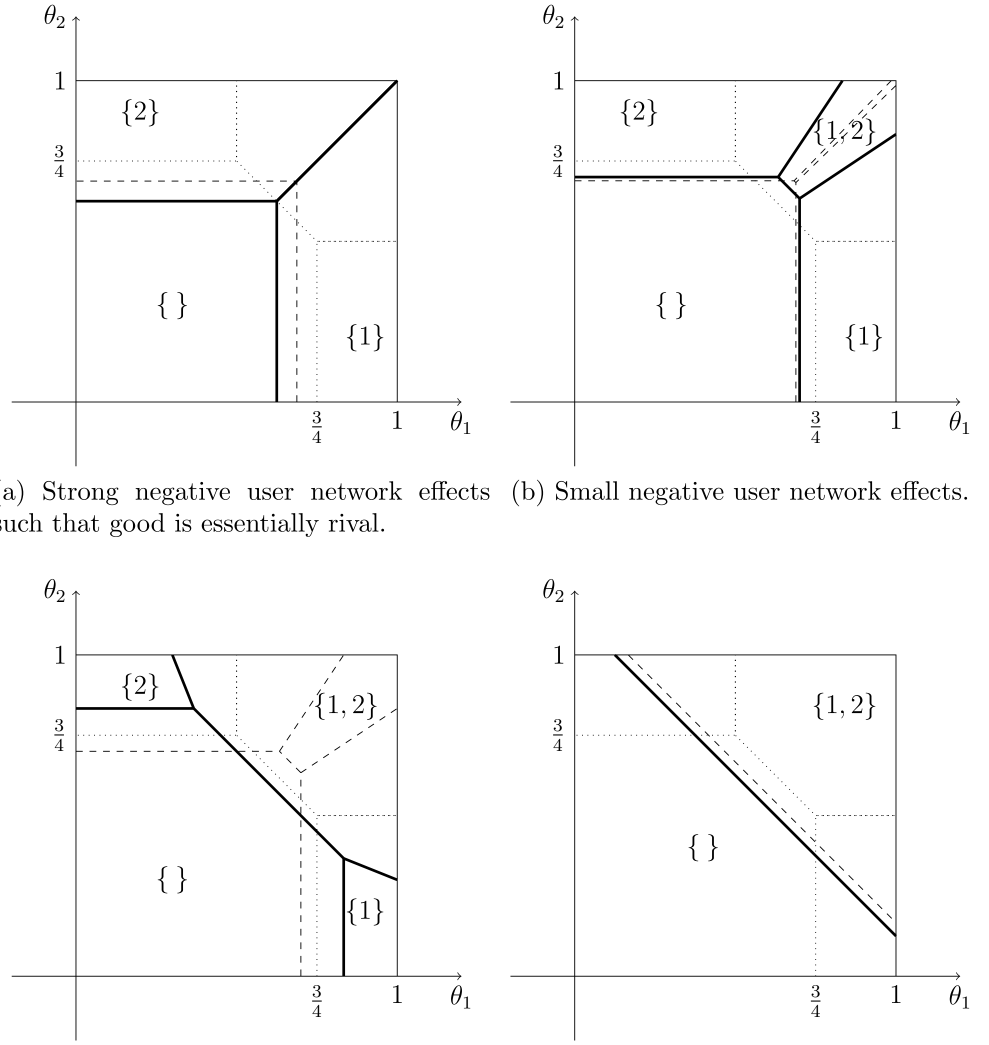
Panel 1b and Panel 1c incorporate creator profit network effects. In Panel 1b, they are constant for any provision, $\varphi^c(2) = \varphi^c(1) = 1/10 > \varphi^c(0) = 0$. The new optimal allocation is visualized by the thick black lines, where as the dotted lines represent the optimal allocation from Panel 1a. Incorporating these constant creator profit network effects, we see a uniform shift of the adjusted cost $C^b(k) = C^a(k) - 1/10$ and a shift in the thresholds for good provision to a single user and to two users. Panel 1c includes variable profit network effects, specifically $\varphi^c(2) = 1/5 > \varphi^c(1) = 1/10 > \varphi^c(0) = 0$. There is an additional shift just for the provision for two users compared to Panel 1b. There is, however, no effect on the threshold of providing to one user only as the extractable valuation while providing to one user does not change. Only when provided to both users does the network effect allow for lower types to be included in the allocation.

Panel 1d depicts the optimal allocation in a setting with the same variable network effects as in 1c and additionally incorporates (positive) consumer network effects $v^d(\theta_i, k) = (2+k)\theta_i/3$, where $v^d(\theta_i, 1) = v^a(\theta, k)$ for all k. The dashed line characterizes the optimal allocation from Panel 1c. Because the consumer network effects are positive, this line is shifted to the south-west. Additionally, the lines separating the allocation of providing to one instead of two consumers is tilted. In the other panels, a consumer type left/below this line has an insufficient virtual value and is excluded in Cornelli (1996) purely for incentives, i.e., to maintain lower information rents for higher types. In Panel 1d, however, the designer does not want to exclude all such types because an exclusion also reduces the valuation of the other consumer. The higher the valuation of this other consumer the higher the loss in extractable value of removing a consumer.

Figure 2 focuses on the user network effects (no profit network effects), and it shows how our model nests benchmarks from the literature. We assume uniformly distributed types on $[0, 1]$, and cost $c = 1/4$ with the utility function $v(\theta, 1) = \pi\theta$ and $v(\theta, 2) = (1 - \pi)\theta$. That is, an decrease in π makes consumption alone less valuable and consumption together more valuable. The optimal allocation without network effects ($\pi = 1/2$) is depicted by the dotted lines in each panel.

Panel 2a essentially represents the case of an indivisible, excludable, rival good (Myerson, 1981). The thick line in that Panel represents the optimal allocation with $\pi = 1$, i.e., a user only garners a payoff if he consumes the good alone and otherwise negative network effects destroy all value. Our optimal direct mechanism of Proposition 1 collapses to a second-price auction with a reserve price. A decrease in π up to $2/3$ (dashed line) does not qualitatively change the structure of the allocation. Only the reserve price changes as the good becomes less valuable when consumed alone while consumption together still destroys too much value that it is never optimal to have both users share the good.

Panel 2b shows smaller negative network effects. For parameters π slightly smaller than $2/3$ (dashed line), it becomes optimal to have some types close to



(a) Strong negative user network effects such that good is essentially rival.

(b) Small negative user network effects.

(c) Small positive user network effects.

(d) Positive negative user network effects such that good is essentially public.

Figure 2: Optimal allocations compared to benchmark cases. If user network effects are sufficiently strongly negative, the good is provided to at most one user (Myerson, 1981). If user network effects are sufficiently strongly positive, the good is provided to at least two users (Clarke, 1971; Groves, 1973). The dotted lines represent the settings without network effects (Cornelli, 1996).

the 45-degree line share the good. However, if the types are too far apart it is optimal to award the good only to the higher type. Further decreasing π enlarges the area of type combinations for which the good is allocated to both users. Panel 2c shows smaller positive network effects with the thick line compared to no network effects (dotted line) and negative network effects (dashed line).

In essence, **Panel 2d** represents the case of a public good (non-excludable and non-rival). In that Panel, a user only garners a payoff if no users are excluded and otherwise exclusion destroys all value,[3] i.e., $\pi = 0$. Qualitatively, the optimal allocation looks the same for all π smaller than $1/4$ (dashed line), where the good is also only provided to both users or not at all. Our optimal direct mechanism of Proposition 1 collapses to a standard VCG mechanism (Clarke, 1971; Groves, 1973).

Cost externalities: More broadly, our settings can resemble that of a public good regardless of the direction of user network effects and despite the possibility of exclusion. The reason is that the fixed production cost c creates a positive externality among the users, even when user network effects are slightly negative. To see this, it is easiest to consider the dotted lines in Figure 2, which represent the setting without any network effects (Cornelli, 1996). Here, user 1 may benefit from an increase in user 2's type when this increase pushes the extractable revenue above the cost. However, this is only relevant for types $\theta_1 \in [1/2, 3/4]$ because lower types never get the good, and higher types always get the good and are indifferent between consumption alone and together. With network effects, these boundaries vary with θ_2. When network effects are positive, this cost externality still exists, and the positive network externality is added on top, making the total externality strictly positive in expectation. In the allocation depicted in Panel 2b, an increase in θ_2 can add user 2 to the audience or even kick out user 1 from the audience. However, there is an interval of types θ_1 such that the good is not produced for low θ_2, only provided to user 2 for high θ_2, and provided to both for intermediate θ_2, i.e., the total externality is not monotone. User network effects have to be strongly negative for the positive externality due to production cost to vanish.

3.3 Optimal interim allocation and indirect Bayesian implementation

In the previous subsection, we dealt with the optimal allocation under a dominant-strategy incentive constraint. Because we arrived at our results using pointwise maximization of the relaxed problem in Section 3.1, this allocation is also optimal given a weaker Bayesian incentive constraint. That is, the allocation characterized by Proposition 1 is also the optimal interim allocation. Because direct mechanisms are rarely seen in practice, we now consider a Bayesian implementation of this allocation through an indirect mechanism. Indeed, we can rationalize commonly seen features of monetisation schemes in the creator economy.

Trivial economies: First, in Lemma 8, we determine conditions leading to a trivial outcome, i.e., never or always supplying the good. The first applies if the cost is sufficiently high, and the second applies if it is sufficiently low and there are sufficiently strong positive network effects. In the following, we call any economy satisfying the conditions of Lemma 8 "trivial" and all other settings "non-trivial." Second, in Lemma 9, we characterize the optimal interim expected

[3]To be precise, here it is not that the designer is unable to exclude users, but does not want to because value to extract can only exits without exclusion. Our model would also allow to model exclusion costs directly through profit network effects.

payment function in non-trivial cases, and the notation introduced here is helpful to state the indirect implementation.

Lemma 8. *The following settings constitute trivial economies.*
a) The good is never provided, i.e., $(\overline{q}, \overline{m})(\boldsymbol{\theta}) = (\mathbf{0}, \mathbf{0})$ *for all* $\boldsymbol{\theta}$ *if and only if*

$$k\psi(\overline{\theta}, k) < C(k) \quad \forall k. \tag{19}$$

b) The good is always provided for free to all, i.e., $(\overline{q}, \overline{m})(\boldsymbol{\theta}) = (\mathbf{1}, \mathbf{0})$ *for all* $\boldsymbol{\theta}$ *if and only if*

$$\begin{array}{ll} (i) & N\psi(0,N) \geq C(N), \quad \text{and} \\ (ii) & (N-k)\psi(0,N) \geq \gamma(k, N, \mathbf{0}_{\leq k}) \quad \forall k. \end{array} \tag{20}$$

Outside of the cases outlined above, the good is only sometimes provided, and maybe only to some users. For our indirect implementation, we now introduce the optimal interim allocation,

$$Q_i(\theta_i) := \mathbb{E}_{\boldsymbol{\theta_{-i}}}[q_i(\theta_i, \boldsymbol{\theta_{-i}})], \qquad \text{and } M_i(\theta_i) := \mathbb{E}_{\boldsymbol{\theta_{-i}}}[m_i(\theta_i, \boldsymbol{\theta_{-i}})],$$

where $\langle \overline{Q}, \overline{M} \rangle$ corresponds to $\langle \overline{q}, \overline{m} \rangle$.

We define $\theta^{\underline{p}}$ as the lowest type to get the good for some type combination,

$$\underline{y} = \min_{\theta_i} \{\theta_i : \exists \boldsymbol{\theta_{-i}} : \overline{q}(\theta_i, \boldsymbol{\theta_{-i}}) = 1\} = \min_{\boldsymbol{\theta_{-i}}} \underline{x}(\boldsymbol{\theta_{-i}}). \tag{21}$$

By optimality, it must be that $\overline{Q}_i(\theta_i) = \overline{M}_i(\theta_i) = 0$ for all $\theta_i < \underline{y}$, i.e., all these types are excluded. Incentive compatibility—i.e., Lemma 2—implies that type $\overline{\theta}$ must get the "best contract" in the sense that he always consumes the good with the most preferred audience size among all types. However, there may be other types who get the same allocation for all $\boldsymbol{\theta}_{-i}$. Let the smallest of these types be

$$\overline{y} = \min_{\theta_i} \{\theta_i : \overline{q}(\theta_i, \boldsymbol{\theta_{-i}}) = \overline{q}(\overline{\theta}, \boldsymbol{\theta_{-i}}) \text{ and } \overline{k}(\theta_i, \boldsymbol{\theta_{-i}}) = \overline{k}(\overline{\theta}, \boldsymbol{\theta_{-i}}) \forall \boldsymbol{\theta_{-i}}\}. \tag{22}$$

By incentive compatibility, $Q_i(\theta_i) = Q_i(\overline{\theta})$ and $M_i(\theta_i) = M_i(\overline{\theta})$ for all $\theta_i \in [\overline{y}, \overline{\theta}_i]$. It is possible that $\overline{y} = \overline{\theta}$. The following lemma shows that higher types pay more in expectation, and two types only have the same interim expected payment if they get exactly the same contract.

Lemma 9. *Consider any non-trivial economy. For all users* i*, the optimal interim expected transfer* $\overline{M}_i$ *is weakly increasing in the type, and whenever* $\overline{M}_i(x) = \overline{M}_i(y)$ *for two types* $x \neq y$*, it must be that* $\overline{k}(x, \boldsymbol{\theta}_{-i}) = \overline{k}(y, \boldsymbol{\theta}_{-i})$ *and* $\overline{q}_i(x, \boldsymbol{\theta}_{-i}) = \overline{q}_i(y, \boldsymbol{\theta}_{-i})$ *for all* $\boldsymbol{\theta}_{-i}$*.*

Indirect implementation: We exploit this structure to propose an indirect implementation with a simple structure. In an **all-pay contribution mechanism**, each agent i selects a price $p_i \in [0, \infty)$, and then an allocation decision is taken based on the price vector $\mathbf{p} = (p_i)_{i\in\mathcal{N}}$, while all agents have to pay their selected price independent of the allocation decision. In the formal language of Section 2, any all-pay contribution mechanism has an action set $\mathcal{A}_i = [0, \infty)$, and the outcome is pinned down by an outcome function g such that for any $\mathbf{p}$ and any user i, $g_i(\mathbf{p}) \in \{(0, p_i), (1, p_i)\}$. A (pure) strategy for a user i in the all-pay contribution game induced by such a mechanism is a function $\rho_i : \Theta_i \to \mathcal{A}_i$ that maps a type into a price, and a strategy profile consists of all players' strategies $\boldsymbol{\rho} = (\rho_i)_{i\in\mathcal{N}}$.

Proposition 2. *There is an all-pay contribution mechanism $\langle \widetilde{\mathcal{A}}, \widetilde{g} \rangle$ that implements the optimal allocation in a Bayesian Nash equilibrium $\boldsymbol{\rho}^*$.*

The idea behind this implementation is simple. Essentially, it follows a reverse revelation principle. We know that the optimal direct mechanism $\langle \overline{q}, \overline{m} \rangle$ is Bayesian incentive compatible and that $\overline{M}$ is either invertible or, where it is constant, gives all types in this region the same contract. Hence, the game induced by outcome function $\widetilde{g}_i(p_i, \mathbf{p}_{-i}) = (\widetilde{q}_i(p_i, \mathbf{p}_{-i}), p_i) = (\overline{q}_i(\overline{M}_i^{-1}(p_i), \overline{\mathbf{M}}_{-i}^{-1}(\mathbf{p}_{-i})), p_i)$ has a Bayesian Nash equilibrium in which any type θ_i selects price $\rho_i(\theta_i) = \overline{M}_i(\theta_i)$ to get expected utility for this type

$$\begin{aligned} \widetilde{u}_i(\rho_i(\theta_i), \theta_i) &= \mathbb{E}_{\boldsymbol{\theta}_{-i}} \left[\widetilde{q}_i(\rho_i(\theta_i), \boldsymbol{\rho}_{-i}(\boldsymbol{\theta}_{-i})) v(\theta_i, \widetilde{k}(\rho_i(\theta_i), \boldsymbol{\rho}_{-i}(\boldsymbol{\theta}_{-i})) - \rho_i(\theta_i)) \right] \\ &= \mathbb{E}_{\boldsymbol{\theta}_{-i}} \left[\overline{q}_i(\theta_i, \boldsymbol{\theta}_{-i}) v(\theta_i, \overline{k}(\theta_i, \boldsymbol{\theta}_{-i})) - \overline{m}_i(\theta_i, \boldsymbol{\theta}_{-i})) \right], \end{aligned}$$

where the incentive compatibility of $\langle \overline{q}, \overline{m} \rangle$ ensures that no type θ_i finds it profitable to select a price designated to another type. If agent i chooses a price outside of the range of the interval $[\overline{M}_i(\underline{y}), \overline{M}_i(\overline{y})]$, he never gets the good such that these deviations are also not profitable.

Cost externality→ voluntary payments: Self-selected contributions are a common feature in monetization schemes in the creator economy, e.g., "cheering" on Twitch or "donations" on other platforms. In our model, a high-type user volunteers to pay more than other users to jointly consume exactly the same good. Through the lens of our model, the rational behind this behavior is to increase the probability that the content is provided. That is, next to the transfer, also the threat of not producing any content can be used to incentivize users. The cost externality discussed in the previous section is the reason why this implementation resembles the equilibria in games of private provision of public goods. In the benchmark of Myerson (1981), there is no joint consumption and thus no cost externality such that the above implementation corresponds to an all-pay auction that awards the good to a single user.

Often users only have to pay their contribution when they actually consume the content. In some settings, a similar construction is possible, in which the user only has to pay if he is a member of the audience, i.e., we reverse engineer the optimal allocation and payments thorough the equation $M_i(\theta_i) = Q_i(\theta_i)p_i$ rather than $M_i(\theta_i) = p_i$. For instance, Cornelli (1996) suggests such a scheme or, as another example, the optimal allocation in Myerson (1981) can be implemented with a first-price auction rather than an all-pay auction (or the second-price auction essentially suggested by Proposition 1). However, it is not feasible in the example below, where ${}^{M_i}/_{Q_i}$ is not invertible.

Positive network effects→ community gifts: Twitch also allows gifted subscriptions ("community subs") by one user to another one. In some sense, we incorporate this kind of "altruism" as positive network effects into the payoff function v. If users have an intrinsic preference for joint consumption with a larger audience, there is an implicit motive to subsidize other users. Figure **??** (details in the appendix) illustrates how to employ community gifts to implement the optimal allocation in a 2-user setting with positive user network effects as depicted in Panel 2c. Here, the creator offers each user to voluntarily choose a reduced price $p_i \in [0, \underline{p})$ or a regular price $p_i \in [\underline{p}, \overline{p}]$ with an additional subsidy $s_i \in [0, \overline{s}]$. The regular prices and the subsidies are paid unconditionally as in Proposition 2. In contrast, the reduced price only has to be paid when the user obtains the good which happens only if the other user paid a large enough subsidy. Figure **??** shows the equilibrium strategy that implements the optimal allocation in the appropriate implementation game constructed in the appendix.

"Altruistic" features: To put it in a nutshell, our model offers a rationalization for two elements of monetization schemes in the creator economy, voluntary payments ("donations") and community support features. On a first glance, the efficacy of these features appears to be be driven by generosity. However, our model shows that both implementation features can work perfectly fine with purely self-interested agents when two important aspects of the creator economy are accounted for. First, the non-rivalry of digital content with production costs entails that a user may want to pay more to increase the probability that the digital content is provided (or, alternatively, that the creator can continue her career rather than leaving the creator economy for a job in another industry). Moreover, a user can benefit from subsidizing other users to garner positive network effects. Although we do not deny the importance (or existence) of altruism in small digital communities, we believe our model contributes to a better understanding of the full picture.

Negative network effects→ exclusive content: Some digital content is provided in a more exclusive fashion that is more akin to standard private-good provision. For example, chess streamers may discuss games in front of a larger audience of subscribers, but may also offer to give some students private and exclusive feedback on their own games. As an alternative example, live-streamers offering adult content may perform in front of a larger audience, but may change to a more private setting for additional payments. Figure **??** (details in the appendix) illustrates how to employ exclusive content to implement the optimal allocation in a 2-user setting with negative user network effects as depicted in Panel 2b.

Here, users are sorted into three different tiers depending on their types. Low types are "regular users" who may only consume the content in a group, and they only have to pay their selected price if they get to consumer the content.

Intermediate types are "subscribers." They unconditionally pay an entry fee for a first-price auction to bid on consuming the good alone, but may also consume the good as a group or not at all. High types are "premium subscribers." They unconditionally pay a higher entry fee for a first-price auction to bid on consuming the good alone, but get to consume the good with certainty (either alone or in a group).

Large markets: Because we focus on how to finance the careers of smaller content creators, oligopsony is a central ingredient of this paper. In small markets, each individual user can have a strategic impact. However, as discussed in the introduction, this industry has also spawned superstars with very large audiences. A natural question to ask is whether the optimal monetization of such creators' content is structurally different. For example, many large video platforms impose an identical cost on every user to consume the cost by showing the same amount of ads to all users (and only those with a sufficiently large value watch the add to consume to content) rather than screening types. Such a monetization turns out to be optimal in large markets.

The next result shows that if network effects are everywhere positive and concave, the optimal selling procedure converges to a simple posted price: all users whose value (in the large-audience limit) exceeds the price accept it to consume the content with certainty, while all others do not get the good and do not pay anything. Here, not only the network externalities disappear due to the market size, also the the cost externality vanishes. In a large market, the fixed production cost becomes negligible and, consequently, so does a single user's impact on the production decision. When user network externalities can also become negative, there may be a finite optimal audience size k^*, which, in the limit, can also be implemented with a posted price, but rationing is needed.

Lemma 10. *Consider the large-market limit as $N \to \infty$.*

a) Suppose

$$\lim_{k\to\infty} \varphi(k) = \varphi^* \quad \text{and} \quad \lim_{k\to\infty} v(\theta_i, k) = v^*(\theta_i) \quad \forall\theta_i. \tag{23}$$

Then, there is a cutoff p^a such that $\lim_{N\to\infty} \overline{q}_i(\theta_i, \boldsymbol{\theta_i}) = \mathbb{1}_{\theta_i \geq p^a}$ for all $\boldsymbol{\theta}$. That is, the optimal allocation is implementable by a simple posted price p.

b) Suppose $k^* = \arg\max_k\{k\psi(\overline{\theta}, k) + \varphi(k)\}$ is finite. Then, in the limit as $N \to \infty$, the optimal allocation is implementable by a simple posted price $p^b = v(\overline{\theta}, k^*)$ with random rationing in case more than k^* agents accept the price.

An alternative perspective on result a) above is to consider it as the limit of the implementation suggested before. As the extractable revenue increases as the market grows, the constant costs c becomes essentially irrelevant for the production decision. In the limit, the good is produced with probability one and also the marginal network effects vanishes. Hence, a user's extra contribution has no impact, and each user either selects the minimum price to be eligible for consumption or not to pay at all. Result b) is an implication of the fact that the optimal k^*-unit auction converges to posting the a price equal to the highest possible value as the number of bidders approaches infinity.

4 Conclusion

We design the profit-maximizing monetization schedule for a monopolistic digital content creator. Her content is an excludable and non-rival good that is produced once at a fixed cost. Users have heterogeneous private values that depend on the audience size, i.e., how many other users also consume the content. Moreover, the audience size maps into additional creator profits. We construct an allocation algorithm to implement the optimal allocation while truthful revelation of the types is a dominant strategy for users. From a purely technical perspective, our contribution can be helpful in other contexts as well. For instance, the user network effects in our model also apply in situations in which an authority is procuring, say, taxi licences, where the number of firms obtaining a license substantially influence the profit in the market, i.e., the value for a license.

On top of the direct mechanism, we discuss indirect implementations of this allocation. A more specific and applied contribution of our paper is to rationalize features of monetization schemes in the creator economy. For instance, we can explain donations to the creator or even other users, which seem to be based on users' generosity. The rational that makes such implementation features optimal in our setting is that users take into account a cost externality and a user network externality. They are willing to pay more than others for the same good to increase the probability of production, and they are willing to subsidize others when enlarging the audience is of inherent value to them. In contrast, when users dislike larger audiences they are willing to pay extra to have others excluded. We believe that our mechanism-design approach is helpful to shed light on other instruments in this industry.

Appendix

5 Proofs

Proof of Lemma 1. We first show that in any incentive-compatible mechanism (2), (3), and (4) must hold.

Fix any user i and two types $x > y$. (IC) requires that for each $\boldsymbol{\theta}_{-i}$ (replaced by $\cdot$ below)

$$\begin{aligned} U_i(x,\cdot) = v(x,k(x,\cdot))q_i(x,\cdot) - m_i(x,\cdot) \quad &\geq v(x,k(y,\cdot))q_i(y,\cdot) - m_i(y,\cdot) \qquad \text{and} \\ U_i(y,\cdot) = v(y,k(y,\cdot))q_i(y,\cdot) - m_i(y,\cdot) \quad &\geq v(y,k(x,\cdot))q_i(x,\cdot) - m_i(x,\cdot). \end{aligned}$$

Subtracting the inequalities yields

$$q_i(x,\cdot)\big(v(x,k(x,\cdot)) - v(y,k(x,\cdot))\big) \geq q_i(y,\cdot)\big(v(x,k(y,\cdot)) - v(y,k(y,\cdot))\big).$$

Hence, for all $x, y, \boldsymbol{\theta}_{-i}$ with $q_i(x,\boldsymbol{\theta}_{-i}) = q(y,\boldsymbol{\theta}_{-i}) = 1$, we obtain (3). Because $x > y$ and v is increasing in θ_i for all k, we obtain for all other $x, y, \boldsymbol{\theta}_{-i}$ that q_i must be weakly increasing in θ_i,

$$q_i(x,\cdot) \geq q_i(y,\cdot)\left(\frac{v(x,k(y,\cdot)) - v(y,k(y,\cdot))}{v(x,k(x,\cdot)) - v(y,k(x,\cdot))}\right).$$

Because the second factor on the right-hand side is positive, it must be that $q_i(x,\cdot) = 1$ when $q_i(y,\cdot) = 1$, and it must be that $q_i(y,\cdot) = 0$ when $q_i(x,\cdot) = 0$. Therefore, for all $\boldsymbol{\theta}_{-i}$, q_i is either constant or has exactly one jump upwards, i.e., (2) holds.

Moreover, considering $x = y + \delta$ with $q_i(x,\cdot) = q_i(y,\cdot) = 1$, we obtain

$$\begin{aligned}
\lim_{\delta\to 0}\frac{U_i(y+\delta,\cdot) - U(y,\cdot)}{\delta} &\geq \lim_{\delta\to 0}\frac{1\big(v(y+\delta,k(y+\delta,\cdot)) - v(y,k(y,\cdot))\big)}{\delta}\\
&= \frac{\partial v(y,k(y,\cdot))}{\partial y},\\
\lim_{\delta\to 0}\frac{U_i(x,\cdot) - U(x-\delta,\cdot)}{\delta} &\leq \lim_{\delta\to 0}\frac{1\big(v(x,k(x,\cdot)) - v(x-\delta,k(x-\delta,\cdot))\big)}{\delta}\\
&= \frac{\partial v(x,k(x,\cdot))}{\partial x},
\end{aligned}$$

which implies that $U_i'(x,\boldsymbol{\theta}_{-i}) = \frac{\partial v(x,k(x))}{\partial x}$ wherever q_i is equal to one. Since (IC) also implies that U_i is Lipschitz-continuous, it is differentiable almost everywhere and equals the integral over its derivative. Hence, (4) follows from rearranging.

Now, suppose (2), (3), and (4) hold. Then, also (IC) holds, because

$$\begin{aligned}
U_i(x,\cdot) &\geq u_i(y,\cdot|x) = q_i(y,\cdot)v(x,k(y,\cdot) - m_i(y,\cdot)\\
U_i(x,\cdot) &\geq u_i(y,\cdot|x) + q_i(y,\cdot)v(y,k(y,\cdot) - q_i(y,\cdot)v(y,k(y,\cdot)\\
U_i(x,\cdot) &\geq q_i(y,\cdot)\big(v(x,k(y,\cdot) - v(y,k(y,\cdot))\big) + U_i(y,\cdot)\\
\int_y^x q_i(t,\cdot)\frac{\partial v(t,k(t,\cdot))}{\partial t}dt &\geq \int_y^x q_i(y,\cdot)\frac{\partial v(t,k(y,\cdot))}{\partial t}dt
\end{aligned}$$

is implied by these conditions. □

Proof of Lemma 2. By (3) of Lemma 1, (IC) requires that for all $x > y > \underline{x}(\boldsymbol{\theta}_{-i})$,

$$((v(x,k(x,\cdot)) - v(y,k(x,\cdot))) - (v(x,k(y,\cdot)) - v(y,k(y,\cdot))) \geq 0. \tag{24}$$

By contradiction suppose that $k(x,\cdot) = k_j$ and $k(y,\cdot) = k_{j'}$ with $j > j'$, i.e., the lower type y receives a "better" audience size. (SC) implies that

$$((v(x,k(x,\cdot)) - v(y,k(x,\cdot))) - (v(x,k(y,\cdot)) - v(y,k(y,\cdot))) < 0, \tag{25}$$

a contradiction.

Vice versa, suppose that for all $x > y$, $k(x, \cdot) = k_j$ and $k(y, \cdot) = k_{j'}$ with $j \leq j'$. The negation of (25) implies (24). $\square$

Proof of Lemma 3. Let $\overline{J}(\boldsymbol{\theta})$ be the optimal audience in the relaxed problem given type vector $\boldsymbol{\theta}$. That is, $q_i(\boldsymbol{\theta}) = \mathbb{1}_{i \in \overline{J}(\boldsymbol{\theta})}$.

Contradicting the lemma, suppose $\overline{J}(\widehat{\boldsymbol{\theta}}) \neq \varnothing$ although (8) is violated for all J. The creator makes a loss that can be avoided by setting $\overline{J}(\widehat{\boldsymbol{\theta}}) = \varnothing$. Analogously, $\overline{J}(\widehat{\boldsymbol{\theta}}) = \varnothing$ cannot be optimal for type vector $\widehat{\boldsymbol{\theta}}$ if a profitable J satisfying (8) exists. $\square$

Proof of Lemma 4. The right-hand side of (9) is zero when $|J| = |J'|$, while the left-hand side is positive when $J \neq J'$ has structure (10). $\square$

Proof of Lemma 5. Fix any $\boldsymbol{\theta}$. By Lemma 4, the optimal audience J^* is either empty or it is the most profitable candidate set, $J^* = J_{k^*}$ for some k^*. The algorithm compares only candidate sets, and it considers all such sets. Hence, it cannot select a set with another structure. We show that our algorithm selects $K_{N+1} = J_{k^*}$ and only provides the good if and only if $J^* = J_{k^*}$.

If J^* is empty, Lemma 3 implies that for all non-empty candidate sets J_k, $\Psi(\boldsymbol{\theta}|J_k) < C(k)$. Hence, the algorithm never chooses to provide the good in the final step.

Similarly, if J^* is non-empty, optimality implies that $\Psi(\boldsymbol{\theta}|J_{k^*}) \geq C(k^*)$. Hence, the algorithm chooses to provide the good in the final step if $K_{N+1} = J_{k^*}$. By optimality, $J_{k^*} = \overline{\max}_{\boldsymbol{\theta}}\{J_{k^*}, K_k\}$ for all k. Hence, J_{k^*} becomes the comparison set at step k^*, $K_{k^*+1} = \overline{\max}\{J_{k^*}, K_{k^*}\} = J_{k^*}$, and it stays the comparison set until the final step, $K_{N+1} = K_{k^*+1} = J_{k^*}$.

Suppose there is a candidate set $J_k \neq K_{N+1}$ such that $\Psi(\boldsymbol{\theta}|J_k) \geq C(k)$, and let $k^* = k + j$ for some $j > 0$. Because $J_{k^*} = \overline{\max}_{\boldsymbol{\theta}}\{J_k, J_{k^*}\}$, it holds that

$$\begin{aligned}
\sum_{i=1}^{j} \psi(\theta_{k+i}, k+j) &\geq \varphi(k) - \varphi(k+j) + \sum_{i=1}^{k} [\psi(\theta_i, k) - \psi(\theta_i, k+j)] \\
\Psi(\boldsymbol{\theta}|J_{k+j}) &\geq \varphi(k) - \varphi(k+j) + \Psi(\boldsymbol{\theta}|J_k) \\
\Psi(\boldsymbol{\theta}|J_{k^*}) &\geq \varphi(k) - \varphi(k+j) + C(k) = C(k^*),
\end{aligned}$$

and an analogous argument holds when $j < 0$. Hence, if the algorithm chooses not to provide the good at the final step, there is no other set for which the adjusted total cost is covered. $\square$

Proof of Lemma 6. Fix any $\boldsymbol{\theta}$. By definition of $\widehat{k}$ in (16), for all $k > \widehat{k}$, it holds that $\psi(\theta_k, k) < \gamma(k-1, k, \boldsymbol{\theta}_{\leq k-1})$. Hence, for all $j \geq 1, j \leq n - \widehat{k}$, we have

$$\sum_{i=1}^{j} \psi(\theta_{\widehat{k}+i}, \widehat{k}+i) < \sum_{i=1}^{j} \gamma(\widehat{k}+i-1, \widehat{k}+i, \boldsymbol{\theta}_{\leq \widehat{k}+i-1}), \tag{26}$$

which is equivalent to the statement that adding any number of j users to audience $J_{\widehat{k}}$ reduces profit. The reason is that the right-hand side of the inequality above can be rewritten as

$$\sum_{i=1}^{j}\left(\varphi(\widehat{k}+i-1)-\varphi(\widehat{k}+i)+\sum_{\ell=1}^{\widehat{k}+i-1}(\psi(\theta_\ell,\widehat{k}+i-1)-\psi(\theta_\ell,\widehat{k}+i))\right)$$
$$=(\varphi(\widehat{k})-\varphi(\widehat{k}+j))+\sum_{i=1}^{j}\sum_{\ell=1}^{\widehat{k}+i-1}(\psi(\theta_\ell,\widehat{k}+i-1)-\psi(\theta_\ell,\widehat{k}+i))$$
$$=(\varphi(\widehat{k})-\varphi(\widehat{k}+j))+\sum_{i=1}^{\widehat{k}}\psi(\theta_i,\widehat{k})+\sum_{i=1}^{\widehat{j}-1}\psi(\theta_{\widehat{k}+i},\widehat{k}+i)-\sum_{i=1}^{\widehat{k}+j-1}\psi(\theta_i,\widehat{k}+j)$$

such that (26) becomes

$$\sum_{i=1}^{j}\psi(\theta_{\widehat{k}+i},\widehat{k}+j)<(\varphi(\widehat{k})-\varphi(\widehat{k}+j))+\sum_{i=1}^{\widehat{k}}(\psi(\theta_i,\widehat{k})-\psi(\theta_i,\widehat{k}+j))$$
$$\sum_{i=1}^{j}\psi(\theta_{\widehat{k}+i},\widehat{k}+j)<\gamma(\widehat{k},\widehat{k}+j,\boldsymbol{\theta}_{\leq\widehat{k}}).$$

□

Proof of Lemma 7. Fix any $\boldsymbol{\theta}$, and let $\widehat{k}=\widehat{k}(\boldsymbol{\theta})$.

Suppose (18) holds. For any $j<\widehat{k}$, we have

$$\sum_{i=\widehat{k}-j+1}^{\widehat{k}}\psi(\theta_i,\widehat{k})\geq j\psi(\theta_{\widehat{k}},\widehat{k})\geq j\gamma(\widehat{k}-1,\widehat{k},\boldsymbol{\theta}_{\leq\widehat{k}-1})\geq\gamma(\widehat{k}-j,\widehat{k},\boldsymbol{\theta}_{\leq k}),$$

where the first inequality follows from our relabeling of subscripts and regularity, the second inequality follows from the definition of $\widehat{k}$ in (16), and the third inequality follows from (18). Consequently, for any candidate set $J_{\widehat{k}-j}$, adding all users up to user $\widehat{k}$ is profitable and removing any user from $J_{\widehat{k}}$ is unprofitable. Lemma 6 implies that adding more users is also unprofitable. The resulting allocation is the most profitable candidate set, i.e., $\widehat{q}(\boldsymbol{\theta})=\overline{q}(\boldsymbol{\theta})$.

Suppose (18) is violated for some j. Construct another type vector $\boldsymbol{\theta}'$ such that $\theta'_i=\theta_i$ for all $i\notin\{\widehat{k}-j+1,\widehat{k}-j,\dots,\widehat{k}\}$ and $\psi(\theta'_i,\widehat{k})=\gamma(\widehat{k}-1,\widehat{k},\boldsymbol{\theta}_{\leq\widehat{k}-1})$ for all $i\in\{\widehat{k}-j+1,\widehat{k}-j,\dots,\widehat{k}\}$. That is, the types of users outside of the subset remain unchanged from $\boldsymbol{\theta}$, while the types of all users in the subset are set equal to a local cutoff type. Because we only changed the types of users $i\leq\widehat{k}$, the definition of $\widehat{k}$ (16) implies that $\widehat{k}(\boldsymbol{\theta}')=\widehat{k}=\widehat{k}(\boldsymbol{\theta})$.

The violation of (18) implies the following inequality

$$\sum_{i=\widehat{k}-j+1}^{\widehat{k}} \psi(\theta_i', \widehat{k}) = j\psi(\theta_{\widehat{k}}', \widehat{k}) = j\gamma(\widehat{k}-1, \widehat{k}, \boldsymbol{\theta}_{\leq \widehat{k}-1})$$
$$< \gamma(\widehat{k}-j, \widehat{k}, \boldsymbol{\theta}_{\leq \widehat{k}-j}) = \gamma(\widehat{k}-j, \widehat{k}, \boldsymbol{\theta}'_{\leq \widehat{k}-j}).$$

Hence, for type vector $\boldsymbol{\theta}'$, removing users from audience $J_{\widehat{k}}$ increases profit. Since the inequality is strict, there is a positive measure of type vectors such that the above holds. Hence, allocation rule $\widehat{q}$ cannot be optimal. □

Proof of Lemma **??**. Clearly, (18) holds if the inequality holds for both added components separately, i.e., if

$$j(\varphi(k-1) - \varphi(k)) \geq \varphi(k-j) - \varphi(k) \quad \text{and} \tag{27}$$
$$j\sum_{i=1}^{k-1}(\psi(\theta_i, k-1) - \psi(\theta_i, k)) \geq \sum_{i=1}^{k-j}(\psi(\theta_i, k-j) - \psi(\theta_i, k)). \tag{28}$$

(27) is equivalent to

$$(\varphi(k-1) - \varphi(k)) + (j-1)(\varphi(k-1) - \varphi(k)) \geq$$
$$\varphi(k-j) - \varphi(k-j+1) + \varphi(k-j+1) - \varphi(k),$$

and we show that the inequality holds for both added components separately. Concavity (27) implies that

$$\varphi(k-1) - \varphi(k) \geq \varphi(k-2) - \varphi(k-1) \geq \cdots \geq \varphi(k-j+1) - \varphi(k-j).$$

The missing second part,

$$(j-1)(\varphi(k-1) - \varphi(k)) \geq \varphi(k-j+1) - \varphi(k),$$

is the original inequality (27) with j replaced by $(j-1)$. Therefore, (27) holds because it holds for $(j-1)$ by induction as it is just a restatement of concavity (**??**) when $j = 2$.

Next, we consider (28) with a case distinction.
a) Suppose that $\psi(\theta_i, k-1) - \psi(\theta_i, k) \geq 0$ (we assume the same sign for all θ_i in our model). Then we can use the same argument as above because

$$j\sum_{i=1}^{k-1}(\psi(\theta_i, k-1) - \psi(\theta_i, k)) \geq j\sum_{i=1}^{k-j}(\psi(\theta_i, k-1) - \psi(\theta_i, k))$$
$$= \sum_{i=1}^{k-j}(\psi(\theta_i, k-1) - \psi(\theta_i, k)) + (j-1)\sum_{i=1}^{k-j}(\psi(\theta_i, k-1) - \psi(\theta_i, k)).$$

b) Suppose that $\psi(\theta_i, k) < \psi(\theta_i, k-1)$ □

Proof of Proposition 1. To show that $\overline{q}$ is incentive-compatible, we verify that it always satisfies the three conditions of Lemma 1, i.e., (2), (3), and (4). The allocation is individually rational if it is incentive-compatible and the participation constraint binds for type $\theta_i = 0$, which is true.

(4) is satisfied by construction (15).

We fix a user i and consider two possible types $x > y$ for him. We also fix the types of all other users $\boldsymbol{\theta}_{-i}$ and order them, $\theta_j > \theta_{j+1}$ for all j.

To show (2), we need that

$$\overline{q}_i(y, \boldsymbol{\theta}_{-i}) = 1 \implies \quad q_i(x, \boldsymbol{\theta}_{-i}) = 1, \text{ and } \quad \overline{q}_i(x, \boldsymbol{\theta}_{-i}) = 0 \implies \quad q_i(y, \boldsymbol{\theta}_{-i}) = 0.$$

The first part cannot be violated because if the algorithm added type y (and possibly others) to the audience at some step k, it would also add type x. That is, the inclusion condition at step k is

$$\psi(y, k) + \sum_{j=1}^{k-1} \psi(\theta_j, k) - \sum_{j=1}^{k'-1} \psi(\theta_j, k') \geq \varphi(k') - \varphi(k)$$

for some k'. Because of regularity, this condition also holds for type x. Hence, if type x was not added at an earlier step, it would have been added at step k as well. Similarly, if the condition does not hold for type x, it cannot hold for type y. Also because of regularity, we have that if there exists an audience set including type y for which the provision condition (8), the condition also holds for the same audience set including type $x > y$. Therefore, given any $\boldsymbol{\theta}_{-i}$, $\overline{q}_i$ is weakly increasing in type θ_i for all i such that (2) holds.

Suppose that $\overline{q}_i(x, \cdot) = \overline{q}_i(y, \cdot) = 1$, but $\overline{k}(x, \cdot) = k_x > \overline{k}(y, \cdot) = k_y$. It must be that

$$\begin{aligned}
\psi(x, k_x) + \sum_{j=1}^{k_x-1} \psi(\theta_j, k_x) + \varphi(k_x) \quad &\geq \psi(x, k_y) + \sum_{j=1}^{k_y-1} \psi(\theta_j, k_y) + \varphi(k_y), \text{ and} \\
\psi(y, k_y) + \sum_{j=1}^{k_y-1} \psi(\theta_j, k_y) + \varphi(k_y) \quad &\geq \psi(y, k_x) + \sum_{j=1}^{k_x-1} \psi(\theta_j, k_x) + \varphi(k_x).
\end{aligned}$$

Subtracting these two inequalities, we obtain

$$\begin{aligned}
\psi(x, k_x) - \psi(y, k_x) \quad &\geq \psi(x, k_y) - \psi(y, k_y) \\
(v(x, k_x) - v(y, k_x)) - (v(x, k_y) - v(y, k_y)) \quad &\geq h(x, k_x, k_y) + h(y, k_x, k_y), \qquad (29)
\end{aligned}$$

where

$$h(x, k_x, k_y) := \frac{1 - F(x)}{f(x)} \left(\frac{\partial v(x, k_x)}{\partial x} - \frac{\partial v(x, k_y)}{\partial x} \right).$$

The hazard rate is non-negative, and, by (SC), the second factor is positive if and only if $v(x, k_x) > v(x, k_y)$. That is, the left-hand side of (29) is positive if and only if type x prefers k_x over k_y. Thus, (3) holds.

Hence, $\langle \overline{q}, \overline{m} \rangle$ is incentive compatible and individually rational, and thus also solves the full constrained problem. $\square$

Proof of Lemma 8. Suppose (19) holds. For any $k \leq N$, the revenue from accepting the highest k types does not cover the adjusted cost $C(k)$ even when the highest k types are all $\overline{\theta}$. By regularity, this is also true for all other type vectors. Suppose (19) does not hold for some k. Then there exists a type vector realization such that the profit from producing the good for the highest k types is profitable. Hence, Lemma 3 implies statement a).

Suppose (20) holds, and suppose all types are zero, $\boldsymbol{\theta} = \mathbf{0}$. Part (i) implies that accepting all N users covers the adjusted total cost , and part (ii) implies that removing any user from this full audience reduces profits. Hence $\overline{q}_i(\mathbf{0}) = 1$ for all i. The incentive compatibility of $\langle \overline{q}, \overline{m} \rangle$—Lemma 1(i)—implies the audience cannot be smaller for any other type vector. Individual rationality and $v(0, k) = 0$ for all k imply $\overline{m}_i(\mathbf{0}) = 0$ for all i.

Suppose one part of (20) does not hold. Either accepting all users does not cover the adjusted cost or excluding some users for some type vector increases profit. Hence, accepting all users is not optimal. Combined with the previous paragraph, this implies statement b). $\square$

Proof of Lemma 9. This lemma follows from applying the expectation operator to the payment function $\overline{m}_i$.

$\square$

Proof of Proposition 2. Take any setting and the corresponding optimal direct mechanism $\langle \overline{Q}, \overline{M} \rangle$. We construct the corresponding all-pay contribution mechanism $\langle \widetilde{\mathcal{A}}, \widetilde{g} \rangle$ and the Bayesian Nash equilibrium in the corresponding game that implements the optimal allocation, $\boldsymbol{\rho}^* = (\rho_i^*)_{i \in \mathcal{N}}$ where $\rho_i^* : \Theta_i \to \mathcal{A}_i = [0, \infty)$ is a strategy in the game.

First, set $\widetilde{g}_i(p_i, \boldsymbol{p}_{-i}) = (0, p_i)$ for all $p_i \notin [\overline{M}_i(\underline{y}), \overline{M}_i(\overline{y})]$ such that choosing any such price is weakly dominated by selecting price $p_i = 0$. Suppose all types $\theta_i < \underline{y}$ select $\rho_i^*(\theta_i) = 0$.

Second, suppose any type $\theta_i \in [\underline{y}, \overline{y})$ selects price

$$\rho_i^*(\theta_i) = \overline{M}_i(\theta_i),$$

and any type $\theta_i \geq \overline{y}$ selects price $\rho_i^*(\theta_i) = \overline{M}_i(\overline{\theta})$.

Third, whenever $\overline{M}_i$ is invertible, construct

$$\widetilde{g}_i(p_i, \mathbf{p}_{-i}) = (\overline{q}(\overline{M}_i^{-1}(p_i), \overline{\mathbf{M}}_{-i}^{-1}(\mathbf{p}_{-i})), p_i)$$

such that under the supposed strategy profile $\widetilde{g}_i(p_i, \boldsymbol{p}_{-i}) = (\overline{q}_i(\boldsymbol{\theta}), M_i(\theta_i))$ for all i and all $\boldsymbol{\theta}$. If for several types $x \neq y$ we have $\overline{M}_i(x) = M_i(y)$, Lemma 9 ensures that $q_j(x, \boldsymbol{\theta}_{-i}) = q_j(y, \boldsymbol{\theta}_{-i})$ for all $j \in \mathcal{N}$ such that it does not matter which types is used as input among types in a flat region of $\overline{M}_i$.

Under the proposed strategy profile, user i's expected utility from following strategy ρ_i is

$$\begin{aligned}\widetilde{u}_i(\rho_i(\theta_i), \theta_i) &= \mathbb{E}_{\boldsymbol{\theta}_{-i}}\left[\widetilde{q}_i(\rho_i(\theta_i), \boldsymbol{\rho}_{-i}(\boldsymbol{\theta}_{-i}))v(\theta_i, \widetilde{k}(\rho_i(\theta_i), \boldsymbol{\rho}_{-i}(\boldsymbol{\theta}_{-i})) - \rho_i(\theta_i))\right] \\ &= \mathbb{E}_{\boldsymbol{\theta}_{-i}}\left[\overline{q}_i(\theta_i, \boldsymbol{\theta}_{-i})v(\theta_i, \overline{k}(\theta_i, \boldsymbol{\theta}_{-i})) - \overline{m}_i(\theta_i, \boldsymbol{\theta}_{-i}))\right],\end{aligned}$$

which by incentive compatibility and individual rationality of $\langle \overline{q}, \overline{m}\rangle$ is at least as large as the expected utility from any deviation to any price $p_i \in \{0\} \cup [\overline{M}_i(\underline{y}), \overline{M}_i(\overline{y})]$, which dominate all other prices. Hence, $\boldsymbol{\rho}^*$ is a Bayesian Nash equilibrium of the constructed all-pay contribution and implements the optimal allocation for all $\boldsymbol{\theta}$. □

Proof of Lemma 10. a) Consider an increasing sequence of market sizes $N^{(n)} = (1, 2, \dots)$. Let $\boldsymbol{\theta}^{(n)}$ be the sequence of type profile realizations with $\boldsymbol{\theta}^n = (\boldsymbol{\theta}^{n-1}, \theta_n^n)$ (before ordering types according to virtual values), i.e., one more type draw is added each step along the sequence and the other types remain fixed $\theta_i^n = \theta_i^{n'}$ for all $i < n < n'$. Let $\overline{k}^n = |\overline{J}(\boldsymbol{\theta}^n)|$ be the optimal audience size at market size n along the sequence. Let $\psi^*(\theta_i) := \lim_{k\to\infty} \psi(\theta_i, k) = v^*(\theta_i) - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial v^*(\theta_i)}{\partial \theta_i}$.

$\overline{k}^n$ is weakly increasing in n because in the inequality below the left-hand side (after ordering types) is increasing in n and j, while the right-hand side converges to a constant as n and j increase (and the production cost c is also constant)

$$\sum_{i=\overline{k}^n}^{\overline{k}^n+j} \psi(\theta_i^n, \overline{k}^n + j) \geq \gamma(\overline{k}^n, \overline{k}^n + j, \boldsymbol{\theta}_{\leq \overline{k}^n})$$

$$\gamma(\overline{k}^n, \overline{k}^n + j, \boldsymbol{\theta}_{\leq \overline{k}^n}) = \underbrace{\varphi(\overline{k}^n) - \varphi(\overline{k}^n + j)}_{\to \varphi(\overline{k}^n) - \varphi^*} + \sum_{i=1}^{\overline{k}^n} \underbrace{[\psi(\theta_i, \overline{k}^n) - \psi(\theta_i, \overline{k}^n + j)]}_{\to \psi(\overline{\theta}, \overline{k}^n) - \psi^*(\overline{\theta})}.$$

That is, as the market grows it becomes profitable to increase the audience. Finally, for all j

$$\lim_{k\to\infty} \gamma(k, k+j, \boldsymbol{\theta}_{\leq k}) := \overbrace{\varphi(k) - \varphi(k+j)}^{\to 0} + \sum_{i=1}^{k} \overbrace{[\psi(\theta_i, k) - \psi(\theta_i, k+j)]}^{\to 0} = 0.$$

That is, as the audience grows the cutoff for the virtual value to be admitted to the audience converges to zero. The Consequently, in the large-market limit, the optimal allocation $\overline{q}_i^*(\boldsymbol{\theta}) = \mathbb{1}_{\psi^*(\theta_i)\geq 0}$ can be implemented by posting a price p^a such that $\psi^*(p^a) = 0$.

b) By assumption, the profit $k^*\psi(\overline{\theta}, k^*) + \varphi(k^*) - c$ is an upper bound on the extractable revenue for all $n > k^*$. Because $\theta_{k^*}^n \to \overline{\theta}$, the probability that at least k^* users accept a price $p^b \approx v(\overline{\theta}, k^*)$ approaches 1. Hence, any rule that rations trade at p^b to at most k^* users approaches the profit upper bound arbitrarily closely. □

References

Belleflamme, P., Omrani, N., Peitz, M., 2015. The economics of crowdfunding platforms. Information Economics and Policy 33, 11–28.

Bierbrauer, F.J., 2011. Incomplete contracts and excludable public goods. Journal of Public Economics 95, 553–569.

Birulin, O., 2006. Public goods with congestion. Journal of Economic Theory 129, 289–299.

Buchanan, J.M., 1965. An economic theory of clubs. Economica 32, 1–14.

Clarke, E.H., 1971. Multipart pricing of public goods. Public choice , 17–33.

Cornelli, F., 1996. Optimal selling procedures with fixed costs. Journal of Economic Theory 71, 1–30.

Csorba, G., 2008. Screening contracts in the presence of positive network effects. International Journal of Industrial Organization 26, 213–226.

Deb, R., Razzolini, L., 1999. Auction-like mechanisms for pricing excludable public goods. Journal of Economic Theory 88, 340–368.

Ellman, M., Hurkens, S., 2019. Optimal crowdfunding design. Journal of Economic Theory 184, 104939.

Groves, T., 1973. Incentives in teams. Econometrica: Journal of the Econometric Society , 617–631.

Hellwig, M.F., 2003. Public-good provision with many participants. The Review of Economic Studies 70, 589–614.

Hellwig, M.F., 2005. A utilitarian approach to the provision and pricing of excludable public goods. Journal of public Economics 89, 1981–2003.

Hellwig, M.F., 2007. The provision and pricing of excludable public goods: Ramsey–boiteux pricing versus bundling. Journal of Public Economics 91, 511–540.

Imas, A., Madarász, K., 2021. Dominance-seeking and the economics of exclusion .

Jarman, F., Meisner, V., 2017. Deterministic mechanisms, the revelation principle, and ex-post constraints. Economics Letters 161, 96–98.

Jehiel, P., Moldovanu, B., Stacchetti, E., 1996. How (not) to sell nuclear weapons. The American Economic Review , 814–829.

Moulin, H., 1994. Serial cost-sharing of excludable public goods. The Review of Economic Studies 61, 305–325.

Moulin, H., Shenker, S., 1992. Serial cost sharing. Econometrica: Journal of the Econometric Society , 1009–1037.

Myerson, R.B., 1981. Optimal auction design. Mathematics of operations research 6, 58–73.

Norman, P., 2004. Efficient mechanisms for public goods with use exclusions. The Review of Economic Studies 71, 1163–1188.

Ostrizek, F., Sartori, E., 2023. Screening while controlling an externality. Games and Economic Behavior 139, 26–55.

Schmitz, P.W., 1997. Monopolistic provision of excludable public goods under private information. Public Finance 52, 89–101.

Segal, I., 1999. Contracting with externalities. The Quarterly Journal of Economics 114, 337–388.

Strausz, R., 2017. A theory of crowdfunding: A mechanism design approach with demand uncertainty and moral hazard. American Economic Review 107, 1430–76.