Optimal Membership Design *

Changhwa Lee, Scott Kominers, Piotr Dworczak, Marco Reuter

February 20, 2024

Abstract

Many allocation problems can be recast as designing membership. The defining feature of membership as an economic good is that its value depends on who is a member. We introduce a framework for optimal membership design by combining an otherwise standard mechanism-design model with allocative externalities that depend flexibly on agents' observable and unobservable characteristics. Our main technical result characterizes how the optimal mechanism depends on the pattern of externalities. Specifically, we show how the number of distinct membership tiers—differing in prices and potentially involving rationing—is increasing in the complexity of the externalities. This insight may help explain a number of mechanisms used in practice to sell membership goods, including artists charging below-market-clearing prices for concert tickets, heterogeneous pricing tiers for access to digital communities, the use of vesting and free allocation in the distribution of network tokens, and certain admission procedures used by colleges concerned about the diversity of the student body.

*We thank Ebehi Iyoha, Valet Jones, Ray Kluender, Tim Roughgarden, SAFA, as well as presentation audiences at a16z crypto, HBS, and Tokenomics for helpful comments. Dworczak gratefully acknowledges the support received under the ERC Starting grant IMD-101040122. Kominers gratefully acknowledge support from the Ng Fund and the Mathematics in Economics Research Fund of the Harvard Center of Mathematical Sciences and Applications. Part of this work was conducted during the Simons Laufer Mathematical Sciences Institute Fall 2023 program on the Mathematics and Computer Science of Market and Mechanism Design, which was supported by the National Science Foundation under Grant No. DMS-1928930 and by the Alfred P. Sloan Foundation under grant G-2021-16778. Kominers is a Research Partner at a16z crypto, which reviewed a draft of this article for compliance prior to publication and is an investor in crypto projects and online platforms (for general a16z disclosures, see `https://www.a16z.com/disclosures/`). Notwithstanding, the ideas and opinions expressed herein are those of the authors, rather than of a16z or its affiliates. Kominers also holds digital assets, including both fungible and non-fungible tokens, and advises a number of companies on marketplace and incentive design. The views expressed herein are those of the authors and do not necessarily represent the views of the EU or ERC, or the IMF, the IMF Executive Board, or IMF management. Any errors or omissions remain the sole responsibility of the authors.

1 Introduction

Many allocation problems can be recast as designing membership. The defining feature of membership as an economic good is that its value for any agent depends on who else is a member. For example, a ticket to a concert gives the owner the right to attend it; but the value of attending may depend on the concert atmosphere, which is influenced by the composition of the audience. NFT projects allocate digital assets to consumers; but it is often the community of other owners of a given NFT collection that creates demand for the asset in the first place.[1] Finally, designing rules for admission to a school or an academic program such as executive education requires understanding how prospective students' values for enrolling depend on peer effects and opportunities for networking.

In short, membership goods differ from conventional goods in that they feature potentially complex externalities between the owners. In this paper, we provide a simple framework for studying optimal design of membership that allows for a rich pattern of externalities that can depend on both observable and unobservable agent characteristics. We show how viewing a diverse set of allocation problems as membership design leads to a unifying explanation for a number of real-life selling practices such as lotteries and group-specific price discounts. Finally, our framework can be used to guide the design of a platform or project that must account for the interactions between various members to correctly price access.

Our model combines a standard mechanism-design framework—featuring a large population of agents, one-dimensional private information, quasi-linear utilities, and labels capturing observable information—with allocative externalities. We introduce allocative externalities by assuming that the value of membership for an agent is obtained as a product of her type with a term determined by the distribution of other members' characteristics. The main simplifying assumption is that—conditional on observables—agents care about the *average* characteristics of other members. The designer is assumed to know the distribution of agent characteristics and chooses an individually-rational and incentive-compatible mechanism to maximize an objective such as revenue or welfare.

We show that optimal membership can always be implemented by a mechanism with label-dependent tiered pricing. Dependence on labels means that different groups—sets of agents sharing the same observable characteristics—could face different price schedules, allowing incentives to join to vary with group identity and the associated externalities. Tiered pricing means that agents within a group are offered a menu of different levels of access at different prices. Level of access could simply be the probability of becoming a member but—depending on the application—it could also capture factors such as a higher quality of service or additional membership privileges. The key idea behind tiered pricing in our framework is that it allows the designer to achieve a more valuable composition of members by relying on self-selection to screen for unobserved characteristics associated with membership externalities.

[1] Examples include NFT projects such as Bored Ape Yacht Club, Pudgy Penguins, SupDucks, Tycoon Tigers, and 1337 Skulls, all of which have actively fostered community engagement and co-creation around their tokens (see Kaczynski and Kominers (2024)).

Our main technical result connects the number of distinct membership tiers offered by the optimal mechanism to the complexity of the structure of externalities. We measure complexity as the dimension of the image of an operator relating an allocation rule to one group to the value generated for all other groups. In particular, a single price is optimal for a group that does not exert any allocative externalities. When a group exerts a uniform externality (i.e., its magnitude is the same for all agents), two prices may be needed for that group: a high price with the highest level of access, and a discounted price with lower level of access. Finally, when agents' externalities within a given group vary arbitrarily with their types and have heterogeneous effects on other groups, the number of membership tiers required for optimality may be as high as the number of groups plus one.

The main contribution of our paper is to argue that membership design provides a unifying lens through which a diverse set of allocation problems can be analyzed. First, we explore the problem faced by an artist or sports club selling tickets to their events. Without concerns about the composition of the audience, it is optimal to charge a single monopoly price. However, if the atmosphere at the event depends on the presence of the most dedicated fans (e.g., screaming teenagers at a Taylor Swift concert, or diehard fans cheering for their team at a football game), then our framework indicates that it may become optimal to sell at a below-market-clearing price with rationing. We show that using lotteries is optimal when the fans that contribute the most to the atmosphere cannot be easily identified based on observables and have relatively low ability to pay. This may help explain why pop stars like Taylor Swift and sports associations like UEFA sometimes sell tickets to their events at below-market-clearing prices via lotteries—appearing to leave money on the table. Our explanation suggests that such sales mechanisms may actually be revenue-maximizing, even absent any dynamic or reputational considerations.

Second, we consider the primary sales process of a non-fungible token (NFT) collection. NFTs are digital ownership records that are often associated to imagery or other media; they are often intentionally sold at low prices or given away for free, while later reaching resale prices in the hundreds or even thousands of dollars. We provide an explanation for this phenomenon by recognizing that NFTs often serve as gateways to communities whose value is determined by who is a member. The resulting externality implies that different groups of agents should be charged different prices, depending on their desirability as potential members. We show how the pattern of across-group externalities determines whether prices charged to different groups behave like substitutes or complements. Additionally, this analysis lets us recover some of the classical insights about pricing access to traditional buyer-seller platforms.

Third, we analyze the optimal design of a token-based network, focusing on what determines the value of the network. We examine the tension between securing participation by developers who improve the network's functionalities and limiting the congestion externality generated by other users. We show that when the tokens can be eventually traded in a secondary market, getting the network off the ground may require allocating tokens at prices significantly below the expected resale price—sometimes, taking the form of what are called “free mints”—but with vesting.

Finally, we consider the problem faced by a college designing its admissions procedure in the face of diversity considerations. We show that affirmative action (conditioning on the applicants’ background characteristics) allows the college to implement the optimal student body composition via a system of merit and diversity scholarships. However, if affirmative action is not allowed, it may become optimal to bundle admission and funding decisions by forcing students to choose between non-funded traditional admission versus admission that is partially- or fully-funded but rationed. In this way, the college relies on correlations between willingness to pay and background characteristics to implement the desired student body composition through self-selection.

Beyond the ability to explain certain real-life phenomena, our model provides a simple but flexible conceptual framework for thinking about optimal design and policy in contexts in which externalities due to membership structure play an important role.

1.1 Related literature

Our paper is most closely related to a sizeable literature within mechanism design studying allocative externalities. Broadly, two classes of models have been considered. The first class of models, introduced by Jehiel, Moldovanu, and Stacchetti (1996, 1999), studies allocation to a finite number of agents exerting fully flexible individual-specific externalities that are typically assumed to be unobserved. Such a model is challenging to analyze and often leads to negative results (see, e.g., Jehiel and Moldovanu, 2001). A more recent class of models—studied by Kang (2020), Pai and Strack (2022), Ostrizek and Sartori (2023), and Akbarpour ⓡ Budish ⓡ Dworczak ⓡ Kominers (2024)—focuses instead on a large population of agents contributing (in a potentially heterogeneous way) to a single aggregate externality (which often enters the designer’s objective function additively). Our framework sits in between these two approaches: Relative to the first class of models, we impose the simplifying assumption that externalities depend only on compositional effects (in a model with a continuum of agents). Relative to the second class of models, we allow for a much richer pattern of externalities that depend on group identity and interact with individuals’ willingness to pay. The resulting set of applications also differs: the first class of models is ideal for studying a small number of strategically interacting agents (e.g., the problem of selling nuclear weapons in Jehiel et al., 1996); the second class of models is ideal for studying problems such as regulating the sale of goods that generate externalities affecting *all* agents (e.g., goods whose consumption contributes to climate change, as in Kang, 2020; Pai and Strack, 2022); and our framework is geared towards settings in which agents care about the characteristics of other agents receiving the allocation, such as in the context of selling tickets to events, determining access to digital communities, building multi-sided platforms, or designing policies with compositional effects such as school admission.

Allocative externalities similar to the ones we study appear in the classical literature on two-sided buyer–seller platforms (see Jullien, Pavan, and Rysman (2021) for a comprehensive survey). Most papers in this literature do not adopt a mechanism-design approach,[2] and hence do not study the use of rationing (or lotteries) that play a key role in our results.

The puzzle of why performers often sell tickets to their events at below-market-clearing prices has attracted considerable attention in economics.[3] Becker (1991) put forward an explanation by arguing that willingness to pay for goods consumed in public (such as going to a concert) may directly depend on demand of others (e.g., via image concerns); Mortimer, Nosko, and Sorensen (2012) observed that artists may be motivated by generating demand for complementary products (e.g., recorded music); Che, Gale, and Kim (2013) showed that rationing arises when agents are budget-constrained and the seller cares about their utility; finally, Loertscher and Muir (2022) demonstrated that lotteries may arise in a standard monopolist's problem with capacity constraints whenever ironing would be used in the canonical Myerson (1981) model. To the best of our knowledge, our explanation in terms of allocative externalities and associated concerns related to the composition of the audience is new to this literature.[4]

Finally, our perspective on NFTs as gateways to a community of owners follows a common interpretation—and associated business models—in the NFT industry (see, e.g., Kaczynski and Kominers (2021, 2024); Cassatt (2023); Kraski and Shenkarow (2023)); as such, our framework provides practical advice for how NFT project creators should conceptualize their primary sales. This is complementary with articles that have explored the Veblen-good view of NFTs such as Oh, Rosen, and Zhang (2023). Through an empirical study, Oh et al. (2023) showed that NFT markets exhibit many of the properties predicted by an extension of the Becker (1991) model of consumption complementarities.

2 Model

A *designer* allocates access to a platform or community to a unit mass of agents (potential members). Each agent is characterized by a *private type* θ and a publicly observable *label* i. An agent's *level of access* to the platform is described by a variable $x \in [0, 1]$, and the agent can be charged a transfer $t \in \mathbb{R}$. Our baseline interpretation of x is that it represents the probability of becoming a member but it could also reflect the quality of access to the platform (e.g., which features are available to the agent). To focus on the novel aspects of our framework, we assume that the designer does not incur marginal

[2]Some notable exceptions include (but are not limited to) the work of Damiano and Li (2007) and Gomes and Pavan (2021); however, these papers focus on the optimal matching between the two sides of the market.

[3]Budish and Bhave (2023) provide a discussion and an empirical application (see also, e.g., Baliga and Ely (May 8, 2013); Arslan, Tereyağoğlu, and Yılmaz (2023)).

[4]That said, the idea that ticket pricing impacts the composition of attendees is well recognized in the industry; for example, a US Government Accountability Office (2018, p. 8) report on ticket pricing and resale listed the following among the reasons for ticket underpricing (emphasis in original):

> **Audience mix.** Some artists prefer to have the most enthusiastic fans at their shows, rather than just those able to pay the most, especially in the front rows, where tickets are generally the most expensive.

costs for allowing agents' access and does not face an explicit capacity constraint—an assumption that is most appropriate for the design of digital communities.[5]

Conditional on label i, agents' types have a distribution with a cumulative distribution function F_i and a continuous density f_i fully supported on an interval $[\underline{\theta}_i, \bar{\theta}_i]$. There are finitely many labels $i \in I$, and each group i has mass $\mu_i > 0$, with $\sum_{i \in I} \mu_i = 1$.

The utility an agent receives from access depends on that agent's type and the *composition* of other agents using the platform. Specifically, an agent with label i and type θ has willingness to pay for membership equal to

$$\theta v_i,$$

where v_i is common to all agents in group i. The simplifying assumption here is that heterogeneity in willingness to pay can be decomposed into a vertical component captured by the type θ and a horizontal component that only depends on the label i. The value v_i is determined by the expected externalities exerted by all members of the community. Conceptually, each member has some unobserved "true" externality on agents in group i. Because an externality an agent exerts on others does not affect that agent's payoff, it cannot be elicited by any incentive-compatible mechanism.[6] However, the designer can form the best possible "forecast" of the true externality based on characteristics that are observable (the label i) or elicitable by a mechanism with transfers (the type θ). Thus, we let

$$e_{j \to i}(\theta)$$

denote the *expected* externality exerted by an agent with label j and type θ onto any agent in group i. The function $e_{j \to i}(\theta)$ is assumed to be continuous and bounded for all $i, j \in I$. Finally, if $x_j(\theta)$ denotes the probability of access for an agent (j, θ), we assume that

$$v_i \equiv v_i \left(\sum_{j \in I} \int e_{j \to i}(\theta) x_j(\theta) \mu_j dF_j(\theta) \right),$$

where, with slight abuse of notation, $v_i : \mathbb{R} \to \mathbb{R}_+$ is some strictly increasing upper semi-continuous function, for all $i \in I$. We put no sign restrictions on $e_{j \to i}(\theta)$, i.e., agents can exert both positive and negative externalities on one another, depending on group identity and type. We assume utilities are quasi-linear in the monetary transfer and linear in probability of access: Allocation (x, t) gives an (i, θ) agent utility

$$\theta v_i x - t.$$

We assume that the designer maximizes an objective function—such as revenue or agents' welfare—over all direct-revelation mechanisms, subject to individual-rationality

[5]These features could be added to the model without changing our results in any substantial way.

[6]Jehiel et al. (1996) provide a formal statement and proof of this observation.

and incentive-compatibility constraints.[7] Following the standard approach in the mechanism design literature (Myerson, 1981), we can characterize feasible mechanisms through the envelope formula for transfers and monotonicity conditions on the allocation rules, giving rise to the following optimization problem for the designer:

$$\max_{x_1,\ldots,x_{|I|}} \left\{ \sum_{i\in I} \int v_i \left(\sum_{j\in I} \int e_{j\to i}(\theta) x_j(\theta) \mu_j dF_j(\theta) \right) W_i(\theta) x_i(\theta) \mu_i dF_i(\theta) \right\} \tag{1}$$
$$\text{s.t. } x_i(\theta) \text{ is non-decreasing in } \theta, \text{ for all } i \in I,$$

where $W_i : [\underline{\theta}_i, \bar{\theta}_i] \to \mathbb{R}$ is a continuous function capturing the designer's objective. For example, setting $W_i(\theta) = J_i(\theta) := \theta - (1 - F_i(\theta))/f_i(\theta)$ (the virtual surplus function) corresponds to maximizing revenue; while setting $W_i(\theta) = h_i(\theta) := (1 - F_i(\theta))/f_i(\theta)$ (the inverse hazard rate) corresponds to maximizing agents' total surplus (assuming that the lowest type within each group receives zero utility).

We note from expression (1) that the presence of externalities significantly complicates the designer's problem. Assigning access $x_i(\theta)$ to an agent allows the designer to realize the payoff $W_i(\theta)$ from allocating to that agent. At the same time, however, it affects the composition of agents on the platform and thus potentially changes the values that all other agents have for joining. Mathematically, the problem is non-linear in the allocation rule. We purposefully constructed our framework to have a very simple optimal mechanisms when there are no externalities. Indeed, if $v_i(\cdot)$ is constant for all $i \in I$, it is well-known that, regardless of the properties of the function $W_i(\theta)$, the designer can achieve full optimality by simply posting a price p_i for allocation $x = 1$ to each group i. Thus, tiered membership and lotteries can only arise in our setting due to allocative externalities.[8]

3 Results

Define a linear operator E^i from the set of (non-decreasing) allocation rules on $\mathrm{supp}(F_i)$ to $\mathbb{R}^{|I|}$ by

$$E^i_j(x_i) := \int e_{i\to j}(\theta) x_i(\theta) dF_i(\theta), \quad j = 1, \ldots, |I|.$$

Next, let the dimension of the operator E^i be

$$\dim(E^i) := \dim(\mathrm{Im}(E^i)).$$

That is, the dimension of the operator E^i is simply the dimension of its image as a set in the Euclidean space $\mathbb{R}^{|I|}$. For a simple example, suppose that agents in group i do

[7]Restricting attention to direct-revelation mechanisms in our framework is without loss of generality because the revelation principle applies.

[8]Of course, tiered pricing also arises if agent's values or seller's costs are non-linear in the allocation x, as in the celebrated product-quality model of Mussa and Rosen (1978). We turned off these forces in order to focus on the role of externalities.

not exert any externalities on agents in any group j: $e_{i\to j}(\theta) = 0$ for all j and θ. Then, E^i has dimension 0. Next, suppose that externalities do not depend on an agent's type: $e_{i\to j}(\theta) = e_j$, for all j and θ. Then, E^i has dimension 1. When $e_{i\to j}(\theta) = e^1_j\,\theta + e^0_j$, for all j and θ, E^i has dimension 2. Of course, the dimension of E^i is bounded above by the number of groups $|I|$.

Returning to our interpretation of $e_{i\to j}(\theta)$ as the *expected* externality that an agent with type θ in group i exerts on agents in group j, we can relate the dimension of E^i to the sophistication of the econometric model that the designer could be using in practice to estimate the dependence of the externalities on labels and types (e.g., after observing true externalities for a sample). For example, a linear regression of externalities on type θ within each group would result in the estimated operator E^i having dimension 2 (unless the regression co-efficient would turn out to be zero, in which case the dimension would be 1). Fitting a $k-$degree polynomial would make E^i have dimension $k+1$, while a fully non-parametric regression would most likely result in an $|I|-$dimensional E^i.

We will show that the dimension of E^i is closely connected to *how many prices* the optimal mechanism may be required to use. To that end, we define a class of mechanisms that we refer to as k-tiered pricing mechanisms:

Definition 1. An allocation rule $x_i(\theta)$ for group i is a *k-tiered pricing mechanism* if $x_i(\theta)$ takes on at most k non-zero values on $\mathrm{supp}(F_i)$.

A k-tiered pricing mechanism can be implemented by offering agents in a given group at most k different prices for k different levels of access. For example, a 1-tiered pricing mechanism could be a posted-price mechanism, which we define as offering a single price for full access ($x = 1$). But a 1-tiered pricing mechanism could also allow agents to access the platform for free (or at a positive price) with some interior probability. Similarly, a 2-tiered pricing mechanism could allow agents to choose between a low price that offers a low level of access and a higher price that offers full access.

Our main technical result establishes that the optimal mechanism is a k-tiered pricing mechanism for every group, where k can be related to the complexity of the externalities as measured by the dimension of the operator E^i.

Theorem 1. *It is optimal to use a k^i-tiered pricing mechanism for each group i, where $k^i = \dim(E^i) + 1$ and the highest-price tier provides full access ($x = 1$). Moreover, the optimal mechanism uses at most $|I| + K$ prices in total, where $K = \dim(\sum_{i\in I} \mu_i E^i) + 1$.*

The proof of Theorem 1 can be found in Appendix A.1. The intuition for the theorem is simple: While problem (1) faced by the designer is non-linear in the allocation rule x_i for group i, it would become linear if we fixed (and treated as exogenous) the values v_j for all groups. Moreover, fixing the value v_j while letting $x_i(\theta)$ vary corresponds to imposing a single linear constraint on $x_i(\theta)$ since, by assumption, the value v_j only depends on the average externality exerted by members from group i. As a result, the auxiliary problem in which we only optimize over the allocation rule $x_i(\theta)$ for a single group subject to inducing the same target values v_j for all groups j is a linear problem with $|I|$ linear constraints. Furthermore, the dimension of the operator E^i

determines whether some of these constraints are co-linear and hence can be dropped. Once we represent the auxiliary problem as a linear program with $\dim(E^i)$ active linear constraints, an application of an extension of the Carathéodory's theorem yields that the optimal mechanism is a k^i-tiered pricing mechanism with $k^i = \dim(E^i) + 1$ prices. Moreover, the top price can always be assumed to be associated with full access.[9]

The second part of Theorem 1 imposes an upper bound on the total number of prices used by an optimal mechanism. The reason why the total bound is typically much lower than the sum of bounds for all groups i is that our proof strategy of looking at a single group at a time leads to "multiple counting" of the constraints in the auxiliary problem. When all allocation rules can be flexibly chosen, achieving a target value v_j for group j does not require fixing separately the average contributions from each group i; rather, it requires fixing the average contribution *across* all groups, which is a single linear constraint. Thus, we can find an optimal mechanism that charges at most $K = \dim(\sum_{i \in I} \mu_i E^i) + 1$ additional prices (on top of charging a price to each group).

Theorem 1 establishes our main general result which is that the number of prices offered by an optimal mechanism depends on the complexity of externalities. In the remainder of this section, we explore a few special cases of that insight.

First, we confirm an earlier statement that a posted price would be optimal in the absence of externalities.

Corollary 1. *It is optimal to sell to group i at a posted price if group i does not exert externalities on any other group j, i.e., $e_{i \to j}(\theta) = 0$ for all $j \in I$ and θ.*

The absence of externalities exerted by agents from a certain group can be seen as an extreme case; next, we offer a more permissive condition under which the optimal mechanism still sells access at a posted price.

Proposition 1. *Suppose that the designer's objective function $W_i(\theta)$ is strictly increasing in θ, and that $\int W_j(\theta) x_j(\theta) dF_j(\theta) \geq 0$, for all $j \in I$, holds in the optimal mechanism.*[10] *Then, it is optimal to sell to group i at a posted price if the externality that group i exerts on any group j, i.e., $e_{i \to j}(\theta)$, is non-decreasing in θ.*

Under the conditions of Proposition 1, there is an alignment between the designer's objective (for a fixed profile of values v_j) and how the values v_j vary with the allocation—both considerations point towards optimality of a threshold allocation rule that grants access to all agents with a sufficiently high type. Proposition 1 thus implies that the additional prices predicted by Theorem 1 will only be needed when there is a conflict

[9]Note that—except for the case $k^i = 1$—the formulation of Theorem 1 does not preclude the possibility that all agents choose an interior level of access; this is because the price of full access can be set to be so high that no one chooses it.

[10]That last assumption holds automatically for revenue and surplus maximization when types are non-negative (because the required inequality holds for all feasible allocation rules). The assumption could fail in other cases if the designer wanted to allocate membership to some group j to generate positive externalities for other groups at the cost of achieving a negative payoff from group j; in such a case, the designer's objective would become *decreasing* in the value v_j.

between the designer's objective for fixed v_j's and the desire to maximize v_j's over the composition of members.

Finally, we state a simple corollary of Theorem 1 identifying two cases in which two prices—a high price with full access and a low price with lower level of access—are sufficient for achieving optimality.

Corollary 2. *It is optimal to use a 2-tiered pricing mechanism in group i (with the higher price providing full access) if either (1) there is only one group, or (2) group i exerts a uniform externality: $e_{i\to j}(\theta) = e_i(\theta)$ for all $j \in I$.*

4 Applications

4.1 Ticket Sales

We apply our framework to the problem of an artist selling concert tickets or a sports team selling tickets to the game, aiming to maximize revenue. At such events, the atmosphere is often a crucial part of the experience. For example, screaming teenagers are somewhat essential for the atmosphere at a Taylor Swift concert, while the atmosphere at a football match depends on the most dedicated fans cheering for their teams.

We assume that there is a single group of agents (and hence drop the subscript i throughout), reflecting an assumption that the seller cannot easily identify the most enthusiastic fans based on observables.[11] It is then natural to assume that the lower bound of the support of types θ is zero. To model how the atmosphere at an event is generated, we use the externalities that agents exert on others. We assume that the value of externalities is captured by

$$v = v_0 + \int_{\underline{\theta}}^{\overline{\theta}} e(\theta)x(\theta)dF(\theta).$$

We do not explicitly model capacity constraints, such as limited seating availability, but assume that $e(\theta)$ contains a congestion externality that serves as a proxy for capacity considerations. Even though our analysis is valid for a general function $e(\theta)$, we provide the following simple way of deriving $e(\theta)$ from more primitive elements. Suppose that agents are either high-value or low-value audience members (e.g., high-value members cheer for teams at a football game while low-value members sit quietly), and this is not observed. All members generate a congestion externality of -1. However, high-value members additionally generate a positive externality $\Delta > 0$. The high-value members' type is drawn from $F_H(\theta)$, while the low-value members' type is drawn from $F_L(\theta)$, with respective densities $f_H(\theta)$ and $f_L(\theta)$. A fraction μ of all agents are high-value. Then,

[11] Our results go through qualitatively if the seller can observe an imperfect signal, as long as that signal is not very precise. Ticketmaster uses the "Verified Fan" program to sell some concert tickets but the "verification" part is mostly making sure that the potential buyer is a human being (not a bot).

by a simple application of Bayes' rule, we have

$$v = v_0 + \int_{\underline{\theta}}^{\bar{\theta}} \underbrace{\left(\frac{\mu f_H(\theta)}{\mu f_H(\theta) + (1-\mu) f_L(\theta)} \Delta - 1 \right)}_{e(\theta)} x(\theta) dF(\theta).$$

Intuitively, the seller is trying to infer the unobserved value of a fan by conditioning on her type. As a result, the shape of the externality $e(\theta)$ depends on the joint distribution of the unobserved value and willingness to pay.

To simplify exposition, we assume that the virtual surplus function $J(\theta)$ is strictly increasing, and that $e(\theta) < 0$, that is, the negative congestion externality dominates the positive externality even for the most dedicated fans (corresponding to $\Delta < 1$ in the above example). The case when some members have a positive net contribution would make our key insights *easier* to obtain (we comment on this further at the end).

First, relying on Proposition 1, we show the following result:[12]

Result 1. *A posted price is optimal when $e(\theta)$ is increasing.*

When the externality $e(\theta)$ is increasing, selecting the most valuable audience members is aligned with the direct revenue motive of the seller. In our example, the externality is increasing if high-value members have a higher willingness to pay than low-value members (in the monotone likelihood ratio order). This scenario is empirically plausible for cases such as selling tickets to a networking event, where it is natural to expect that higher willingness-to-pay individuals are also more likely to hold professional positions that make them attractive to connect with.

In general, applying Corollary 2 allows us to restrict attention to 2-tiered pricing mechanisms. The seller is thus maximizing

$$\underbrace{\left(v_0 + x_0 \int_{\theta_0}^{\theta_1} e(\theta) dF(\theta) + \int_{\theta_1}^{\bar{\theta}} e(\theta) dF(\theta) \right)}_{V} \underbrace{\left(x_0 \int_{\theta_0}^{\theta_1} J(\theta) dF(\theta) + \int_{\theta_1}^{\bar{\theta}} J(\theta) dF(\theta) \right)}_{R} \tag{2}$$

over $x_0 \in [0, 1]$ and $\theta_0, \theta_1 \in [\underline{\theta}, \bar{\theta}]$, where V represents the base value of the event given the audience composition, and R represents the normalized revenue from selling tickets. Because we assumed that externalities $e(\theta)$ are negative, the term V is *decreasing* in the allocation, while the term R is *increasing* in the allocation (for types for whom virtual surplus is positive)—leading to a trade-off. Our next observation is that a single tier is sufficient if that trade-off admits a certain monotone structure.

Result 2. *If $J(\theta)/(-e(\theta))$ is strictly increasing, then it is optimal to offer a 1-tiered pricing mechanism that offers level of access $x_0^\star$ to all types above $\theta_0^\star$ satisfying*

$$\frac{J(\theta_0^\star)}{-e(\theta_0^\star)} = \frac{R}{V} \quad \textit{and} \quad x_0^\star = \min\left\{ 1, \, \frac{v_0}{2\int_{\theta_0^\star}^{\bar{\theta}} (-e(\theta)) dF(\theta)} \right\},$$

[12]The calculations underlying the results presented in this section can be found in Appendix A.3.

where R and V are functions of $\theta_0^\star$ and $x_0^\star$ as defined in (2) *(with $\theta_1^\star = \bar{\theta}$).*

The ratio of virtual surplus to (the negative of) the externality, $J(\theta)/(-e(\theta))$, can be seen as measuring the trade-off between extracting revenue and preserving the value of the event: Ideally, the seller wants to target agents with high virtual surplus $J(\theta)$ and low (in magnitude) negative externality $e(\theta)$. Under the monotoniticy assumption made in Result 2, the trade-off is always resolved in favor of allocating to all agents with a type above a certain threshold $\theta_0^\star$, pinned down by the first-order condition. Unlike in Result 1, however, it may be optimal to ration access to the event, especially when the negative externalities are substantial.[13] Intuitively, a reduction in congestion can be achieved either through rationing or through a price increase. These two tools, however, are not perfect substitutes: While an increase in price excludes the *marginal* attendee, rationing excludes the *average* attendee. Therefore, the optimal combination of price and rationing balances out the ratio of revenue to externalities of the marginal attendee (excluded by the price) against the ratio of revenue to externalities of the average attendee (excluded by the lottery). Regulating access via lotteries is thus useful precisely when average and marginal characteristics of potential attendees are not aligned: for example, when some lower-willingness-to-pay individuals exert a smaller negative externality than some higher-willingness-to-pay individuals.

The most interesting case arises when the ratio $J(\theta)/(-e(\theta))$ is non-monotone. This can happen when—in the language of our example—the probability of being a high-value member is strongly negatively correlated with willingness to pay in some range of the distribution. For example, blue collar workers who often make for the most dedicated fans of local soccer clubs may have lower ability to pay than wealthy tourists but they are essential to creating the atmosphere at the game. Mathematically, if a majority of high-value members have willingness to pay bounded above by a relatively low number, there might be a sharp drop in $e(\theta)$ around that threshold. This could in turn introduce a non-monotonicity in $J(\theta)/(-e(\theta))$: The ratio could be large for small θ due to low $(-e(\theta))$ and large for high θ due to high $J(\theta)$, but relatively low for intermediate θ. The following result describes the optimal mechanism in this case:

Result 3. *Suppose that $\theta_1^\star$ maximizes* (2) *when x_0 is constrained to be zero. It is then profitable to introduce a second tier with rationing if there exits $\theta_0 < \theta_1^\star$ such that*

$$\frac{J(\theta_1^\star)}{-e(\theta_1^\star)} < \frac{\int_{\theta_0}^{\theta_1^\star} J(\theta)dF(\theta)}{\int_{\theta_0}^{\theta_1^\star}(-e(\theta))dF(\theta)}. \tag{3}$$

If such θ_0 exists, then, in the optimal mechanism, assuming an interior solution,

$$\frac{J(\theta_0^\star)}{-e(\theta_0^\star)} = \frac{J(\theta_1^\star)}{-e(\theta_1^\star)} = \frac{R}{V},$$

where R and V are functions of $\theta_0^\star$, $\theta_1^\star$, and $x_0^\star$ as defined in (2).

[13]Note that a 1-tiered pricing mechanism with rationing is payoff-equivalent to a 2-tiered pricing mechanism in which the top tier has full access but is degenerate (the price is equal to the willingness to pay of the highest type $\bar{\theta}$)—this is why Result 2 is consistent with Theorem 1.

Result 3 predicts that it may be optimal to offer two prices: a low price with rationing and a high price with full access. Intuitively, the low-price option is designed to attract agents who contribute to the atmosphere of the event but do not have a very high willingness to pay.

To understand condition (3), recall that the seller is trying to target agents with a high ratio of virtual surplus to (the negative of) the externality. If the seller is constrained to choosing a single price, she will choose a cutoff type $\theta_1^\star$ at which the ratio $J(\theta_1^\star)/(-e(\theta_1^\star))$ equates R/V, as in Result 2. Moreover, the second-order condition implies that the ratio must be locally increasing around $\theta_1^\star$. However, if $J(\theta)/(-e(\theta))$ is non-monotone, there may exist θ_0 such that agents with types between θ_0 and $\theta_1^\star$ achieve an even higher ratio on average, on the margin (i.e., when they enter with small probability). Thus, the seller finds it optimal to let these types in—with appropriate rationing to limit the congestion externality. At the optimal $\theta_0^\star$, the ratio of virtual surplus to (the negative of) the externality must be the same at $\theta_0^\star$ and $\theta_1^\star$, implying that the seller does not want to further adjust the composition of the rationed and full-access groups.

Because we assumed that externalities are always negative, both cutoffs $\theta_0^\star$ and $\theta_1^\star$ must belong to the region in which virtual surplus is positive. Suppose, instead, that some low-willingness-to-pay agents have a positive net externality—in the context of our example, this corresponds to assuming that $\Delta > 1$ and there is negative correlation between being a high-value member and willingness to pay. Then, condition (3) could be satisfied at θ_0 with $J(\theta_0) < 0$ and $e(\theta_0) > 0$. That is, the seller could decide to use a lottery to allocate to agents who contribute negatively to profits as long as they contribute positively to the willingness to pay of agents selecting the full-access option.

Overall, Results 2 and 3 predict that performers selling tickets to their events may find it optimal to charge below-market-clearing prices and use lotteries to ration. The revenue-maximizing mechanism may take the form of a single lottery, or a combination of a lottery with a high-price full-access option. In the latter case, the lottery allows the seller to strike a better balance between improving the atmosphere at the event (by securing participation of high-value members) and extracting more revenue from agents with high willingness to pay.

4.2 Digital Communities

Next, we apply our model to the problem of designing digital communities where membership is gated by non-fungible tokens (NFTs). We assume that there is a set I of groups of potential members, which one might think of as content creators, advertisers, marquee members (i.e., celebrities), the general public, and so forth. We assume that, for each group, the types θ_i are distributed according to a CDF $F_i(\theta_i) = 1 - (1-\theta_i)^{(1-\beta_i)/\beta_i}$ on the interval $[0, 1]$, for $\beta_i \in (0, 1)$. This family of distributions is conveniently parameterized so that β_i is the optimal monopoly price in the model without externalities (i.e., when $v_i \equiv 1$). To focus on across-group interactions, we assume that externalities are constant in agent's type: $e_{i \to j}(\theta) = e_{i \to j}$ for all $i, j \in I$. Furthermore, we assume that the designer maximizes revenue, and that v_i is the identity function for all i. That is, the value of joining the community is directly proportional to total expected externalities:

$$v_i = \sum_{j \in I} \int_{\underline{\theta}_j}^{\bar{\theta}_j} e_{j \to i} \, x_j(\theta) \mu_j dF_j(\theta).$$

By Proposition 1, the optimal mechanism within each group is a posted price. Assuming an interior solution, so that first-order conditions are valid, we can derive the following formula for the optimal posted prices p_i:

$$p_i = \beta_i v_i - (1 - \beta_i) \sum_{j \in I} e_{i \to j} \mu_j R_j, \tag{4}$$

where $R_j := \frac{p_j}{v_j} \left(1 - \frac{p_j}{v_j}\right)^{(1-\beta_j)/\beta_j}$ is the normalized (by the value v_j) revenue generated from selling to group j.[14] Note that if there are no within-group externalities for some group i ($e_{i \to i} = 0$), then formula (4) provides a closed-form solution for the optimal price for that group; in general, it establishes a fixed point problem that can be solved to determine optimal prices. Note also that, in line with Corollary 1, if group i does not exert any externalities, it is charged the monopoly price $\beta_i v_i$.

For the remainder of this subsection, we exploit comparative statics results that are a consequence of formula (4). We begin with an observation about when it is optimal to charge a group more or less than the monopoly price $\beta_i v_i$.

We refer to $\sum_{j \in I} e_{i \to j} \mu_j R_j > 0$ as the *aggregate revenue externality* exerted by group i.

Result 4. *If group i exerts a positive aggregate revenue externality, i.e., $\sum_{j \in I} e_{i \to j} \mu_j R_j > 0$, then it is charged less than the monopoly price. If it exerts a negative aggregate revenue externality, then it is charged more than the monopoly price.*

In practice, there are groups that are particularly important for the success of a digital community. For example, creators contribute to the community by providing content for its members to consume; celebrities may generate value to other members who are interested in interacting with them; and particularly enthusiastic consumers (as identified, e.g., through their prior engagement in similar communities) may serve as brand evangelists who drive up the visibility and value of membership. As a result, certain groups of agents may be incentivized to join the community through low prices. On the other hand, groups like advertisers may be charged even more than the monopoly price to compensate for the negative externality that their presence exerts.

To conceptualize the connection between the prices for two groups i and j and their corresponding externalities, we consider how the optimal price for group j reacts if the price for group i were perturbed upwards from its optimal level. We say that the prices p_i and p_j *behave like substitutes* if an upward perturbation of p_i leads to a

[14] That is, holding fixed the cutoff type $\theta_j = p_j/v_j$ buying at price p_j, R_j would be the revenue raised by selling to types above θ_j in a model in which $v_j \equiv 1$.

downwards adjustment of p_j (and vice versa). Similarly, we say that p_i and p_j *behave like complements* if an upward perturbation of p_i leads to an upwards adjustment for p_j (and vice versa).

Result 5. *The prices p_i and p_j behave like **substitutes** if*

1. *group i exerts a positive externality on group j, and group j exerts a positive aggregate revenue externality; or*
2. *group i exerts a negative externality on group j, and group j exerts a negative aggregate revenue externality.*

*The prices p_i and p_j behave like **complements** if*

1. *group i exerts a positive externality on group j, and group j exerts a negative aggregate revenue externality; or*
2. *group i exerts a negative externality on group j, and group j exerts a positive aggregate revenue externality.*

For intuition, suppose that group i exerts a positive externality on group j and group j exerts a positive aggregate revenue externality. Suppose that the price p_i increases. As a result, fewer members of group i become members. Since they exert a positive externality onto group j, also fewer members of group j join for a given price p_j. However, since j exerts a positive aggregate revenue externality, it is a valuable group for the community and losing members of this group would cause members in other groups to leave. Thus, it is optimal to lower the price for group j, to prevent its members from leaving, even if it comes at the cost of lowering the revenue extracted from group j.

As a final consideration, suppose some group i exerts a positive externality on all other groups, i.e., $e_{i \to j} > 0$ for all $j \in I$. How should the designer resolve the conflict between generating revenue from selling access to this group and encouraging users of this group to join the community? We can analyze this trade-off by varying the parameter β_i that governs the distribution of willingness to pay within group i. As β_i decreases, the distribution of agents' types shifts toward smaller values. Using formula (4), it is easy to see that $p_i = 0$ when β_i is small enough. Without externalities, the optimal posted price for group i would always be strictly positive (albeit small as $\beta_i \to 0$). In contrast, with externalities, it may be optimal to allow group i to join the community for free.

These observations motivate a number of mechanisms used for NFT allocation in practice: In the context of NFT communities built around emergent digital brands (Kaczynski and Kominers (2021, 2024)), it is common to manage primary sales through some mixture of an open public sale and a private sale for people on an "allow list." The allow-list sale is typically conducted at a lower price than the public sale, and/or with a higher probability of actually being able to purchase the good. And the allow list itself comprises prospective buyers who the founders believe are likely to be particularly value-generating for the community—often assessed as a function of participation in similar communities, or through a survey or other costly and informative signal of interest (see, e.g., Kominers and Roughgarden (September 10, 2022); Kominers and 1337 Skulls Sers (April 6, 2023)).

More broadly, these insights apply to anything we might think of as a club good, or a conspicuous good where the set of owners is important for determining one's own consumption value. At the same time, when we specialize to the case of labels that reflect distinct sides of a multi-sited platform, we can recover classic intuitions on subsidizing marketplace participants with significant cross-platform externalities (see, e.g., Eisenmann, Parker, and Van Alstyne (2006)).

4.3 Token-Based Networks

In this application, we investigate the optimal design of a blockchain project that allocates tokens in the primary market to raise revenue and regulate access to the network. We focus on conditions under which the token can have any value at all (as measured by the price it achieves in the secondary market), the role of vesting in getting the network off the ground, and optimal pricing.

We consider two groups of agents. First, there are potential developers, denoted D, who might actively contribute to building or improving the network. Second, there are agents who are part of the "general public," denoted P, who may decide to hold the token for speculative or investment purposes or use the network's functionalities but do not contribute to its development. We assume that agents belonging to the general public exert a network congestion externality

$$e_P(\theta) = -1$$

which can be interpreted as occupying the network's computational resources.[15] The developers potentially contribute to the network development, in addition to exerting a congestion externality:

$$e_D(\theta) = -1 + e(\theta),$$

where the dependence of $e(\theta)$ on θ reflects potential correlation between willingness to pay and the likelihood of contributing. Finally, we interpret the value v as the resale price of the token that will arise in the secondary market once it opens (hence, v is not indexed by group label i). We assume that the value is given by

$$v = \max\left\{ \sum_{i \in \{D, P\}} \int e_i(\theta) x_i(\theta) \mu_i dF_i(\theta),\ 0 \right\}.$$

Intuitively, the value of the token is a reflection of the network functionality—given by the total expected non-congestion externalities of all developers who hold tokens—net

[15] For example, congestion leads to increased transaction fees on the bitcoin blockchain (Huberman, Leshno, and Moallemi, 2021) or increased gas prices on the Ethereum network.

of the network congestion. Since only developers can provide positive externalities, it is crucial to incentivize them to join. At the same time, giving the token to too many agents undermines the value of the network through congestion—a network that is too congested (relative to its functionalities) becomes useless and has no value.

We interpret θ as capturing an agent's belief about the (unmodeled) long-term resale value of the token. It is then natural to assume that the median of the unconditional distribution of θ is 1. If we assume that tokens can be costlessly liquidated at v once the secondary market opens, then agents with $\theta \geq 1$ are the ones who believe the token will appreciate in value and hence decide to hold it. In contrast, agents with $\theta < 1$ want to sell as soon as the secondary market opens.

We assume that the designer is the network founder who controls the allocation of tokens in the primary market (with total quantity normalized to 1). The founder thus maximizes

$$\sum_{i \in \{D,P\}} \int \left(vJ_i(\theta)x_i(\theta) + v(1 - x_i(\theta)) \right) \mu_i dF_i(\theta),$$

under the assumption that the tokens kept by the founder are valued at v (and do not contribute to the network congestion). The allocation rule $x_i(\theta)$ represents holding tokens, which is required for access to the network. In particular, only agents with $x_i(\theta) > 0$ can contribute to network development.

Solution without vesting. Under our interpretation of θ as the agent's optimism about the future value of the network (relative to the resale price of the token), the designer is constrained to allocation rules satisfying $x_i(\theta) = 0$ for $\theta < 1$ (since agents with $\theta < 1$ will not hold tokens once the secondary market opens). We can then make the following simple observation:

Result 6. *The network token has a strictly positive resale price if and only if*

$$\exists p \geq 1 \text{ such that } \int_p^{\bar{\theta}_D} e_D(\theta) dF_D(\theta) > 0. \tag{5}$$

It is optimal to allocate tokens by posting a single price for the general public, and at most two prices for developers (where there could be rationing at the lower price).

A strictly positive value of the network can be generated if there exists a price weakly above the resale value with the property that—among the developers who decide to buy—the average positive externality $e(\theta)$ exceeds the congestion externality 1. In other words, the founder must attract the high-value developers through self-selection based on optimism about the network value. We view Result 6 as a negative result. Compared to the general public, the beliefs of the well-informed high-skill developers are more likely to be concentrated around 1. Moreover, developers with unrealistically high expectations may not be the best contributors—$e(\theta)$ may be decreasing in θ above $\theta = 1$. At the same time, the founder cannot engage the potentially valuable developers with $\theta < 1$ because they would prefer to sell their tokens as soon as the secondary market opens.

Solution with vesting. Next, we observe that if we allow for vesting (an allocation of tokens with a temporary restriction on reselling them in the secondary market), then the condition from Result 6 can be relaxed. We model vesting simply by allowing the designer to choose any allocation rule, where $x_i(\theta) > 0$ for $\theta < 1$ is now possible due to vesting.

Result 7. *With vesting, the network token has a strictly positive resale price if and only if*

$$\exists p \geq 0 \text{ such that } \int_p^{\bar{\theta}_D} e_D(\theta) dF_D(\theta) > 0. \tag{6}$$

It is optimal to sell tokens to the general public at a single price that strictly exceeds v and to developers at at most two prices; if condition (5) *fails but condition* (6) *holds, then at least one of these prices is below v, involves vesting, and may involve rationing.*

With vesting, the designer can access the potentially high–externality developers with lower beliefs θ about the network value. This expands the set of parameters for which the token can have a strictly positive price.

Rationing schemes do not play a role in determining whether the token can have positive value—in conditions (5) and (6), it is enough to consider the effect of charging different posted prices to developers. However, rationing may play an important role in the revenue-maximizing mechanism. Rationing at a low price with vesting may allow the designer to ensure sufficient participation by high-externality developers with relatively low θ, without putting too much downward pressure on the price she charges for the non-rationed option to more optimistic developers. Unlike in the ticket sales example in Subsection 4.1, it may be optimal to have a rationing scheme with a price of zero for developers, since the designer can still make money by selling to the general public (or by keeping the tokens and reselling them in the secondary market once it opens). In practice, such mechanisms sometimes take the form of "free mints." Free mints may be optimal when $e(\theta)$ exceeds the congestion externality even for θ close to zero. Such a case is plausible if the network founders believe that initially skeptical developers are likely to engage with the network once they gain access to it; this for example might be a reasonable interpretation of promising token grants to developers who commit to building around a new blockchain in advance of launch.

In token-based networks where the goal is to create a decentralized software ecosystem—such as with open source projects—our results suggest it makes sense to reserve a large share of the initial tokens for developers, and implement vesting schedules to incentivize ongoing participation. While our model is static, intuitively, vesting should be calibrated so that token holders' ability to sell kicks in after the ecosystem value has already been well established.

4.4 College Admissions

Our final application is to designing admission policy in the context of a college or other academic program in which students' value for participating (and hence, their willingness to pay) depends on the composition of the student body. Specifically, we assume that students' values for attending are higher when the incoming class is balanced across various characteristics such as socioeconomic background, minority status, geography, or field of interest.

In this setting, on top of the unobserved heterogeneity in willingness to pay parameterized by the type θ, each (prospective) student has a label $i = (b, t)$, where $b \in \mathcal{B}$ captures diversity characteristics (that we refer to as "background") and $t \in \mathcal{T}$ captures observable measures of talent or ability (that we refer to as "test score"). We assume that a student with characteristics (b, t) exerts an externality $e_{(b,\,t)}$ on all other students—for example, students could be benefit from interacting with talented peers. Additionally, students are assumed to value diversity: For any b, let

$$s_b = \sum_{t \in \mathcal{T}} \mu_{(b,t)} e_{(b,\,t)} \int x_{(b,t)}(\theta) dF_{(b,t)}(\theta),$$

be the mass of admitted students with background b, weighted by their externality. Then, a student's willingness to pay for attending the college is θv, where $v = v(s_1, \ldots, s_{|\mathcal{B}|})$ satisfies the usual Inada conditions in each argument. These assumptions imply that—everything else being equal—there is a higher marginal value for adding a student from a background group that is underrepresented. For example, setting $v(s_1, \ldots, s_{|\mathcal{B}|}) = \prod_b s_b^{\mu_b}$ corresponds to assuming that the ideal composition of the student body mimics the population shares μ_b of the respective groups.[16]

The designer (the college) maximizes a weighted sum of revenue (with weight β) and student welfare: $W_i(\theta) = \beta J_i(\theta) + (1-\beta) h_i(\theta)$, where recall that $h_i(\theta) = (1 - F_i(\theta))/f_i(\theta)$ is the inverse hazard rate, which we assume is non-increasing.

The example is a special case of our baseline model except that the value v is now a function of multiple arguments; in Appendix A.4, we explain how the proofs of Theorem 1 and Proposition 1 can be modified to cover this case.

When affirmative action is allowed. When the college is allowed to condition admission decisions on all observable characteristics, we obtain the following result as a corollary of Proposition 1.

Result 8. *When admission decisions can explicitly condition on background characteristics, and $\beta \geq \frac{1}{2}$, it is optimal to use a posted-price mechanism for every group i, with prices varying across the labels i.*

According to Result 8, the optimal admission policy is relatively simple: Each student faces a price for admission that depends on both the measure of their talent t and their background b. This mechanism could be implemented through a constant tuition and a combination of merit and diversity scholarships, along with prohibitively high prices for candidates with low ability t.

[16]See Strack and Yang (2023) for a related analysis of admission policies where compositional considerations are modeled as a hard constraint.

The role of the assumption $\beta \geq \frac{1}{2}$ is to make the college's objective function strictly increasing in θ. For a sufficiently low weight β on revenue, the college's objective function could become decreasing, in which case it would be optimal to drop prices to 0 and admit students using a label-dependent lottery. In this scenario, diversity considerations would lead to higher probabilities of admission for minority groups.

When affirmative action is not allowed. Next, we investigate the optimal admission mechanism when the college is not allowed to explicitly condition on the background characteristics $b \in \mathcal{B}$. Mathematically, this means that the set of labels is taken to be just the values $t \in \mathcal{T}$; the allocation rule $x_t(\theta)$ cannot depend on b directly; and thus

$$s_b = \sum_{t \in \mathcal{T}} \mu_t \int \underbrace{e_{(b,\,t)} \frac{\mu_{(b,\,t)} f_{(b,\,t)}(\theta)}{\mu_t f_t(\theta)}}_{e_{(b,\,t)}(\theta)} x_t(\theta) dF_t(\theta),$$

where we now treat $e_{(b,\,t)}(\theta)$ as the effective externality of group $(b,\,t)$. The key observation is that the "composition externalities" are now a function of the type θ. While θ does not matter for the actual externality, when the college cannot condition admission decisions on background characteristics, willingness to pay becomes useful in inferring a candidate's background. Indeed, $e_{(b,\,t)}(\theta)$ is proportional to the share of candidates with background b among candidates with test score t. As a result, the expected externality may naturally be decreasing in θ for groups who tend to have low baseline willingness to pay (e.g., students coming from disadvantaged backgrounds). This changes the optimal admission policy:

Result 9. *When admission decisions cannot explicitly condition on background characteristics, it is optimal to use a k^t-tiered pricing mechanism for group t, where $k^t \leq |\mathcal{B}| + 1$.*

The main insight of Result 9 is that randomized admission at lower prices may become optimal when affirmative action is not allowed. Moreover, the number of distinct admission options (differing in the price and probability of acceptance) is upper-bounded by the number of distinct background characteristics that matter for assessing diversity.

The intuition for the result is straightforward. Suppose that the college must admit a certain mass of students coming from economically underprivileged backgrounds to ensure a high value v. When affirmative action is available, the college achieves that via targeted admission decisions and scholarships. But when such a policy is not available, the college must rely on self-selection to ensure diversity. Since students coming from economically underprivileged backgrounds tend to have low willingness to pay (due to low ability to pay), the college would have to reduce prices substantially for everyone to ensure enrollment of such students. Randomization relaxes that trade-off: When applying, students can be offered a choice between a high-tuition option and a lottery-assigned low-tuition option. Higher willingness-to-pay students would choose to avoid the additional admission risk and select into the high-tuition option, while low willingness-to-pay students could be admitted via the rationed option with sufficient probability to ensure the desired level of diversity.[17] Note that such a policy necessarily bundles together admission and funding decisions.

Our framework may also help explain the recent admission trends among US colleges. Roughly a decade ago, many colleges moved to stop conditioning admission decisions on standardized test scores because of concerns that this was advantaging more affluent and otherwise better resourced students. More recently, a number of colleges have reversed course on this, reinstating standardized test score consideration because they have found that test scores were less susceptible to socioeconomic bias than other factors such as participation in numerous extracurriculars (see, e.g., Wren (March 28, 2022); Korn (February 5, 2024) and the discussion therein).

Our framework makes it clear that if the distribution of scores t across groups b is heavily skewed towards more advantaged students (e.g., because they have access to private tutoring), then admitting more students with higher t undermines the self-selection mechanism. Thus, the college may find it optimal not to condition admission decisions on t.[18] If we think of the score t as comprising two dimensions—a standardized test score along with a second component accounting for extracurriculars, essays, and so forth—then the preceding logic suggests that colleges would want to adjust the weights on these two components to reflect how they correlate with background characteristics.

References

Akbarpour, M. ⓡ E. Budish ⓡ P. Dworczak ⓡ S. D. Kominers (2024): "An Economic Framework for Vaccine Prioritization," *Quarterly Journal of Economics*, 139, 359–417.

Akbarpour, M. ⓡ P. Dworczak ⓡ S. D. Kominers (2023): "Redistributive Allocation Mechanisms," *Journal of Political Economy (forthcoming)*.

Arslan, H. A., N. Tereyağoğlu, and Ö. Yilmaz (2023): "Scoring a touchdown with variable pricing: Evidence from a quasi-experiment in the NFL ticket markets," *Management Science*, 69, 4435–4456.

Baliga, S. and J. Ely (May 8, 2013): "The Power of Purple Pricing," *Harvard Business Review*.

Bauer, H. (1958): "Minimalstellen von Funktionen und Extremalpunkte," *Archiv der Mathematik*, 9, 389–393.

[17] For a closely related economic mechanism but in the context of allocation under redistributive social preferences, see the work of Dworczak ⓡ Kominers ⓡ Akbarpour (2021) and Akbarpour ⓡ Dworczak ⓡ Kominers (2023).

[18] In our model, it may in fact be optimal to give higher priority to certain intermediate or even low scores t if they are particularly prevalent within some underrepresented groups b. However, (unmodeled) moral-hazard considerations may prevent colleges from implementing admission rules that are non-monotone in test scores (see Sönmez (2013) for discussion of an instance when precisely this moral hazard issue arose in practice).

Becker, G. S. (1991): "A Note on Restaurant Pricing and Other Examples of Social Influences on Price," *Journal of Political Economy*, 99, 1109–1116.

Bonnans, J. F. and A. Shapiro (2000): *Perturbation Analysis of Optimization Problems*, New York, NY: Springer.

Budish, E. and A. Bhave (2023): "Primary-Market Auctions for Event Tickets: Eliminating the Rents of "Bob the Broker"?" *American Economic Journal: Microeconomics*, 15, 142–70.

Cassatt, A. (2023): *Web3 Marketing: A Handbook for the Next Internet Revolution*, Wiley.

Che, Y.-K., I. Gale, and J. Kim (2013): "Assigning Resources to Budget-Constrained Agents," *The Review of Economic Studies*, 80, 73–107.

Damiano, E. and H. Li (2007): "Price Discrimination and Efficient Matching," *Economic Theory*, 30, 243–263.

Doval, L. and V. Skreta (2022): "Constrained Information Design," *Mathematics of Operations Research (forthcoming).*

Dworczak, P. ⓡ S. D. Kominers ⓡ M. Akbarpour (2021): "Redistribution through Markets," *Econometrica*, 89, 1665–1698.

Eisenmann, T., G. Parker, and M. W. Van Alstyne (2006): "Strategies for Two-Sided Markets," *Harvard Business Review*, 84.

Gomes, R. and A. Pavan (2021): "Two-sided Markets, Platforms, and Matching Design," .

Huberman, G., J. D. Leshno, and C. Moallemi (2021): "Monopoly without a Monopolist: An Economic Analysis of the Bitcoin Payment System," *The Review of Economic Studies*, 88, 3011–3040.

Jehiel, P. and B. Moldovanu (2001): "Efficient Design with Interdependent Valuations," *Econometrica*, 69, 1237–1259.

Jehiel, P., B. Moldovanu, and E. Stacchetti (1996): "How (Not) to Sell Nuclear Weapons," *The American Economic Review*, 814–829.

——— (1999): "Multidimensional Mechanism Design for Auctions with Externalities," *Journal of Economic Theory*, 85, 258–293.

Jullien, B., A. Pavan, and M. Rysman (2021): "Two-sided Markets, Pricing, and Network Effects," in *Handbook of Industrial Organization*, Elsevier, vol. 4, 485–592.

Kaczynski, S. and S. D. Kominers (2021): "How NFTs Create Value," *Harvard Business Review.*

——— (2024): *The Everything Token: How NFTs and Web3 Will Transform the Way We Buy, Sell, and Create*, Portfolio.

Kang, Z. Y. (2020): "Optimal Indirect Regulation of Externalities," *Available at SSRN 3586050.*

——— (2023): "The Public Option and Optimal Redistribution," Working paper.

Kominers, S. D. and 1337 Skulls Sers (April 6, 2023): "A new NFT launch strategy: The wave mint," *a16z crypto.*

Kominers, S. D. and T. Roughgarden (September 10, 2022): "NFT Sales: Clearing the Market, Avoiding Gas Wars," *a16z crypto*.

Korn, M. (February 5, 2024): "Dartmouth Reinstates SAT Requirement in First for Ivy League," *Wall Street Journal*.

Kraski, J. and J. Shenkarow (2023): *The Future of Community: How to Leverage Web3 Technologies to Grow Your Business*, Wiley.

Loertscher, S. and E. V. Muir (2022): "Monopoly Pricing, Optimal Randomization, and Resale," *Journal of Political Economy*, 130, 566–635.

Mortimer, J. H., C. Nosko, and A. Sorensen (2012): "Supply responses to digital distribution: Recorded music and live performances," *Information Economics and Policy*, 24, 3–14, the Economics of Digital Media Markets.

Mussa, M. and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic Theory*, 18, 301–317.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–78.

Oh, S., S. Rosen, and A. L. Zhang (2023): "Digital Veblen Goods," *Working Paper.*

Ostrizek, F. and E. Sartori (2023): "Screening While Controlling an Externality," *Games and Economic Behavior*, 139, 26–55.

Pai, M. and P. Strack (2022): "Taxing Externalities Without Hurting the Poor," *Available at SSRN 4180522*.

Sönmez, T. (2013): "Bidding for army career specialties: Improving the ROTC branching mechanism," *Journal of Political Economy*, 121, 186–219.

Strack, P. and K. H. Yang (2023): "Privacy Preserving Signals," Working paper.

Szapiel, W. (1975): "Points Extrémaux dans les Ensembles Convexes(I). Théorie Générale," *Bulletin de l'Académie Polonaise des Sciences. Sériedes Sciences Mathématiques, Astronomiques et Physiques*, 23, 389–393.

US Government Accountability Office (2018): "Event Ticket Sales: Market Characteristics and Consumer Protection Issues," `https://www.gao.gov/assets/gao-18-347.pdf`.

Wren, K. (March 28, 2022): "Q&A: Stuart Schmill on MIT's decision to reinstate the SAT/ACT requirement," *MIT News Office*.

A Proofs

A.1 Proof of Theorem 1

We first note that an optimal solution to the designer's problem exists, by a standard argument (the objective function is upper semi-continuous on a compact set, when we endow the space of non-decreasing allocation rules with the L^1 topology).

Let $N = |I|$. We consider an auxiliary problem for the designer of maximizing her objective over the allocation rule $x_i(\theta)$ for group i only, keeping the other allocation rules fixed, subject to preserving the values v_j for all groups $j \in I$. Formally, we consider the problem

$$\max_{x_i} \quad \int W_i(\theta)x_i(\theta)dF_i(\theta) \tag{7}$$

$$\text{s.t.} \;\; v_j = v_j\left(\sum_{k\in I}\int e_{k\to j}(\theta)x_k(\theta)\mu_k dF_k(\theta)\right), \qquad \text{for all } j \in I$$

$$x_i(\theta) \text{ is non-decreasing.}$$

Because we are keeping fixed all allocation rules other than the one for group i, we can rewrite the value-preservation constraint as

$$c_j = \int e_{i\to j}(\theta)x_i(\theta)dF_i(\theta), \quad \text{for all } j \in I,$$

for some constants c_j, or in vector notation

$$c = E^i(x_i).$$

Let $k^i = \dim(E^i) + 1$. After a change of basis (multiplying the left- and right-hand side of the above equality by an appropriate matrix), we can assume that E^i has an image contained in $\mathbb{R}^{k^i-1}$, and that $c \in \mathbb{R}^{k^i-1}$ (the remaining coefficients are always zero, so they can be dropped).

We now restate (in a slightly modified form) a result from Kang (2023),[19] which is an extension of Carathéodory's Theorem to infinite-dimensional spaces based on the contributions of Bauer (1958) and Szapiel (1975):

Theorem 0. *Let $\mathcal{X}$ be a convex, compact set in a locally convex Hausdorff space, and let $E : \mathcal{X} \to \mathbb{R}^{k-1}$ and $W : \mathcal{X} \to \mathbb{R}$ be continuous affine functions. Suppose that $L \subset \mathbb{R}^k$ is closed and convex, and that $E(\mathcal{X}) \cap L \neq \emptyset$. Then, there exists $x^\star \in \mathcal{X}$ such that $W(x^\star) = \max_{x\in\mathcal{X}:\, E(x)\in L} W(x)$ and*

$$x^\star = \sum_{j=1}^{k} \alpha_j x_j, \quad \sum_{j=1}^{k} \alpha_j = 1, \quad \alpha_j \geq 0, \qquad \textit{and } x_j \in \mathrm{Ex}(\mathcal{X}),\ \forall j = 1, \dots, k,$$

where $\mathrm{Ex}(\mathcal{X})$ denotes the set of extreme points of $\mathcal{X}$.

Intuitively, Theorem 0 states that the problem of maximizing a linear objective function over a compact convex set $\mathcal{X}$ subject to $k-1$ linear constraints admits a solution that is a convex combination of at most k extreme points of $\mathcal{X}$.

We can apply Theorem 0 to our problem by setting $\mathcal{X}$ to be the set of non-decreasing allocation rules (which is convex and compact when we endow it with the L^1 topology), E to be E^i, L to consist of the single point $c \in \mathbb{R}^{k^i-1}$, and $W(x) = \int W_i(\theta)x(\theta)dF_i(\theta)$. By Theorem 0, there exists a solution to the auxiliary problem that is a convex combination

[19] See also Doval and Skreta (2022) for a similar argument used in the context of information design.

of at most k^i extreme points of the set of non-decreasing allocation rules; that is, there exists a solution to (7) that takes the form

$$x_i(\theta) = \sum_{j=1}^{k^i} \alpha_j \mathbf{1}_{\{\theta \geq \theta_j\}},$$

for some weights $\alpha_j \geq 0$ adding up to 1 and cutoff types θ_j, $j = 1, .., k^i$. Applying the same argument for all $i \in I$ establishes existence of a solution in which there is a k^i-tiered pricing mechanism for every group. Additionally, the top tier is granted full access: $x_i(\theta) = 1$ for $\theta \geq \theta_j$ (which does not preclude the possibility that $\theta_j = \bar{\theta}$ in which case removing the top tier from the mechanism would not affect the designer's expected payoff).

To prove the second part of the theorem, we consider an auxiliary problem in which we fix all the externality values v_i but reoptimize over all allocation rules jointly:

$$\begin{aligned}
\max_{x_1,\ldots,x_N} \quad & \sum_{i\in I} v_i \int W_i(\theta) x_i(\theta) \mu_i dF_i(\theta) \\
\text{s.t. } \; & v_i = v_i\left(\sum_{j \in I} \int e_{j\to i}(\theta) x_j(\theta) \mu_j dF_j(\theta)\right) && \text{for all } i \in I, \\
& x_i(\theta) \text{ is non-decreasing} && \text{for all } i \in I.
\end{aligned}$$

The problem can be rewritten, for some vector of constants $c \in \mathbb{R}^N$, as

$$\begin{aligned}
\max_{x_1,\ldots,x_N} \quad & \sum_{i\in I} v_i \int W_i(\theta) x_i(\theta) \mu_i dF_i(\theta) \\
\text{s.t. } \; & c = \sum_{j\in I} \mu_j E^j(x_j), \\
& x_i(\theta) \text{ is non-decreasing} && \text{for all } i \in I.
\end{aligned}$$

Let $K = \dim(\mathrm{Im}(\overline{E})) + 1$ where $\overline{E}(x_1, \ldots, x_N) := \sum_{j\in I} \mu_j E^j(x_j)$. After a change of basis, we can assume that $\overline{E}$ has an image contained in $\mathbb{R}^{K-1}$ and $c \in \mathbb{R}^{K-1}$.

We know from the first part of the proof that there exists an optimal solution in which each allocation rule x_i can be written as $x_i(\theta) = \sum_{j=1}^{N} \alpha^i_j \mathbf{1}_{\{\theta \geq \theta^i_j\}}$, for some weights $\alpha^i_j \geq 0$ (adding up to 1) and cutoff types θ^i_j, $j = 1, .., N+1$ (where we have used the fact that $k^i \leq N+1$). Fixing the cutoff types, and the value $\bar{v}$ of the optimal payoff for the designer, the optimal weights α^i_j must satisfy the following system of inequalities:

$$\bar{v} = \sum_{i \in I} \sum_{j=1}^{N+1} \alpha_j^i \, v_i \, \tilde{W}_j^i$$

$$c_k = \sum_{i \in I} \sum_{j=1}^{N+1} \mu_i \, \alpha_j^i \, \tilde{e}_{i,j,k}, \quad k = 1, \ldots, K-1$$

$$1 = \sum_{j=1}^{N+1} \alpha_j^i, \quad \forall i \in I$$

$$0 \le \alpha_j^i, \quad \forall j = 1, \ldots, N+1, \, i \in I,$$

for some constants $\tilde{W}_j^i$ and $\tilde{e}_{i,j,k}$ that do not depend on the weights α_j^i. Moreover, any set of weights that satisfies this system of equations defines an optimal solution. A solution to the system exists (because there exists an optimal solution to the designer's problem that takes this form). The system of linear equations has $N(N+1)$ variables and $N+K$ equality constraints. By the fundamental theorem of linear algebra, there exists a solution to the system of equations in which (at most) $N+K$ weights are non-zero. It follows that setting $N+K$ cutoff types (equivalently, prices) optimally allows the designer to achieve full optimality in her optimization problem.

A.2 Proof of Proposition 1

Consider the auxiliary problem of reoptimizing over x_i subject to keeping the expected externalities exerted by group i on other groups at least as large as in the original solution. Imposing a weak inequality (as opposed to an equality, as in the proof of Theorem 1) is sufficient because of our assumption that—at optimum—the designer's payoff is non-decreasing in the values v_j (we also rely on our assumption that all functions v_i are strictly increasing). The problem thus becomes:

$$\begin{aligned} \max_{x_i} \quad & \int W_i(\theta) x_i(\theta) dF_i(\theta) \\ \text{s.t. } \; c_j \le & \int e_{i \to j}(\theta) x_i(\theta) dF_i(\theta), \quad \text{for all } j \in I, \\ & x_i(\theta) \text{ is non-decreasing,} \end{aligned} \tag{8}$$

for some constants c_j. We can assign Lagrange multipliers $\lambda_j \ge 0$ to the inequality constraints,[20] and then the problem is to maximize a Lagrangian

$$\int \left(W_i(\theta) + \sum_{j \in I} \lambda_j e_{i \to j}(\theta) \right) x_i(\theta) dF_i(\theta).$$

[20] The existence of Lagrange multipliers follows from Theorem 2.165 in Bonnans and Shapiro (2000) as long as c_j, for each $j \in I$, is in the interior of all possible values that the right-hand side can take as $x_i(\theta)$ varies. Whenever this is not the case, $x_i(\theta)$ must take the form of a threshold allocation rule, which is what we want to show.

By assumption that $W_i(\theta)$ is strictly increasing and $e_{i\to j}(\theta)$ is non-decreasing in θ, for all j, $W_i(\theta) + \sum_{j\in I} \lambda_j e_{i\to j}(\theta)$ is strictly increasing in θ for any set of Lagrange multipliers. It follows that there exists an optimal solution in which $x_i(\theta)$ is a threshold (one-price) allocation rule.

A.3 Supplementary Materials for Section 4.1

In this appendix, we provide supplementary calculations for Results 2 and 3.

First, we note that the first-order conditions for optimality of $\theta_0^\star$ and $\theta_1^\star$ for problem (2) read

$$\frac{J(\theta_0^\star)}{-e(\theta_0^\star)} = \frac{J(\theta_1^\star)}{-e(\theta_1^\star)} = \frac{R}{V}.$$

Thus, we cannot have two non-degenerate membership tiers when $J(\theta)/(-e(\theta))$ is strictly increasing. In particular, this shows that a 1-tiered pricing mechanism is optimal under the assumption of Result 2. The rest of Result 2 follows from standard analysis of first-order conditions for maximizing objective (2) when θ_1 is set to $\bar{\theta}$. The derivative of objective (2) with respect to x_0 is

$$x_0 \int_{\theta_0}^{\bar{\theta}} e(\theta)dF(\theta) \int_{\theta_0}^{\bar{\theta}} J(\theta)dF(\theta) + \left(v_0 + x_0 \int_{\theta_0}^{\bar{\theta}} e(\theta)dF(\theta)\right) \int_{\theta_0}^{\bar{\theta}} J(\theta)dF(\theta).$$

Solving for x_0 yields the second condition characterizing the solution in Result 2.

To obtain Result 3, suppose that $\theta_1^\star$ is the optimal cutoff type if the designer restricts attention to posted-price mechanisms. We can then consider how the expected payoff of the designer changes under the following perturbation of the mechanism: For some type $\theta_0 < \theta_1^\star$, allocate access to types $\theta \in [\theta_0, \theta_1^\star]$ with some small probability $\epsilon > 0$. This perturbation is profitable for small enough ϵ if

$$\int_{\theta_0}^{\theta_1^\star} e(\theta)dF(\theta) \cdot R + V \cdot \int_{\theta_0}^{\theta_1^\star} J(\theta)dF(\theta) > 0 \iff \frac{J(\theta_1^\star)}{-e(\theta_1^\star)} < \frac{\int_{\theta_0}^{\theta_1^\star} J(\theta)dF(\theta)}{\int_{\theta_0}^{\theta_1^\star} (-e(\theta))dF(\theta)},$$

where we have relied on the first-order condition for $\theta_1^\star$ to substitute $R/V = J(\theta_1^\star)/(-e(\theta_1^\star))$.

A.4 Supplementary Materials for Section 4.4

In this appendix, we briefly explain how our arguments extend to the case when the value functions v_i are common to all groups, $v_i = v$ for all $i \in I$, but v is a function of multiple arguments: $v = v(s_1, \ldots, s_K)$, where each s_k depends linearly on the allocation rule to each group.

To extend the first part of Theorem 1 to this case (in order to obtain Result 9), note that we can modify the proof of Theorem 1 by fixing the values of each of the inputs s_k. Doing so implies that the auxiliary problem (7) of optimizing over an allocation rule for a single group will have K linear constraints. Thus, by the same argument as in the proof of Theorem 1, we need at most $K+1$ prices in the optimal mechanism. This gives us Result 9.

Result 8 is implied by an analogous modification applied to the proof of Proposition 1 (instead of imposing a constraint on the values v_j for all groups j, we impose a constraint on the inputs s_k to the single value function v). Note that the designer's payoff is non-decreasing in the value v because—under the objective of weighted revenue and surplus—the designer's payoff from every group is always non-negative.