Collective Upkeep

Erik Madsen[*]   Eran Shmaya[†]

July 27, 2024

Abstract

We design mechanisms for maintaining public goods which require periodic non-monetary contributions. Utilitarian welfare is maximized by concentrating contributions among low-cost group members, but such policies generally induce some members to leave the group or misreport their preferences. To forestall exit, contributions should be shifted from members with intermediate costs to some high-cost members. To deter misreporting, members should be screened using up to two membership tiers, which reward larger contributions with increased access to the good. We apply our results to the design of crowd-sourced recommendation engines hosted by platforms such as Netflix and TikTok, which function as public goods supported by user feedback about new content.

1   Introduction

Social and economic groups commonly band together to provide public goods by pooling contributions from individual group members. A collective action problem arises whenever the group benefits of these contributions exceed their private benefits to the

---

[*]Department of Economics, New York University (Email: `emadsen@nyu.edu`).
[†]Department of Economics, Stony Brook University (Email: `eran.shmaya@stonybrook.edu`).

contributor. In these situations, institutions are needed to coordinate contributions and ensure that group members internalize their collective benefits.

A large literature has studied the design of such institutions when group members make monetary contributions. In some contexts, however, members may make "in-kind" or other nonmonetary contributions. An important example involves crowd-sourced recommendation engines maintained by online platforms such as Netflix and TikTok. Those platforms rely on customer feedback about new content to make informed recommendations to other members. Feedback is costly for customers to provide, both because it requires consuming content of uncertain quality and because evaluating one's experience takes time and energy. Further, some of the value from this feedback accrues to other customers, who benefit from improved recommendations. Feedback is therefore effectively a contribution to a public good, and a platform wishing to efficiently elicit it must solve a collective action problem.

We study the design of mechanisms for solving such problems. We formalize our analysis through a model of *collective upkeep* (Section 2). In our model, a machine provides services to a large group of agents over a long time horizon. The machine periodically breaks, rendering it useless until sufficient contributions have been collected to fix it. All agents benefit from the services provided by a working machine but dislike contributing to fix it, with potentially heterogeneous preferences for both activities. A designer seeking to maximize the group's long-run welfare crafts a mechanism, which regulates access to the machine's services and solicits contributions from group members.

We model the relationship between contributions and the uptime of the machine in a flexible, reduced-form way admitting multiple interpretations. To demonstrate our results and facilitate comparisons with existing work, we describe two interpretations in detail: a fluid setting featuring a large population of long-lived agents participating simultaneously (Section 6), and a Poisson setting featuring a sequence of short-lived agents participating asynchronously (Section 7).

Our setting can be viewed as a variant of classic public goods provision problems, with the twist that contributions are nonmonetary. In particular, they can be made at a limited rate and only when the machine is broken, and they cannot be stored for later use. Since the designer commits to directives based on information (the state of the machine) that is not available to agents, our model additionally has an information-design interpretation. Formally, it can be viewed as an information design problem

2

in which the designer's information and payoff function are endogenously determined by the mechanism. (We elaborate on these relationships in Section 2.3.)

Under a first-best mechanism (Section 3), the designer asks all agents with sufficiently low contribution costs to contribute whenever the machine is broken. Unfortunately, when agents have heterogeneous preferences, this mechanism generally induces some agents with low usage benefits to leave the group. Participation constraints must therefore be incorporated into the mechanism design problem.

If the designer can observe agents' preferences (Section 4), the optimal mechanism ensures participation of all agents by customizing their contribution requirements. In general, optimal contributions drop with an agent's contribution cost and rise with their usage benefit. When agent preferences satisfy a total ordering condition, optimal contributions exhibit a tripartite structure: High types contribute fully and low types contribute nothing, while intermediate types make partial contributions. Notably, participation constraints bind only for intermediate types, and so an agent's payoff is non-monotone in his type.

If the designer cannot observe agents' preferences (Section 5), the mechanism must restrict access to the machine in order to truthfully elicit preferences. We show that the optimal screening mechanism offers at most two distinct participation levels. In one, the agent enjoys full access to to the machine. In the other, the agent enjoys only limited access while incurring a reduced responsibility for contributions. Under this mechanism, some agents opt out of the group.

These results suggest new lessons for the design of recommendation engines. Purely voluntary feedback leaves an unresolved collective action problem, and the average user's experience could be improved by collecting additional mandatory feedback from some users. However, this policy may lead some users to stop participating. Our results indicate that platforms may be able to more efficiently elicit feedback by mandating a small amount from users as a baseline, while simultaneously offering a premium membership tier with better-tailored recommendations in return for additional feedback.

1.1 Related literature

Our collective upkeep model is most closely related to work studying mechanisms for provision of excludable public goods, notably Hellwig (2003) and Norman (2004). Our

3

model differs from those in two key respects. First, group members may differ in their cost of contribution. This heterogeneity creates a utilitarian motive for concentrating contributions on low-cost group members, an objective absent from existing work. Second, group members face bounds on their ability to contribute. This constraint drives the optimality of mechanisms with multiple access and contribution levels. By contrast, a posted price would be optimal in a model with atomistic agents and no contribution limits, as in the large-population limit of Norman (2004).

Recent work by Battaglini and Palfrey (2024b) similarly features heterogeneous contribution costs and bounds on contribution levels. Unlike our model, their agents cannot be excluded from the public good, so must be incentivized to contribute through the possibility that their contribution is pivotal. In the large-population limit, optimal public goods provision therefore converges to zero unless the per-capita cost of the public good vanishes. By contrast, in our setting the additional instrument of excludability allows non-vanishing contributions to be extracted from a large population. An additional distinction is that in our model, the contribution bound is endogenously determined and tightens as more of the public good is provisioned. This feature has no analog in their model.

Our setting also bears similarities to the literature studying dynamic contribution models (Admati and Perry 1991; Battaglini and Palfrey 2024a; Compte and Jehiel 2003; Deb, Oery, and Williams 2024; Fershtman and Nitzan 1991; Marx and Matthews 2000; Matthews 2013). In these models, players make incremental contributions toward the provision of a public good. Unlike in our setting, they cannot commit to contribution plans before observing the state of the public good, creating incentives for dynamic freeriding. We instead focus on how dynamics constrain players' opportunities to contribute.

Our paper additionally joins a recent literature studying mechanism design under redistributionary motives. Existing work has focused on the allocation of private goods: See, for instance, Akbarpour, Dworczak, and Kominers (2023), Condorelli (2013), and Pai and Strack (2024) for one-sided problems; Dworczak, Kominers, and Akbarpour (2021) for a two-sided problem; and Kang (2023) for a problem involving a mechanism run alongside an unregulated market. In these settings, agents have differing, privately observed social values of money, endowing a utilitarian designer with a preference for redistribution. In our setting, agents' heterogeneous contribution costs induce an analogous motive in a public goods provision context.

4

As applied to crowd-sourced recommendation engines, our paper contributes to recent work in economics (Che and Hörner 2018; Kremer, Mansour, and Perry 2014) and computer science (Frazier et al. 2014; Immorlica et al. 2019; Mansour, Slivkins, and Syrgkanis 2015) studying their incentive compatibility. In these studies, recommendation engines face constraints on their ability to foster experimentation when agents can observe the time that new content has been available. In particular, potential early adopters expect the underlying algorithm to have little information and distrust the engine's recommendations. By contrast, the designer in our setting can obscure the time since new content has arrived. She is instead constrained in her ability to preferentially select agents for experimentation on the basis of their preferences.

Methodologically, our analysis shares important features with recent studies of long-run or steady-state mechanism design, especially in queuing (Che and Tercieux 2021; Margaria 2023) and matching (Baccara, Lee, and Yariv 2020). In these settings, agents view themselves as participating in a stationary system with an endogenous state whose distribution is the same for all agents. This assumption contrasts with the prevalent classical framework in which the initial state is exogenous and agents' incentives change systematically with calendar time. The stationary perspective plays a central role in our dynamic model interpretations. (See Sections 6 and 7 for details.)

2 Model

2.1 Setting

A group of agents enjoy the services provided by an excludable public good, which for concreteness we refer to as a *machine*, in continuous time over a doubly infinite horizon. At each moment the machine is either *working* or *broken*. The machine provides services only when it is working, and a working machine breaks periodically. A broken machine must be fixed through contributions from group members. Each agent can contribute only when the machine is broken and faces a limit on his rate of contribution, which without loss we normalize to 1.

Agents value the machine according to the frequency with which they use its services and contribute to fixing it. Agents have heterogeneous preferences over these two outcome metrics, summarized by *types* θ drawn from a probability measure λ

over a Borel space Θ. If an agent of type θ uses the machine a fraction $r \in [0, 1]$ of the time and contributes a fraction $p \in [0, 1]$ of the time, he enjoys total utility

$$r \cdot u(\theta) - p \cdot c(\theta), \tag{1}$$

where $u(\theta) > 0$ is the agent's usage benefit and $c(\theta) > 0$ is his contribution cost. We assume that $u, c \in \mathscr{L}^1(\lambda)$.

2.2 Design problem

A designer oversees the machine's upkeep. To do so, she collects contributions from agents and regulates their access to the machine's services. Concretely, she chooses an *allocation* (R, P) consisting of long-run *usage levels* $R : \Theta \to [0, 1]$ and *contribution levels* $P : \Theta \to [0, 1]$ for all agents. Her goal is to maximize utilitarian social welfare

$$W(R, P) = \int \Big(R(\theta) \cdot u(\theta) - P(\theta) \cdot c(\theta) \Big) \, \lambda(\mathrm{d}\theta). \tag{2}$$

The set of feasible allocations is limited by *physical constraints* related to the process by which the machine breaks and is fixed. We implement these constraints in a reduced-form way consistent with multiple interpretations. Because these interpretations do not add any further formal content to our analysis, we defer describing them until Sections 6 and 7.

Let $Q \in [0, 1]$ denote the long-run fraction of time the machine is working, which we refer to as the machine's *uptime*. Then the physical constraints are given by

$$\rho \cdot Q = \int P \, \mathrm{d}\lambda \tag{B}$$

$$R(\theta) \leq Q, \quad P(\theta) \leq 1 - Q \tag{S-θ}$$

The *balance condition* (B) says that the long-run rate at which the machine breaks (on the left-hand side) balances the long-run rate at which agents contribute to fixing it (on the right-hand side). The parameter $\rho > 0$ can be interpreted as the machine's breakage rate. The *simplex bounds* (S-θ) say that each agent's usage and contributions are bounded by the respective frequencies with which the machine is working and broken.

The set of feasible allocations may additionally be limited by *economic constraints*

6

related to agents' freedom of action. These constraints encode familiar interim participation and incentive compatibility requirements. Formally, the participation constraints are

$$R(\theta) \cdot u(\theta) - P(\theta) \cdot c(\theta) \geq 0. \tag{P-θ}$$

The left-hand side captures an agent's payoff from participating in the mechanism, while the lower bound of 0 is an agent's outside option obtained by never using or contributing to fix the machine. Meanwhile, the incentive-compatibility (IC) constraints are

$$R(\theta) \cdot u(\theta) - P(\theta) \cdot c(\theta) \geq R(\theta') \cdot u(\theta) - P(\theta') \cdot c(\theta). \tag{IC-$\theta\theta'$}$$

That is, each agent can obtain any other agent's usage and contribution levels by misreporting his type. Under incentive compatibility, no agent should prefer to do so.

Given a feasible allocation (R, P), the machine's uptime Q is pinned down by the balance condition (B). Nonetheless, as we will demonstrate later, it is technically convenient to include Q as an additional design variable and enforce (B) by imposing it as a constraint on the designer's welfare maximization problem. We therefore study the designer's problem of choosing a *mechanism* (R, P, Q) to maximize utilitarian welfare (2) subject to the physical constraints (B) and (S-θ), as well as potentially the economic constraints (P-θ) and (IC-$\theta\theta'$).

2.3   Relationship to existing models

Our model is closely related to static models of public goods provision and information design. We now describe the precise relationship between these settings. Both the similarities and differences are instructive for understanding our results and contribution to the literature, in particular the role played by our physical constraints.

2.3.1   Public goods provision

In a static public goods provision problem, a designer must raise contributions simultaneously from a group of heterogeneous agents in order to produce a collective good. As a benchmark, consider the following static model, which is a closest approach to our collective upkeep model. The group is a continuum of mass 1 with preferences described by a measure λ over the type space Θ. Let P be the vector of contributions;

R be the vector of allocations of the public good; and Q be the quantity of the public good, which is determined by total contributions through the linear production function (B). Agent payoffs are linear in allocations and contributions, as in (1). The public good is excludable, so that any R such that $0 \leq R(\theta) \leq Q$ is feasible.

This static benchmark is very close to the setting studied by Hellwig (2003) and Norman (2003). The main distinction between their setting and our benchmark is that here agents are atomistic (i.e., infinitesimal). As a result, the designer faces no aggregate uncertainty over group preferences. Otherwise, the benchmark is a special case of their setting, with linear production and consumption utility.

Payoffs for all players are the same in the static benchmark and our model of collective upkeep. Additionally, the constraints on feasible allocations in the former model are a subset of the ones in the latter. In particular, the economic constraints on participation and incentive compatibility are identical in the two settings. Furthermore, the balance condition and the simplex bound $R(\theta) \leq Q$ are required for feasibility in both settings. The key difference between the settings is that the collective upkeep model requires the additional simplex bound $P(\theta) \leq 1 - Q$.

This extra constraint not only limits each agent's ability to contribute, but it imposes an *endogenous* bound governed by overall provision of the public good. In a static setting, exogenous contribution bounds could be justified through budget constraints or the contribution of a scarce non-monetary resource like labor. (For an example of a public goods problem in which exogenous contribution bounds are imposed, see Battaglini and Palfrey (2024b).) However, in static settings there is no natural microfoundation for contribution bounds which depend on the contributions of other agents. In our collective upkeep setting, by contrast, such bounds arise naturally.

2.3.2 Information design

In a static information design problem, a sender attempts to persuade a receiver to take their preferred action by disclosing information about an underlying state. As a benchmark, consider the following problem, which is a closest approach to our collective upkeep model. The receiver who can take one of two actions, "Interact" (I) or "Avoid" (A). He receives payoff 0 from choosing $a = A$ and a state- and type-dependent payoff from choosing $a = I$, where the state space is $\Omega = \{W, B\}$ and he receives payoff $u(\theta)$ if $\omega = W$ and $-c(\theta)$ if $\omega = B$. The sender's payoffs agree with

the receiver's except when $(\omega, a) = (B, I)$, in which case the sender receives payoff $y - c(\theta)$ for some $y > 0$. The state is W with prior probability Q.

With observable agent types, this problem is similar to the binary-action information design problems studied in Kamenica and Gentzkow (2011). With unobserved types, it is akin to models studied in Kolotilin et al. (2017) and Guo and Shmaya (2019). Standard arguments from these settings imply that the binding incentive constraints involve obedience to a recommendation of $a = I$ and (if types are unobserved) truthful reporting. As a result, the incentive constraints can be reduced to the economic constraints appearing in the collective action model. In particular, the participation constraint ensures obedience to recommendations of $a = I$, while the IC constraint ensures truthful type reporting.

The key difference between the two settings is that in the collective upkeep problem, the variables Q and y are endogenously determined by the balance condition. In particular, the rate at which the receiver chooses $a = I$ when $\omega = B$ pins down Q via the balance condition, and this linkage in turn determines the social value y to the planner of interacting with a broken machine. This feedback loop has no analog in static information design settings, where payoffs and state realizations are exogenous.

3   First-best mechanism

We now solve for the designer's optimal mechanism satisfying the physical constraints, without imposing any economic constraints. Formally, the designer chooses (R, P, Q) to maximize aggregate welfare (2) subject to the balance condition (B) and simplex bounds (S-θ) for all $\theta \in \Theta$.

3.1   Characterizing the optimum

An immediate observation is that welfare is increasing in agents' usage of the machine, while usage levels do not appear in the constraints. As a result, at the optimum all agents should enjoy full access to the machine. Formally, $R(\theta) = Q$ for all types at the optimum, saturating the corresponding simplex bounds. Going forward, we enforce this choice of R and drop it from the optimization problem.

We next establish that the first-best contribution levels exhibit a threshold structure: Low-cost members, whose contribution cost falls below a threshold y^{fb}, con-

tribute the maximum amount, while members with higher costs do not contribute at all. Let $\bar{u} = \int u \, d\lambda$ be the group's aggregate usage benefit from a working machine.

Proposition 1. *Let y^{fb} be the unique y solving*

$$\bar{u} - \rho \cdot y = \int (y - c)^+ \, d\lambda, \tag{3}$$

and let W^{fb} be the common value in (3) evaluated at y^{FB}. Then the mechanism $(P^{\mathrm{fb}}, Q^{\mathrm{fb}})$ defined by

$$\rho \cdot \frac{Q^{\mathrm{fb}}}{1 - Q^{\mathrm{fb}}} = Pr_\lambda(c < y^{\mathrm{fb}}), \quad and \tag{4}$$

$$P^{\mathrm{fb}}(\theta) = (1 - Q^{\mathrm{fb}}) \mathbf{1}_{[c < y^{\mathrm{fb}}]}(\theta)$$

solves the first-best problem and yields welfare W^{fb}.

(All proofs omitted from the body of the paper are collected in Section 8.)

To prove the optimality of a threshold contribution structure, we consider the Lagrangian relaxation of the problem with respect to the balance condition. Assigning a Lagrange multiplier y to that constraint yields the Lagrangian

$$\mathcal{L}(P, Q; y) = Q \cdot (\bar{u} - \rho \cdot y) + \int P(\theta) \cdot (y - c(\theta)) \, \lambda(d\theta) \tag{5}$$

Holding (Q, y) fixed, the Lagrangian factorizes over types, allowing P to be optimized pointwise. Note that this factorization relies on the inclusion of Q as an auxiliary optimization variable, motivating its introduction.

The optimal multiplier y^{fb} measures the social value of fixing the machine. Indeed, y^{fb} is the amount of money the designer would be willing to pay to fix the machine immediately when it breaks down. While this idea can be made formal using dynamic programming principles, such techniques are not easily generalized to incorporate economic constraints. By contrast, the Lagrangian approach taken in our proof is readily adapted to these constraints, as we will demonstrate shortly.

3.2 Comparative statics

We now examine how the optimal mechanism is shaped by the breakage rate ρ. Increasing ρ reduces the machine's average lifespan and decreases the social value of fixing it. Then since y^{fb} tracks this social value, it follows that:

Corollary. *The first-best contribution threshold y^{fb} is decreasing in the breakage rate ρ.*

Proof. The left-hand side of (3) is decreasing in y while the right-hand side is increasing. Since an increase in ρ induces a downward shift in the former, y^{fb} must decrease in ρ. □

A lower contribution threshold means that the machine is fixed more slowly when it breaks. As a result, machine uptime must decline with ρ:

Corollary. *Machine uptime Q^{fb} under the first-best mechanism is nonincreasing in the breakage rate ρ and decreasing whenever $Q^{\text{fb}} > 0$.*

Proof. (4) implies that Q^{fb} is nonincreasing in ρ and decreasing whenever it is positive, while it is and nondecreasing in y^{fb}. Since the previous corollary established that y^{fb} is decreasing in ρ, it follows that Q^{fb} is as well. □

In spite of these two results, the long-run aggregate contribution level $\int P \, d\lambda$ can increase in ρ. Recall that the bound on each agent's contribution level is endogenously determined by the machine's downtime $1 - Q^{\text{fb}}$. Since downtime increases in ρ, all agents asked to contribute do so more often. This force can overwhelm the countervailing reduction in y^{fb}. For instance, if $\Theta = \{\theta_0\}$ is a singleton, then (3) implies that $y^{\text{fb}} = (u(\theta_0) + c(\theta_0))/(1 + \rho) > c(\theta_0)$ for sufficiently small ρ. In that regime, the aggregate contribution level is $1 - Q^{\text{fb}}$, which is increasing in ρ in light of the previous corollary.

4   Mechanisms with participation constraints

When agents are homogeneous, the first-best mechanism of a utilitarian designer delivers all agents a non-negative payoff and therefore satisfies participation. However, when agents have heterogeneous preferences, the first-best mechanism generally violates the participation constraints of some agents. We now solve for an optimal mechanism under the additional economic constraint that all agents must prefer the mechanism to their outside option. Formally, we impose the participation constraints (P-θ) for all $\theta \in \Theta$. For the time being, we continue to assume that the designer observes agents' types.

4.1 Characterizing the optimum

As in the first-best solution, there is no reason to restrict any agent's access to a working machine. Indeed, such restrictions would lower social welfare while simultaneously tightening participation constraints. Hence $R(\theta) = Q$ for all types at the optimum. Going forward, we enforce this choice of R and drop it from the optimization problem.

In this reduced problem, the participation constraint may be rearranged to obtain an upper bound on each agent's contribution level:

$$P(\theta) \leq Q \cdot \nu(\theta),$$

where

$$\nu(\theta) = u(\theta)/c(\theta)$$

captures an agent's rate of substitution between usage and contributions. We will refer to this quantity as the *valuation* of an agent of type θ. This representation allows the two constraints (S-θ) and (P-θ) to be combined into the single constraint

$$P(\theta) \leq \min(1 - Q, Q \cdot \nu(\theta)) \tag{SP-θ}$$

Since participation constraints merely tighten the upper bound on feasible contribution levels, a threshold contribution structure is optimal for the same reasons as in the first-best problem. The new feature of the problem with participation constraints is that agents who are asked to contribute may do so at different levels. Since the constraints (SP-θ) are type-dependent, types for whom $\nu(\theta) < (1 - Q)/Q$ contribute a type-dependent amount less than their feasible maximum. These are precisely the agent types for whom participation constraints bind.

Let

$$\ell(Q; y) = Q \cdot (\bar{u} - \rho \cdot y) + \int \min(1 - Q, Q \cdot \nu(\theta))(y - c(\theta))^{+} \, \lambda(\mathrm{d}\theta) \tag{6}$$

be the designer's reduced Lagrangian under the optimal threshold contribution rule, where recall that y is the Lagrange multiplier on the balance condition. The following theorem establishes that an optimal mechanism under participation constraints can be obtained by solving the minimax problem associated with ℓ.

Theorem 1. *Let W^* be the value and (Q^*, y^*) be the smallest[1] saddle point of the concave-convex minimax problem*

$$\min_{y \geq 0} \max_{Q \in [0,1]} \ell(Q; y) = \max_{Q \in [0,1]} \min_{y \geq 0} \ell(Q; y),$$

with y^* defined as $+\infty$ if the minimum is not attained. Then the mechanism (P^*, Q^*) with

$$P^*(\theta) = \min(1 - Q^*, Q^* \cdot \nu(\theta)) \mathbf{1}_{[c < y^{\text{fb}}]}(\theta)$$

solves the designer's problem under participation constraints and delivers welfare W^*. Moreover, $y^* \geq y^{\text{fb}}$, and this inequality is strict if $W^* < W^{\text{fb}}$.

The cost threshold y^* is the optimized value of the Lagrange multiplier on the balance condition. As a result, it has a natural interpretation as the (marginal) shadow value of fixing the machine. When participation constraints bind, this value is larger than under the first-best mechanism. Indeed, fixing the machine not only generates usage benefits for other agents (as in the unconstrained problem) but additionally loosens their participation constraints by raising the machine's uptime. As a result, the designer can concentrate contributions on a narrower set of low-cost agents.

Surprisingly, this latter force can be powerful enough to induce more frequent repairs under participation constraints than without them. (Of course, this higher uptime cannot improve social welfare over the first-best outcome.) The following example illustrates this possibility.

Example. Consider a three-type environment $\Theta = \{L, M, H\}$ with

	L	M	H
u	3	4	10
c	3	2	1.25
λ	1/3	1/3	1/3

The machine breaks at rate $\rho = 5.5$. In this setting, $y^{FB} = 2.7$ and $Q^{FB} = 4/15$, while $y^* = 45/14$ and $Q^* = 2/7 > Q^{FB}$.

---

[1] The selection of the smallest saddle point is not needed for optimality, but we adopt it as a convention for comparability with the mechanism characterized in Proposition 1.

4.2 Ordered types

In general, the set of participation-constrained agents are those with low contribution costs and valuations. When costs and benefits comove, this set takes a particularly simple form.

Definition 1. Agent types are *ordered* if $\Theta = \mathbf{R}_-$ and ν is increasing in θ while $c(\theta) = -\theta$.

The substantive content of ordered types is that the type structure is one-dimensional and valuations are increasing while costs are decreasing along this dimension. The identification of θ with the agent's cost is simply a convenient normalization. It is without loss, since types can always be relabeled to achieve this identity.

Recall that the first-best mechanism requires agents with sufficiently low costs to contribute. For ordered types c is decreasing, and therefore the first-best mechanism can be characterized by a type threshold $\theta^{\text{fb}} = -y^{\text{fb}}$ above which agents are asked to contribute. Meanwhile, because ν is increasing under ordered types, participation constraints are active on an interval of intermediate types. Indeed, the designer's incentives are aligned with both high-type agents (who don't mind contributing) and low-type ones (for whom contributing is too costly to be socially valuable). A conflict arises only for intermediate-type agents, who generate a large social value from contributing but do not sufficiently enjoy using the machine.

The following proposition formalizes this reasoning. It additionally establishes that the interval on which participation constraints bind includes the first-best contribution threshold.

Proposition 2. *Suppose that types are ordered and first-best welfare is not achievable under participation constraints. Then there exists an optimal mechanism in which participation constraints are active on an interval of types containing θ^{fb}.*

4.3 Comparative statics

Unlike the first-best solution, the optimal threshold y^* under participation constraints does not exhibit unambiguous comparative statics in the breakage rate ρ. There are two contrasting forces here. On the one hand, as in the first-best problem, frequent breakdowns reduce the usage benefits generated by fixing the machine. On the other hand, participation constraints tighten as breakdowns become frequent, and fixing the

machine more often loosens these constraints. Fixing the machine therefore allows the designer to concentrate contributions on lower-cost agents, yielding an aggregate cost savings.

The tension between these forces is illustrated by the limiting behavior of y^* as ρ grows large. Depending on the thickness of the right tail of the valuation distribution, contributions may become narrowly concentrated on low-cost types (i.e., the first force dominates) or else dispersed throughout the population of agents (the second force dominates).

Proposition 3. *Suppose that $\mathbb{E}_\lambda[\nu] = \infty$ and the distribution of c under λ has full support on $\mathbb{R}_+$.*

- *If $\lim_{v \to \infty} v \Pr_\lambda(\nu > v) = 0$, then $\lim_{\rho \to \infty} y^* = \infty$.*

- *If $\lim_{v \to \infty} v \Pr_\lambda(\nu > v) = \infty$ and ν and c are negatively associated under λ, then $\lim_{\rho \to \infty} y^* = 0$.*

The condition that ν have infinite expectation ensures that the optimal mechanism under participation constraints generates positive welfare even for large ρ. It is not a particularly restrictive assumption given that ν is a ratio of usage benefits to costs. It is satisfied, for instance, if $\liminf_{\varepsilon \to 0} \Pr_\lambda(c < \varepsilon)/\varepsilon > 0$ and usage benefits are bounded away from zero. The assumption is also compatible with both tail conditions appearing in the proposition. For instance, if $\Pr_\lambda(\nu > v) \sim 1/(v \cdot (\log v)^\alpha)$ for large v for some $\alpha \geq 1$, then $\mathbb{E}_\lambda[\nu] = \infty$ while $\lim_{v \to \infty} v \Pr_\lambda(\nu > v) = 0$. On the other hand, if $\Pr_\lambda(\nu > v) \sim 1/v^\alpha$ for large v for some $\alpha \in (0, 1)$, then $\mathbb{E}_\lambda[\nu] = \infty$ while $\lim_{v \to \infty} v \Pr_\lambda(\nu > v) = \infty$.

When types are ordered, the tail conditions on ν can be recast in terms of the density of costs at zero. This is because under ordered types, valuations are high when θ is large, which is also when costs are low. As a result, a thick right tail is equivalent to an unbounded cost density near zero, while conversely a thin right tail is equivalent to a vanishing cost density.

Proposition 4. *Under the assumptions of Proposition 3, suppose additionally that types are ordered and $\limsup_{\theta \to 0} u(\theta) < \infty$ while $\liminf_{\theta \to 0} u(\theta) > 0$.*

- *If $\lim_{\varepsilon \to 0} Pr_\lambda(c < \varepsilon)/\varepsilon = 0$, then $\lim_{\rho \to \infty} y^* = \infty$.*

- *If $\lim_{\varepsilon \to 0} \Pr_\lambda(c < \varepsilon)/\varepsilon = \infty$, then $\lim_{\rho \to \infty} y^* = 0$.*

The machine's uptime is a function both of the breakage rate and of aggregate contribution levels. As discussed in our analysis of the first-best mechanism, aggregate contributions need not shrink with ρ. Nonetheless, it turns out that contribution levels never grow fast enough to outweigh the higher breakage rate, and the machine's uptime decreases in ρ.

Proposition 5. *Q^* is nonincreasing in ρ.*

While the assertion is intuitive, we are not aware of any standard results regarding comparative statics of saddle points which directly imply it. The difficulty arises from the fact that the function ℓ of which (Q^*, y^*) is a saddle point is neither sub- nor supermodular in (Q, y, ρ). To prove the result, we identify an appropriate modification of ℓ which is submodular and apply standard results to that function.

5 Incentive-compatible mechanisms

We now solve for an optimal mechanism when the designer must induce agents not just to participate, but also to truthfully report their preferences. Formally, we impose the participation constraints (P-θ) as well as the IC constraints (IC-$\theta\theta'$) for all $\theta, \theta' \in \Theta$. We refer to a mechanism which satisfies both sets of constraints as a *screening* mechanism.

As we have seen, when agents' types are observable, there is no benefit to restricting any agent's access to the machine. However, when agents must self-report their contribution costs, access restrictions become crucial for screening. The usage levels R must therefore be designed in tandem with the contribution levels P.

The following result characterizes the main qualitative features of an optimal screening mechanism.

Theorem 2. *There exists an optimal screening mechanism (R^*, P^*, Q^*) which satisfies*

$$(R^*(\theta), P^*(\theta)) = \begin{cases} (Q^*, 1 - Q^*), & \text{if } \bar{\nu} \leq \nu(\theta) \\ (\hat{R}, \hat{P}), & \text{if } \underline{\nu} \leq \nu(\theta) < \bar{\nu} \\ (0, 0), & \text{if } \nu(\theta) < \underline{\nu} \end{cases}$$

for some $0 < \underline{\nu} \leq \bar{\nu} \leq \infty$ *and* $0 < \hat{R} \leq Q^*$ *and* $0 < \hat{P} \leq 1 - Q^*$. *Moreover, if* $\bar{\nu} = \infty$ *then* $\hat{R} = Q^*$.

In words, the designer screens agents by offering up to two distinct membership tiers: a full-access tier with high contribution requirements, and potentially also a partial-access tier with reduced contributions. Whenever multiple membership tiers are optimal, the high tier demands the maximum contribution from all members in that tier. Finally, all agents have the ability to opt out of the group and receive the $(0,0)$ allocation. Unlike the case of observable preferences, in general some agents do opt out of the group under an optimal screening mechanism.

The proof of this result relies on a reduction of the mechanism design problem to a Myersonian monopoly-sale problem. In this reduction, buyers have quasilinear utility with valuations $\nu(\theta)$ for the good. The quantity $R(\theta)$ becomes the probability with which a buyer of type θ receives the good while $P(\theta)$ is the price he pays for it. The seller's objective is a weighted average of revenue, reflecting a Lagrange multiplier on the balance condition, and the utility of each buyer type. Unlike a classic monopoly-sale problem, the simplex bound $P(\theta) \leq 1 - Q$ imposes an upper bound on the price that each buyer can pay for the good. As a result, a posted-price mechanism may not be optimal and fractional allocations can be beneficial for screening buyers. (We provide an analysis of the monopoly-sale problem under a payment bound in Appendix A.)

We show that an optimal allocation rule corresponds to an extreme point in the set distribution functions which satisfy a moment inequality equivalent to the budget constraint. In general, extreme points may involve one additional point of support for each moment inequality, implying our result. The same argument drives the findings of Le Treust and Tomala (2019) and Doval and Skreta (2024), who show that in an information design context, each additional constraint adds up to one additional point of support in the set of posteriors induced by an optimal mechanism.

Theorem 2 ensures that no more than two membership tiers are required at the optimum. The following example exhibits a setting in which exactly two tiers are required.

Example. Suppose that $\Theta = \{L, M, H\}$, with equal probability of each type. All types have contribution cost $c(\theta) = 1$, while usage utilities are $u(H) = 5, u(M) = 1, u(L) = 0.1$. The breakage rate is $\rho = 1$.

Under the first-best mechanism, $Q^{\text{fb}} = 0.5$, and all types fully contribute, so that $(R^{\text{fb}}(\theta), P^{\text{fb}}(\theta)) = 0.5$ for all types.

Under the optimal screening mechanism, $Q^* = 0.3$. Type H receives full access in return for a full contribution, type M receives partial access in return for a reduced contribution, and type L leaves the group. The optimal allocations are $R^*(H) = 0.3, R^*(M) = 0.2, R^*(L) = 0$ and $P^*(H) = 0.7, P^*(M) = 0.2, P^*(L) = 0$.

6 Model interpretation: Fluid interaction with long-lived agents

In this section, we show how our reduced-form model can be interpreted as a club comprised of a large population of long-lived agents, who use and contribute to the machine in an incremental manner. We describe this interpretation and derive the reduced-form physical constraints from it in Section 6.1. We then relate it to the design of crowdsourced recommendation platforms in Section 6.2.

6.1 Microfounded model

The group is comprised of a mass of atomistic (i.e., infinitesimal), long-lived agents. Agents live for a random period of time drawn from a (type-dependent) distribution $F(\cdot \mid \theta)$ with expectation $1/\kappa$ for each type. (For simplicity, we assume that κ is the same for all types.) Deaths are balanced by an inflow of new agents at rate κ. We assume that the population of agents is in steady state, so that the empirical distribution of agent ages G is time-invariant with density $g(z \mid \theta) = 1 - F(z \mid \theta)$. Agent types are drawn from the distribution λ. Because all types have the same expected lifespan, λ is also the empirical distribution of agent types in steady state.

At each moment when the machine is working, each agent can use its services, enjoying a flow benefit $u(\theta)$ from (full) access. At each moment when the machine is broken, each agent can contribute to fixing it, incurring a flow cost $c(\theta)$ from (fully) contributing. Agents evaluate the payoff of participating in the mechanism according to the expected cumulative usage benefits minus contribution costs over their lifespan.

When the machine breaks, a random *contribution quantum* (i.e., a cumulative contribution aggregated across agents and time) is required to fix it. The size of this quantum is drawn from a fixed distribution with expectation normalized to 1,

independent of the machine's history. When the machine is fixed, it breaks after a random *lifespan* drawn from a fixed distribution with expectation $1/\rho$, independent of the machine's history.

A mechanism assigns personalized (type- and age-dependent) usage and contribution levels, varying based on the current machine state as well as potentially the history of the system. We restrict attention to mechanisms that induce a stationary environment for agents born at different calendar times. Formally, a mechanism must be compatible with a stationary *system process* $X_t = (\omega_t, r_t, p_t)$, where:

- $\omega_t \in \{W, B\}$ is the state of the machine at time t

- $r_t : \Theta \times \mathbf{R}_+ \to [0, 1]$ records the time-t usage level $r_t(\theta, z)$ of each agent of type θ and age z

- $p_t : \Theta \times \mathbf{R}_+ \to [0, 1]$ records the time-t contribution level $p_t(\theta, z)$ of each agent

Note that the information in X is sufficient to derive the entire history of the system, including the lifespans of the machine (using the transition times of ω from W to B) and the contribution quanta (using the transition times of ω from B to W plus the contribution levels r).

Given any stationary system process, let $\bar{r}_t(\theta) = \int r_t(\theta, z)g(z \mid \theta) \, \mathrm{d}\theta$ and $\bar{p}_t(\theta) = \int p_t(\theta, z)g(z \mid \theta) \, \mathrm{d}\theta$ be the group-average usage and contribution levels for each type. The associated *reduced form* (R, P, Q) is well-defined, where $R(\theta) = \mathbb{E}\bar{r}_t(\theta), P(\theta) = \mathbb{E}\bar{p}_t(\theta), Q = \Pr(\omega_t = W)$ and each of these expressions is independent of t by stationarity.

The normalized expected payoff for an agent of type θ born at time t can be written in terms of this reduced form as

$$\kappa \cdot \mathbb{E} \int_0^\infty \mathrm{d}F(\tau \mid \theta) \int_0^\tau \Big(r_{t+z}(\theta, z) \cdot u(\theta) - p_{t+z}(\theta, z) \cdot c(\theta) \Big) \, \mathrm{d}z = R(\theta) \cdot u(\theta) - P(\theta) \cdot c(\theta),$$

as in (1), where the inner integral aggregates the agent's flow payoffs over his lifespan and the outer one averages over possible lifespans. It follows that utilitarian welfare, defined as the average welfare of a randomly-selected agent, equals (2). (By the ergodic theorem, this objective is equivalent to an alternative specification of welfare as the long-run average flow payoffs of group members.)

Markovian mechanisms condition current usage and contribution levels only on the current state of the machine. Formally, a Markovian mechanism is characterized

by a Borel function $\sigma : \Theta \times \{W, B\} \to [0, 1]$, with the interpretation that an agent of type θ is assigned usage level $\sigma(\theta, W)$ when the machine is working and contribution level $\sigma(\theta, B)$ when the machine is broken. We do not formally describe more general mechanisms, as we will show that any system process satisfying certain necessary conditions is payoff-equivalent to one induced by a Markovian mechanism.

Definition 2. A stationary system process is *admissible* if:

1. Machine lifespans are i.i.d. with expectation $1/\rho$.

2. Contribution quanta are i.i.d. with expectation 1.

3. $r_t(\theta) \leq \mathbf{1}_{[\omega_t = W]}$ and $p_t(\theta) \leq \mathbf{1}_{[\omega_t = B]}$ for all t.

Every stationary system process compatible with some (not necessarily Markovian) mechanism is admissible. The first two requirements of admissibility ensure that the process respects the independence of successive lifespans and cumulative contributions. The last requirement ensures that the machine is used only when it is working and contributions are collected only when the machine is broken.

The following result establishes that every reduced form of an admissible system process satisfies the physical constraints. Furthermore, all triples (R, P, Q) satisfying the physical constraints can be achieved via a Markovian mechanism. In particular, the physical constraints are necessary and sufficient for feasibility of an allocation.

Proposition 6. *1. Fix an admissible system process. Then its reduced form satisfies the physical constraints.*

2. *Suppose that the triple (R, P, Q) satisfies the physical constraints. Then there exists a Markovian mechanism σ and a stationary system process X such that σ is compatible with X and (R, P, Q) is the reduced form of X.*

6.2 Application: Crowdsourced recommendation platforms

The fluid interpretation of our model can be applied as a stylized model of digital content platforms such as Netflix and Tiktok, which rely on user feedback to make content recommendations.

In this application, the agents are users of the platform, while the machine is a recommendation algorithm which makes personalized content suggestions to users.

The machine is working when it is well-informed about the attributes of its current content library, for instance each Netflix movie's overall quality as well as its fit for particular user demographics. The machine breaks when the library of available content turns over, for instance because Netflix loses the rights to previously-available shows and movies and licenses a new batch.

The machine is fixed by showing new content to users and eliciting feedback on their experiences. Feedback is costly for users to give for two reasons. First, users do not enjoy untested content on average, for instance because match value is an important component of the user experience. Second, users find it costly to explicitly report their experiences through ratings or more detailed surveys.

This application can be formally mapped onto our abstract model in the following way. Uncertainty about the available content on a recommendation platform is summarized by a state $\omega \in \Omega$ pertaining to a content library X. Each user receives a flow payoff $v(\theta, x, \omega)$ from consuming content $x \in X$ from the library when the library's state is ω. When ω is known, the algorithm recommends $x^*(\theta, \omega) \in \arg\max_{x \in X} v(\theta, x, \omega)$, generating a usage benefit $u(\theta) = v(\theta, x^*(\theta, \omega), \omega)$ for each user. Since consumers may consume content repeatedly over time, a choice x from the content library should be interpreted as a bundle of individual pieces of content.

When the machine breaks, ω is drawn anew from some known distribution μ^ω. (For notational simplicity, the set X of content is held fixed, with the understanding that elements of X are merely labels that carry no informational content.) A mass of content of random size must be consumed by users before ω is revealed, at which point the algorithm can once again make informed recommendations.

While the algorithm is learning ω, users are served content from the library according to some (potentially type-dependent) distribution $\mu^x(\theta)$. Hence in the learning regime, users receive a usage benefit $v^0(\theta) = \mathbb{E}_{\omega, x}(v(\theta, x, \omega) \mid \theta)$. We assume that $v^0(\theta) < 0$ to capture the costs of consuming poorly-targeted content. Direct costs of providing feedback are captured by an additional flow cost $c^0 \geq 0$ of providing ratings. Each user's net cost of contributing is $c(\theta) = c^0 - v^0(\theta) > 0$.

A user's commitment to provide feedback in return for access could be implemented in several ways. The platform could bundle tested and untested content together in its recommendations, preventing a user who accepts the platform's recommendations from using the algorithm without simultaneously contributing. The

platform could additionally prompt the user to periodically provide ratings in a way which is difficult to skip.

7 Model interpretation: Sequential interaction with short-lived agents

In this section, we show how our model can be interpreted as a sequence of interactions with a population of short-lived agents, who each use and contribute to the machine exactly once. We describe this interpretation and derive the reduced-form constraints from it in Section 7.1. We then relate it to the design of incentive-aware bandit exploration models in Section 7.2.

7.1 Microfounded model

Short-lived agents arrive according to a Poisson process with rate normalized to 1. Arriving agents' types are drawn i.i.d. from the distribution λ. An arriving agent can use the machine if it is working and can contribute to fix the machine if it is broken, after which he departs forever. Using a working machine gives an agent of type θ usage benefit $u(\theta)$, while contributing to fix a broken machine incurs a cost $c(\theta)$. A contribution fixes the machine with a probability we normalize to 1. When the machine is fixed, it breaks after a random lifespan drawn from a fixed distribution with expectation $1/\rho$, independent of the machine's history.

As in our fluid interpretation (section 6), we focus on mechanisms which are compatible with a stationary system process. In a sequential arrival environment, the system process tracks agent arrivals, their types, the actions they took, and the state of the machine. Every stationary system process induces a well-defined reduced form (R, P, Q), where:

- Q is the long-run proportion of time that the machine is working. Equivalently, it is the probability that the machine is working at any given time.

- $P : \Theta \to [0, 1]$ is the long-run density of machine usage. That is, for every Borel set $B \subset \Theta$, the long-run frequency of agents who are of types θ and use the machine is $\int_B P \, d\lambda$. Equivalently, the probability that an arriving agent has a type drawn from B and uses the machine is $\int_B P \, d\lambda$.

- $R : \Theta \to [0, 1]$ is the long-run density of contributions. That is, for every Borel set $B \subset \Theta$, the long-run frequency of agents who are of types θ and contribute is $\int_B R \, \mathrm{d}\lambda$. Equivalently, the probability that an arriving agent has a type drawn from B and contributes is $\int_B P \, \mathrm{d}\lambda$.

Given these definitions, the expected payoff for an agent of type θ who is born at type t can be written in terms of this reduced form as in (1), and the long-run expected welfare of an arriving agent is as in (2).

Markovian mechanisms stochastically direct arriving agents to use or fix the machine with probabilities that depend only on their type and the current state of the machine. Formally, a Markovian mechanism is characterized by a Borel function $\sigma : \Theta \times \{W, B\} \to [0, 1]$, with the interpretation that an arriving agent of type θ uses the machine with probability $\sigma(\theta, W)$ if it is working and fixes it with probability $\sigma(\theta, B)$ if it is broken, with independent randomization across agents.

As in our analysis of the fluid interpretation, we explicitly impose only a few necessary conditions on admissible system processes. These conditions will imply the physical constraints and ensure implementability using Markovian mechanisms.

Definition 3. A system process is *admissible* if:

1. Conditional on the history of the system process before time 0, agent arrivals after time 0 follow a Poisson process with arrival rate 1 and arriving agent types are i.i.d. with distribution λ.

2. Machine lifespans are i.i.d. with expectation $1/\rho$.

3. Arriving agents use the machine only when it is working and contribute only when it is broken.

The first requirement ensures ensures that the agent arrival process is exogenous, and in particular that it is independent of the machine state. This assumption can also be viewed as a nonanticipation constraint on the designer, who does not know the future arrival process when deciding how to direct an arriving agent. The second requirement ensures that the system process encodes the physical breakage process described above. The final requirement ensures that agents can use and fix the machine only when it is working or broken, respectively.

The following result is an analog of Proposition 6 for the sequential arrival environment. It establishes that every admissible system process induces a reduced-form

satisfying the physical constraints. Further, all triples (R, P, Q) satisfying the physical constraints can be achieved via some Markovian mechanism. As a result, the physical constraints are necessary and sufficient for feasibility of an allocation.

Proposition 7. *1. Fix a stationary process that satisfies Assumption 3. The reduced form (R, P, Q) induced by this process satisfy (B) and (S-θ) for all $\theta \in \Theta$.*

2. Suppose that the triple (R, P, Q) satisfies (B) and (S-θ) for all $\theta \in \Theta$. Then there exists a Markovian mechanism that induces the reduced-form (R, P, Q).

7.2 Connection to incentivized bandit exploration

The sequential interpretation described in Section 7.1 can be used to build a bridge between our reduced-form setting and models of bandit exploration, for instance Kremer, Mansour, and Perry (2014) (hereafter KMP). To complete the bridge, we supplement the interpretation with a microfoundation of payoffs and state transitions.

The machine is a bandit arm, which each arriving agent can choose to pull or not. The payoff to pulling the arm is controlled by a state variable $\omega \in \Omega$, which renews at random times following a Poisson process with rate ρ. When renewal occurs, a new state is drawn from a probability distribution π over states, independent of the history of the system. The designer does not directly observe the state process. However, whenever an agent pulls the arm, its current state is revealed to the designer. Additionally, the designer observes when renewals occur.

An agent of type θ who pulls the arm in state $\omega \in \Omega$ obtains a type- and state-dependent payoff $v(\omega, \theta)$, where $v \in L^1(\pi \times \lambda)$. If the agent does not pull the arm, he receives a payoff of 0. Let $u, c : \Theta \to \mathbf{R}$ be given by

$$u(\theta) = \int (v(\omega, \theta))^+ \, \pi(\mathrm{d}\omega), \quad c(\theta) = - \int v(\omega, \theta) \, \pi(\mathrm{d}\omega).$$

The quantity $u(\theta)$ is the agent's ex-interim payoff from observing the arm's state and making an optimal decision to pull it or not. Meanwhile, the quantity $-c(\theta)$ is the agent's ex-interim payoff from pulling the arm when its state is unknown.

With this microfoundation, upkeep of the machine takes the form of periodic exploration of the bandit arm to learn its state. The machine is working when the arm's state is known, in which case an agent of type θ can achieve payoff $u(\theta)$ by making an optimal state-contingent decision about whether to pull the arm. The

machine breaks when the state of the arm renews. When the machine is broken, some agent must make a costly contribution by pulling the arm to learn its state.

In this interpretation, agents cannot distinguish between usage and contribution, since both involve pulling the arm. As a result, the ex-interim participation constraint of our reduced-form setting is equivalent to an ex-post obedience requirement that each agent must prefer to obey a recommendation by the designer to pull the arm. This setting can therefore be viewed as an infinite-horizon variant of KMP, with the key distinctions that the arm's state renews periodically and agents may have heterogeneous preferences.

KMP observe that when homogeneous agents do not know their order of arrival in a finite-horizon model, obedience constraints are nonbinding for a utilitarian designer. In our setting, agents arrive in steady state and calendar time is uninformative about the state of the arm and the designer's knowledge. A similar conclusion therefore obtains: The first-best outcome is achievable when agents are homogeneous. However, when agents have heterogeneous preferences, incentive constraints may bind for some agents. Our results indicate how the informativeness of recommendations to different agents should be adjusted to restore obedience.

8 Proofs

8.1 Proof of Proposition 1

We first observe that the designer's problem is equivalent to the relaxed problem in which the balance condition is enforced as the inequality constraint

$$\rho \cdot Q \leq \int P \, \mathrm{d}\lambda.$$

Indeed, any mechanism which satisfies this inequality strictly can be improved by lowering $P(\theta)$ for a positive measure of types without violating the simplex bounds or the relaxed balance condition. Hence any optimum of the relaxed problem satisfies the balance condition as an equality. For the remaining of the proof, we pass to the relaxed problem.

Let C^S be the set of all (P, Q) that satisfy the simplex bounds. For each Q, let $\bar{C}^S(Q) = \{P \; : \; (P, Q) \in C^S\}$. Consider the Lagrangian relaxation of the designer's

problem with respect to the balance condition. Let y be the dual variable. The Lagrangian is given by (5) with the domain $C^S \times \mathbf{R}_+$. The problem

$$\text{maximize}_{P \in \bar{C}^S(Q)} \mathscr{L}(Q, P; y)$$

factorizes over types and has the solution $P(\theta) = (1 - Q) \cdot \mathbf{1}_{[c(\theta) < y]}$. Let

$$\ell^{\text{fb}}(Q; y) = \max_{P \in \bar{C}^S(Q)} \mathscr{L}(Q, P; y) = Q \cdot (\bar{u} - \rho \cdot y) + (1 - Q) \int (c - y)^+ \, \mathrm{d}\lambda$$

be the reduced Lagrangian with domain $[0, 1] \times \mathbf{R}_+$. Since the domain $Q \in [0, 1]$ is compact, it follows from Sion's minimax theorem that the primal problem admits an optimal solution and strong duality holds. The dual Lagrange function is given by

$$\varphi^{\text{fb}}(y) = \max_{Q \in [0,1]} \ell^{\text{fb}}(Q; y) = \max \left(\bar{u} - \rho \cdot y, \int (y - c(\theta))^+ \lambda(\mathrm{d}\theta) \right). \qquad (7)$$

Since the function $y \mapsto \bar{u} - \rho \cdot y$ is decreasing while the function $y \mapsto \int (y - c)^+ \, \mathrm{d}\lambda$ is nondecreasing, the dual problem $\min_{y \geq 0} \varphi^{\text{fb}}(y)$ attains its optimum at y^{fb}, and its value is W^{fb} as stated. Finally, the balance condition implies (4).

8.2   Proof of Theorem 1

As in the proof of Proposition 1, we pass to the relaxed problem in which the balance condition is enforced as an inequality. Let C^{SP} be the set of all (P, Q) that satisfy the simplex bounds and participation constraints (SP-θ). For each Q, let $\bar{C}^{SP}(Q) = \{P : (P, Q) \in C^{SP}\}$. Consider the Lagrangian relaxation of the designer's problem with respect to the balance constraint. Let y be the dual variable. The Lagrangian is given by (5) with domain restricted to $(P, Q) \in C^{SP}$.

For every (Q, y), the problem

$$\text{maximize}_{P \in \bar{C}^{SP}(Q)} \mathscr{L}(Q, P; y)$$

factorizes over types and has the solution $\pi(\theta; Q, y) = \min(1 - Q, Q \cdot \nu(\theta)) \cdot \mathbf{1}_{[\theta > y]}$. The function $\ell(Q, y)$ in the theorem statement is therefore the reduced Lagrangian:

$$\ell(Q; y) = \max_{P \in \bar{C}^{SP}(Q)} \mathscr{L}(Q, P; y).$$

26

Since the domain $Q \in [0, 1]$ is compact, Sion's minimax theorem implies that strong duality holds, the primal problem

$$\max_{Q \in [0,1]} \min_{y \geq 0} \ell(Q; y)$$

admits an optimum, and the optimal value is given by W^*, as stated. If, further, the dual problem admits a (finite) solution y^*, then strong duality ensures that the mechanism (P^*, Q^*) with $P^*(\theta) = \pi(\theta; Q^*, y^*)$ solves the designer's problem.

Suppose instead that the dual problem admits no optimum, i.e., $y^* = \infty$. We claim that $\rho \cdot Q \geq \int_\theta \min(1 - Q, Q \cdot \nu(\theta)) \lambda(\mathrm{d}\theta)$ for every Q. Indeed, if there existed some $Q \in [0, 1]$ such that $\rho \cdot Q < \int_\theta \min(1 - Q, Q \cdot \nu(\theta)) \lambda(\mathrm{d}\theta)$, then there would exist a feasible mechanism satisfying Slater's condition. In that case, the dual problem would admit an optimum (Borwein and Lewis 2006, Theorem 3.2.8), a contradiction. This claim implies that any $(P, Q) \in C^{SP}$ satisfying the balance condition must satisfy $P(\theta) = \min(1 - Q, Q \cdot \nu(\theta))$ for every θ. Hence (P^*, Q^*) with $P^*(\theta) = \pi(\theta; Q^*, y^*) = \min(1 - Q, Q \cdot \nu(\theta))$ must solve the designer's problem.

It remains to prove that $y^* \geq y^{\mathrm{fb}}$. If $y^* = \infty$ then this inequality trivially holds as a strict inequality, so suppose that $y^* < \infty$. Let φ^{fb} and φ^* be the dual Lagrange functions of the two problems, with φ^{fb} as defined in (7) and

$$\varphi^*(y) = \max_{Q \in [0,1]} \ell(Q; y).$$

Then $\varphi^{\mathrm{fb}}(y) = \varphi^*(y)$ for every $y \leq y^{\mathrm{fb}}$, because for $y \leq y^{\mathrm{fb}}$ the maximum in (7) is achieved by $(P, Q) = (0, 1)$, which satisfies the participation constraints. Therefore, for every $y < y^{\mathrm{fb}}$ we have that

$$\varphi^*(y) = \varphi^{\mathrm{fb}}(y) > \varphi^{\mathrm{fb}}(y^{\mathrm{fb}}) = \varphi^*(y^{\mathrm{fb}}),$$

where the strict inequality follows from the fact that φ^{fb} is uniquely maximized at y^{fb}. Since y^* is the solution to the dual problem $\min_{y \geq 0} \varphi^*(y)$, it follows that $y^* \geq y^{\mathrm{fb}}$ as desired. Finally, if $W^* < W^{\mathrm{fb}}$ and $y^* < \infty$, then $\varphi^*(y^*) < \varphi^{\mathrm{fb}}(y^{\mathrm{fb}})$. Since $\varphi^*(y^{\mathrm{fb}}) = \varphi^{\mathrm{fb}}(y^{\mathrm{fb}})$, this implies that $y^{\mathrm{fb}} \neq y^*$, as desired.

8.3 Proof of Proposition 2

Let (Q^*, P^*) be the mechanism characterized in Theorem 1, with y^* the corresponding solution to the dual problem. Let Θ^* be the set of types with active participation constraints under this mechanism. By definition, $\theta \in \Theta^*$ if and only if $P^*(\theta) = Q^* \cdot \nu(\theta)$. If $Q^* = 0$, then $\Theta^* = \mathbf{R}_-$ is an interval. Going forward, we assume that $Q^* > 0$.

Theorem 1 implies that $\theta \in \Theta^*$ if and only if $c(\theta) < y^*$ and $\nu(\theta) \le (1 - Q^*)/Q^*$, i.e., Θ^* is the intersection of $I_r = c^{-1}([0, y^*))$ and $I_l = \nu^{-1}([0, (1 - Q^*)/Q^*])$. Since c and ν are monotone functions, I_r and I_l are both intervals. Therefore, Θ^* is itself an interval.

Let Θ^{fb} be the set of types that contribute under the mechanism characterized in Proposition 1. We claim that $\Theta^{\mathrm{fb}} \cap \Theta^* \ne \emptyset$. Indeed, in any convex optimization problem, all inactive constraints may be removed without affecting the optimality of the solution. But first-best welfare is achievable in the problem without the participation constraints in Θ^{fb}, and so $\Theta^{\mathrm{fb}} \cap \Theta^* = \emptyset$ would contradict our hypothesis.

Recall that $c(\theta) = -\theta$ under ordered types , so that $\Theta^{\mathrm{fb}} = (\theta^{\mathrm{fb}}, 0)$ with $\theta^{\mathrm{fb}} = -y^{\mathrm{fb}}$. We claim that $\theta^{\mathrm{fb}} \in \Theta^*$. Indeed, from Theorem 1 we know that $y^* > y^{\mathrm{fb}}$ and therefore $y^{\mathrm{fb}} \in [0, y^*)$ and $\theta^{\mathrm{fb}} = -y^{\mathrm{fb}} \in I_r$. Meanwhile, $\Theta^{\mathrm{fb}} \cap \Theta^* \ne \emptyset$ implies that $\Theta^{\mathrm{fb}} \cap I_l \ne \emptyset$. Since Θ^{fb} is an interval with lower endpoint θ^{fb} and I_l is a left half-line, it follows that $\theta^{\mathrm{fb}} \in I_l$.

8.4 Proof of Proposition 3

Let (Q^*, y^*) be as defined in Theorem 1. We first prove that if $\mathbb{E}_\lambda \nu = \infty$, then $Q^* \in (0, 1)$ and $y^* < \infty$, so that in particular $Q^* \in \arg\max_{Q \in (0,1)} \ell(Q; y^*)$. The balance condition implies the inequality $Q^*/(1 - Q^*) \le 1/\rho$, so that necessarily $Q^* < 1$. Next, the monotone convergence theorem implies that

$$\lim_{Q \downarrow 0} \int \min(\nu(\theta), (1 - Q)/Q) \, \lambda(\mathrm{d}\theta) = \mathbb{E}_\lambda \nu = \infty.$$

Hence for $Q > 0$ sufficiently small, the mechanism (Q, P) with $P(\theta) = \min(1 - Q, Q \cdot \nu(\theta))$ lies in C^* and satisfies the balance condition as a strict inequality, verifying Slater's condition and therefore (as established in the proof of Theorem 1) implying that $y^* < \infty$.

By construction, this mechanism satisfies participation for all agents and therefore delivers non-negative aggregate welfare. Lowering $P(\theta)$ for a positive measure of types until the balance condition holds with equality therefore yields a feasible mechanism generating positive payoffs for a positive measure of types, i.e., generating positive aggregate welfare. Since any feasible mechanism satisfying $Q = 0$ yields zero welfare, it must be that $Q^* > 0$.

We now establish the claimed limits. Since the reduced Lagrangian ℓ is a concave function of Q, its one-sided derivatives exist everywhere. Its right derivative may be calculated by writing ℓ as

$$\ell(Q; y) = Q \cdot (\bar{u} - \rho \cdot y) + Q \cdot \int \nu(\theta) \cdot (y - c(\theta))^+ \lambda(\mathrm{d}\theta)$$
$$+ \int \mathbf{1}_{[\nu \geq (1-Q)/Q]}(\theta) \cdot \left(1 - (\nu(\theta) + 1) \cdot Q\right) \cdot (y - c(\theta))^+ \lambda(\mathrm{d}\theta),$$

Since $(1 - Q)/Q$ is decreasing and continuous,

$$\frac{d}{d_+Q} \mathbf{1}_{[\nu \geq (1-Q)/Q]}(\theta) = 0$$

everywhere, meaning that

$$\frac{\partial \ell}{\partial_+ Q} = \bar{u} - \rho \cdot y + \int \nu(\theta) \cdot (y - c(\theta))^+ \lambda(\mathrm{d}\theta)$$
$$- \int \mathbf{1}_{[\nu \geq (1-Q)/Q]}(\theta) \cdot (\nu(\theta) + 1) \cdot (y - c(\theta))^+ \lambda(\mathrm{d}\theta).$$

Regrouping terms allows this derivative to be equivalently written

$$\frac{\partial \ell}{\partial_+ Q} = \bar{u} - \int \mathbf{1}_{[c \leq y]}(\theta) \cdot \left(\mathbf{1}_{[\nu < (1-Q)/Q]}(\theta) \cdot u(\theta) - \mathbf{1}_{[\nu \geq (1-Q)/Q]}(\theta) \cdot c(\theta)\right) \lambda(\mathrm{d}\theta)$$
$$- y \cdot \left(\rho + \Pr_\lambda(c \leq y \wedge \nu \geq (1-Q)/Q) - \int \mathbf{1}_{[c \leq y \wedge \nu < (1-Q)/Q]}(\theta) \cdot \nu(\theta) \lambda(\mathrm{d}\theta)\right).$$

The optimality condition $Q^* \in \arg\max_{Q \in (0,1)} \ell(Q; y^*)$ implies the one-sided first-order condition $\frac{\partial \ell}{\partial_+ Q}(Q^*; y^*) \leq 0$. Meanwhile, for every $Q > 0$ the balance condition may be rewritten

$$\rho = \int \mathbf{1}_{[c \leq y \wedge \nu < (1-Q)/Q]}(\theta) \cdot \nu(\theta) \lambda(\mathrm{d}\theta) + \frac{1-Q}{Q} \cdot \Pr(c \leq y \wedge \nu \geq (1-Q)/Q).$$

29

Since the balance condition is satisfied at (Q^*, y^*) and $Q^* > 0$, the first-order condition implies

$$\bar{u} - \int \mathbf{1}_{[c \leq y^*]}(\theta) \cdot \left(\mathbf{1}_{[\nu < \nu^*]}(\theta) \cdot u(\theta) - \mathbf{1}_{[\nu \geq \nu^*]}(\theta) \cdot c(\theta) \right) \lambda(\mathrm{d}\theta)$$
$$\leq y^* \cdot (\nu^* + 1) \cdot \mathrm{Pr}_\lambda(c \leq y^* \wedge \nu \geq \nu^*),$$

where $\nu^* = (1 - Q^*)/Q^*$. Since $c > 0$, this bound in turn implies the weaker bound

$$\int \mathbf{1}_{[c > y^*]}(\theta) \cdot u(\theta) \, \lambda(\mathrm{d}\theta) \leq y^* \cdot (\nu^* + 1) \cdot \mathrm{Pr}_\lambda(\nu \geq \nu^*).$$

Now, suppose that $\lim_{v \to \infty} v \, \mathrm{Pr}_\lambda(\nu > v) = 0$. Since $\lim_{\rho \to \infty} Q^* = 0$, it must be that $\lim_{\rho \to \infty} \nu^* = \infty$ and therefore $\lim_{\rho \to \infty} (\nu^* + 1) \cdot \mathrm{Pr}_\lambda(\nu \geq \nu^*) = 0$. Let $\underline{y} = \liminf_{\rho \to \infty} y^*$, and suppose by way of contradiction that $\underline{y} < \infty$. Then

$$\liminf_{\rho \to \infty} y^* \cdot (\nu^* + 1) \cdot \mathrm{Pr}_\lambda(\nu \geq \nu^*) = 0.$$

Meanwhile, viewed as a function of y, $\mathbf{1}_{[c > y]}(\theta)$ is an indicator function on an open set, meaning that it is lower semicontinuous. Hence

$$\liminf_{\rho \to \infty} \mathbf{1}_{[c > y^*]} \geq \mathbf{1}_{[c > \underline{y}]}.$$

Fatou's lemma therefore implies

$$\liminf_{\rho \to \infty} \int \mathbf{1}_{[c > y^*]}(\theta) \cdot u(\theta) \, \lambda(\mathrm{d}\theta) \geq \int \mathbf{1}_{[c > \underline{y}]}(\theta) \cdot u(\theta) \, \lambda(\mathrm{d}\theta)$$

Since c has full support on $\mathbf{R}_+$, the latter integral must be positive when $\underline{y} < \infty$, a contradiction.

Meanwhile, the left derivative of ℓ may be calculated by writing the reduced Lagrangian as

$$\ell(Q; y) = Q \cdot (\bar{u} - \rho \cdot y) + \int \mathbf{1}_{[\nu \leq (1 - Q)/Q]}(\theta) \cdot \left((\nu(\theta) + 1) \cdot Q - 1 \right) \cdot (y - c(\theta))^+ \lambda(\mathrm{d}\theta)$$
$$+ (1 - Q) \int (y - c(\theta))^+ \lambda(\mathrm{d}\theta).$$

Since $(1-Q)/Q$ is continuous and decreasing,

$$\frac{d}{d_- Q} \mathbf{1}_{[\nu \leq (1-Q)/Q]}(\theta) = 0$$

everywhere, meaning that

$$\frac{\partial \ell}{\partial_- Q} = \bar{u} - \rho \cdot y + \int \mathbf{1}_{[\nu \leq (1-Q)/Q]}(\theta) \cdot (\nu(\theta) + 1) \cdot (y - c(\theta))^+ \, \lambda(\mathrm{d}\theta)$$
$$- \int (y - c(\theta))^+ \, \lambda(\mathrm{d}\theta).$$

This derivative may be equivalently written

$$\frac{\partial \ell}{\partial_- Q} = \bar{u} - \int \mathbf{1}_{[c < y]}(\theta) \cdot \left(\mathbf{1}_{[\nu \leq (1-Q)/Q]}(\theta) \cdot u(\theta) - \mathbf{1}_{[\nu > (1-Q)/Q]}(\theta) \cdot c(\theta) \right) \lambda(\mathrm{d}\theta)$$
$$- y \cdot \left(\rho + \Pr_\lambda(c < y \wedge \nu > (1-Q)/Q) - \int \mathbf{1}_{[c < y \wedge \nu \leq (1-Q)/Q]} \cdot \nu(\theta) \, \lambda(\mathrm{d}\theta) \right).$$

The optimality condition $Q^* \in \arg\max_{Q \in (0,1)} \ell(Q; y^*)$ implies the one-sided first-order condition $\frac{\partial \ell}{\partial_- Q}(Q^*; y^*) \geq 0$. Meanwhile, for every $Q > 0$ the balance condition may be rewritten

$$\rho = \int \mathbf{1}_{[c < y \wedge \nu \leq (1-Q)/Q]} \cdot \nu(\theta) \, \lambda(\mathrm{d}\theta) + \frac{1-Q}{Q} \cdot \Pr(c < y \wedge \nu > (1-Q)/Q).$$

Since the balance condition is satisfied at (Q^*, y^*) and $Q^* > 0$, the first-order condition implies

$$\bar{u} - \int \mathbf{1}_{[c < y^*]}(\theta) \cdot \left(\mathbf{1}_{[\nu \leq \nu^*]}(\theta) \cdot u(\theta) - \mathbf{1}_{[\nu > \nu^*]}(\theta) \cdot c(\theta) \right) \lambda(\mathrm{d}\theta)$$
$$\geq y^* \cdot (\nu^* + 1) \cdot \Pr_\lambda(c < y^* \wedge \nu > \nu^*).$$

Let $\bar{c} = \int c \, \mathrm{d}\lambda$. Then the left-hand side of this inequality is in turn bounded above by $\bar{u} + \bar{c}$, so that

$$\bar{u} + \bar{c} \geq y^* \cdot (\nu^* + 1) \cdot \Pr_\lambda(c < y^* \wedge \nu > \nu^*).$$

Note that $\bar{u} + \bar{c}$ is independent of Q^*, y^*, and ρ.

Now, suppose that ν and c are negatively associated under λ and $\lim_{v \to \infty} v \Pr_\lambda(\nu >$

$v) = \infty$. The first condition implies the inequality

$$\Pr_\lambda(c < y^* \wedge \nu > \nu^*) \geq \Pr_\lambda(c < y^*) \cdot \Pr_\lambda(\nu > \nu^*),$$

yielding the bound

$$\bar{u} + \bar{c} \geq y^* \cdot \Pr_\lambda(c < y^*) \cdot (\nu^* + 1) \cdot \Pr_\lambda(\nu > \nu^*).$$

Since $\lim_{\rho \to \infty} Q^* = 0$, it must be that $\lim_{\rho \to \infty} \nu^* = \infty$. Hence $\lim_{v \to \infty} v \Pr_\lambda(\nu > v) = \infty$ implies that $\lim_{\rho \to \infty} (\nu^* + 1) \cdot \Pr_\lambda(\nu > \nu^*) = \infty$. Suppose by way of contradiction that $\limsup_{\rho \to \infty} y^* > 0$. Since c has full support on $\mathbf{R}_+$, the expression $y \cdot \Pr_\lambda(c < y)$ is an increasing function of y which is positive for all $y > 0$. Hence $\limsup_{\rho \to \infty} y^* \cdot \Pr_\lambda(c < y^*) > 0$. It follows that

$$\limsup_{\rho \to \infty} y^* \cdot \Pr_\lambda(c < y^*) \cdot (\nu^* + 1) \cdot \Pr_\lambda(\nu > \nu^*) = \infty.$$

But this limit must also be bounded above by $\bar{u} + \underline{u}$, which is finite given that $u, c \in \mathscr{L}^1(\lambda)$, a contradiction.

8.5 Proof of Proposition 4

Let $f = \lim_{\varepsilon \to 0} \Pr_\lambda(c < \varepsilon)/\varepsilon$, and assume that this limit exists. Also let $u_+(0) = \limsup_{t \to 0} u(t)$ and $u_-(0) = \liminf_{t \to 0} u(t)$.

Lemma 1. *If $u_-(0) > 0$ and $u_+(0) < \infty$, then $\limsup_{v \to \infty} v \Pr_\lambda(\nu > v) \leq u_+(0) \cdot f$ and $\liminf_{v \to \infty} v \Pr_\lambda(\nu > v) \geq u_-(0) \cdot f$.*

Proof. For any $v > 0$, let $\chi(v) = -\inf\{\theta : \nu(\theta) > v\}$, with the convention that $\chi(v) = 0$ if $v \geq \sup_\Theta \nu(\theta)$. Note that χ is nonincreasing given that ν is increasing. The assumption $u_-(0) > 0$ implies that $\nu(\theta) = -u(\theta)/\theta$ is unbounded above, hence $\chi(v) \in (0, \infty)$ for sufficiently large $v > 0$ and $\lim_{v \to \infty} \chi(v) = 0$. Going forward, we restrict attention to v large enough that $\chi(v) < \infty$.

Fix $\varepsilon \in (0, 1)$ small enough that $\chi((1 - \varepsilon) \cdot v) < \infty$. For all v, the bound $c < (1 - \varepsilon) \cdot \chi(v)$ implies that $\nu > v$, while $\nu > v$ implies that $c < (1 + \varepsilon) \cdot \chi(v)$. Hence

$$\Pr_\lambda(c < (1 - \varepsilon) \cdot \chi(v)) \leq \Pr_\lambda(\nu > v) \leq \Pr_\lambda(c < (1 + \varepsilon) \cdot \chi(v)).$$

Equivalently,

$$(1 - \varepsilon) \cdot v \cdot \chi(v) \cdot \frac{\mathrm{Pr}_\lambda(c < (1-\varepsilon) \cdot \chi(v))}{(1-\varepsilon) \cdot \chi(v)} \leq v \cdot \mathrm{Pr}_\lambda(\nu > v) \leq (1 + \varepsilon) \cdot v \cdot \chi(v) \cdot \frac{\mathrm{Pr}_\lambda(c < (1+\varepsilon) \cdot \chi(v))}{(1+\varepsilon) \cdot \chi(v)}.$$

Now let $v \to \infty$. Since χ is nonincreasing in v and satisfies $\lim_{v \to \infty} \chi(v) = 0$, we have

$$\lim_{v \to \infty} \frac{\mathrm{Pr}_\lambda(c < (1-\varepsilon) \cdot \chi(v))}{(1-\varepsilon) \cdot \chi(v)} = \lim_{v \to \infty} \frac{\mathrm{Pr}_\lambda(c < (1+\varepsilon) \cdot \chi(v))}{(1+\varepsilon) \cdot \chi(v)} = f.$$

Meanwhile, $\nu(-(1 + \varepsilon) \cdot \chi(v)) < v < \nu(-\chi(v)/(1 + \varepsilon))$ for all v, and so

$$\frac{v \cdot \chi(v)}{1 + \varepsilon} > \frac{\nu(-(1 + \varepsilon) \cdot \chi(v)) \cdot \chi(v)}{1 + \varepsilon} = \frac{u(-(1 + \varepsilon) \cdot \chi(v))}{(1 + \varepsilon)^2},$$

implying

$$\liminf_{v \to \infty} \frac{v \cdot \chi(v)}{1 + \varepsilon} \geq \frac{u_-(0)}{(1 + \varepsilon)^2}.$$

Since $u_-(0)$ is positive and finite by assumption, it follows that

$$\liminf_{v \to \infty} v \cdot \mathrm{Pr}_\lambda(\nu > v) \geq \frac{f \cdot u_-(0)}{(1+\varepsilon)^2}.$$

A very similar calculation yields the bound

$$\limsup_{v \to \infty} v \cdot \mathrm{Pr}_\lambda(\nu > v) \leq \frac{f \cdot u_-(0)}{(1-\varepsilon)^2}.$$

Since these bounds hold for any $\varepsilon \in (0, 1)$, taking $\varepsilon \to 0$ yields the claimed bounds. $\quad \square$

Combining this result with Proposition 3 immediately implies the desired result in the first case. As for the second case, recall that under ordered types, c is decreasing in θ while ν is increasing, so that ν and c are mechanically negatively associated under ordered types. Proposition 3 therefore also implies the result for the second case.

8.6   Proof of Proposition 5

We prove the Proposition using the following lemma, which may be useful in other situations to establish comparative statics of saddle points.

Lemma 2. *Let X be a sublattice of $\mathbf{R}_+ \times \mathbf{R} \times \mathbf{R}$ and $h : X \to \mathbf{R}$ be submodular and*

nonincreasing in its final argument. Then $\arg\max_x \min_y x \cdot h(x, y, z)$ *is nonincreasing[2].*

Proof. Let $H(y, z) = \min_y h(x, y, z)$, and consider the objective function

$$g(x, z) = \min_y x \cdot h(x, y, z) = x \cdot H(x, z).$$

Since h is submodular, it follows that H is as well (Topkis 1998, Theorem 2.7.6). Since H is also nonincreasing in z, Claim 1 below implies that g is also submodular. Its maximizers are therefore nonincreasing in z. $\qquad\square$

Claim 1. *Let X be a sublattice of $\mathbf{R}_+ \times \mathbf{R}$ and $f : X \to \mathbf{R}$ be submodular (supermodular) and nonincreasing (nondecreasing) in its second component. Then $g : X \to \mathbf{R}$ given by $g(x, y) = x \cdot f(x, y)$ is submodular (supermodular).*

Proof. The claim follows from general results about the product of supermodular functions (Topkis 1998, Lemma 2.6.4), but we provide a simpler direct proof here as an alternative.

We prove the submodularity case, with supermodularity following from the submodularity result applied to $-f$. Let $x < x'$. Then

$$g(x', y) - g(x, y) = (x' - x) \cdot f(x, y) + x \cdot (f(x', y) - f(x, y))$$

is nonincreasing in y since f is nonincreasing in y and submodular. $\qquad\square$

To complete the proof of Proposition 5 using Lemma 2, consider the designer's primal problem $\max_{Q \in [0,1]} \min_{y \geq 0} \ell(Q; y)$, with $\ell(Q; y) = Q \cdot h(Q, y, \rho)$ and

$$h(Q, y, \rho) = \bar{u} - \rho \cdot y + \int \min\left((1 - Q)/Q, \nu(\theta)\right) (y - c(\theta))^+ \, \lambda(\mathrm{d}\theta).$$

The function h is submodular, because it is a sum of components which are each a product of a nonincreasing function in one variable and a nondecreasing function in another. It is also nonincreasing in ρ. Monotonicity of the solution therefore follows from Lemma 2.

---

[2]Monotonicity is in the induced set ordering (Topkis 1998, Section 2.4).

8.7 Proof of Theorem 2

As in the proof of Proposition 1, we pass to the relaxed problem in which the balance condition is enforced as an inequality. Let C^{IC} be the set of all (R, P, Q) that satisfy the simplex bounds along with the participation and IC constraints. For each Q, let $\bar{C}^{\mathrm{IC}}(Q) = \{(R, P) : (R, P, Q) \in C^{\mathrm{IC}}\}$.

Consider the Lagrangian relaxation of the designer's problem with respect to the balance constraint. Let y be the dual variable. The Lagrangian is given by (5) with domain restricted to $(P, Q) \in C^{SP}$. Let $\ell^{\mathrm{IC}}(Q, y)$ be the reduced Lagrangian:

$$\ell^{\mathrm{IC}}(Q; y) = \max_{(R,P) \in \bar{C}^{\mathrm{IC}}(Q)} \mathscr{L}^{\mathrm{IC}}(R, P, Q; y),$$

where

$$\mathscr{L}^{\mathrm{IC}}(R, P, Q; y) = \int (R(\theta) \cdot \nu(\theta) - P(\theta)) \cdot c(\theta) \, \lambda(\mathrm{d}\theta) + y \cdot \int P \, \mathrm{d}\lambda - \rho \cdot Q \cdot y.$$

Note that the IC constraints may be rewritten

$$R(\theta) \cdot \nu(\theta) - P(\theta) \geq R(\theta') \cdot \nu(\theta) - P(\theta').$$

These are the familiar IC constraints in a single-good monopoly sale problem with valuations $\nu(\theta)$, probabilities $R(\theta)$ of receiving the good, and payments $P(\theta)$. Maximizing $\mathscr{L}^{\mathrm{IC}}$ with respect to $(R, P) \in \bar{C}^{IC}(Q)$ is therefore equivalent to solving the monopoly sale problem for a seller who maximizes a weighted sum of buyer welfare and revenue, under the uniform payment bound $P(Q) \leq 1 - Q$. (The remaining simplex bounds $R(\theta) \leq Q$ can be normalized to yield a quantity 1 of the good by redefining ν, and so they do not impact the form of the solution.)

As we show in Appendix A, the optimal mechanism with bounded payments takes one of two forms. One possibility is a posted price which delivers the good with the maximum probability, i.e., with $R(Q) = Q$. This case corresponds to $\bar{\nu} = \infty$ in the theorem statement. The second is a menu with two options: pay $1 - Q^*$ and receive the item with the maximum probability Q, or pay $\hat{P}$ and receive the item with some probability $\hat{R}$. This case corresponds to $\bar{\nu} < \infty$ in the theorem statement.

We next observe that, since ℓ^{IC} is the partially-optimized Lagrangian of a convex optimization problem, standard arguments imply that it is a concave-convex function.

(Convexity in y follows from the fact that ℓ^{IC} is the Lagrangian dual of the designer's problem enforcing a fixed Q. Concavity in Q follows from the fact that ℓ^{IC} is the value of the convex maximization problem with objective $\mathscr{L}^{\text{IC}}$, choice variables (R, P), and parameter Q.) Since the domain $Q \in [0, 1]$ is compact, Sion's minimax theorem therefore implies that strong duality holds and the primal problem

$$\max_{Q \in [0,1]} \min_{y \geq 0} \ell(Q; y)$$

admits an optimum Q^*. If, further, the dual problem

$$\min_{y \geq 0} \max_{Q \in [0,1]} \ell(Q; y)$$

admits a (finite) optimum y^*, then strong duality ensures that the mechanism (P^*, R^*) that maximizes $\mathscr{L}(R, P, Q^*; y^*)$ solves the designer's problem.

Suppose instead that the dual problem admits no optimum. We claim that $\rho \cdot Q \geq \int P \, d\lambda$ for every mechanism in C^{IC}. Indeed, if there existed some mechanism $(R, P, Q) \in C^{\text{IC}}$ satisfying $\rho \cdot Q < \int P \, d\lambda$, then Slater's condition would be satisfied and the dual problem would admit an optimum, a contradiction. Therefore, any mechanism in C^{IC} satisfying the balance condition also maximizes seller revenue among all mechanisms in C^{IC}. We again appeal to the characterization in Appendix A to conclude that in this case an optimal mechanism must have the structure stated in the Theorem.

8.8   Proof of Proposition 6

Assume first that the system process is stationary and let (R, P, Q) be its reduced form. We first derive the balance condition. If $Q = 0$, then the machine is always broken, agents never contribute, and the balance condition is satisfied. So assume that $Q > 0$.

For $T > 0$, let $n(T)$ be the number of times the machine breaks down during the interval time interval $[0, T]$. We prove the balance condition by showing that

$$\rho \cdot Q = \int P \, d\lambda = \lim_{T \to \infty} \frac{n(T)}{T} \quad \text{a.s.}$$

Since $Q > 0$ it must be that $\lim_{T \to \infty} n(T) = \infty$ a.s. Indeed, if $n(T)$ were bounded,

then from some point onward the machine would never be fixed, else it would eventually break again. But then from some later point onward the machine would remain broken forever, implying $Q = 0$, a contradiction.

Let $L_0 \geq 0$ be residual lifespan of the machine at time 0 (so $L_0 = 0$ if the machine is broken at time 0), and for $i \geq 1$ let L_i be the lifespan of the machine after the i-th time it is being fixed after time 0. For $T > 0$, let $W(T) = \int_0^T \mathbf{1}_{[\omega_t = G]} \, \mathrm{d}t$. Then for every $T > 0$,

$$L_1 + \cdots + L_{n(T)-1} \leq W(T) \leq L_0 + L_1 + \cdots + L_{n(T)}. \tag{8}$$

By the strong law of large numbers, $\lim_{n \to \infty} n^{-1} \sum_{i=1}^n L_n = 1/\rho$ a.s. Since $\lim_{T \to \infty} n(T) = \infty$ a.s., it follows from (8) that $\lim_{T \to \infty} W(T)/n(T) = 1/\rho$ a.s. Meanwhile, $\lim_{T \to \infty} W(T)/T = Q$ a.s. by the definition of Q. Therefore $\lim_{T \to \infty} n(T)/T = Q \cdot \rho$ a.s.

Let $M_0 \geq 0$ be the residual contribution quantum at time 0 (so $M_0 = 0$ if the machine is working at time 0), and for $i \geq 1$ let M_i be the contribution quantum after the i-th time breakdown subsequent to time 0. For $T > 0$, let $C(T)$ be the total amount of contribution in the time interval $[0, T]$. Then for every $T > 0$,

$$M_1 + \cdots + M_{n(T)-1} \leq C(T) \leq M_0 + M_1 + \cdots + M_{n(T)}. \tag{9}$$

By the strong law of large numbers, $\lim_{n \to \infty} n^{-1} \sum_{i=1}^n M_n = 1/\rho$ a.s. Since $\lim_{T \to \infty} n(T) = \infty$ a.s., it follows from (9) that $\lim_{T \to \infty} C(T)/n(T) = 1$ a.s. Meanwhile, $\lim_{T \to \infty} C(T)/T = \int P \, \mathrm{d}\lambda$ a.s. by the definition of P. Therefore $\lim_{T \to \infty} n(T)/T = \int P \, \mathrm{d}\lambda$ a.s.

We now establish the simplex bounds. At every point in time the realized usage of each agent is bounded by the 1 if the machine is working and by 0 if it is broken. Therefore for the population of agents of each type, expected usage is bounded by Q, the probability that the machine is working. In other words, $R(\theta) \leq Q$. Similarly, the realized contribution of each agent is bounded by the 0 if the machine is working and by 1 if it is broken. Therefore for the population of agents of each type, expected contributions are bounded by $1 - Q$, the probability that the machine is broken. In other words, $P(\theta) \leq 1 - Q$.

Assume now that (R, P, Q) satisfies the physical constraints, and consider the Markovian mechanism given by $\sigma(\theta, W) = R(\theta)/Q$ and $\sigma(\theta, B) = P(\theta)/(1 - Q)$, with the convention that $\sigma(\theta, W) = 0$ if $Q = 0$ and $\sigma(\theta, B) = 0$ if $Q = 1$. The machine state process under this mechanism is an alternating renewal process: When

37

the machine becomes working, it remains working for the lifespan with expectation $\mu_W = 1/\rho$. When the machine becomes broken, it remains broken for a random time with expectation $\mu_B = 1/\int \sigma(\theta, B) \, \lambda(d\theta)$.

The balance condition implies that $Q/(1-Q) = \mu_W/\mu_B$, which may be rearranged to find that the long-run proportion of time that the machine is working is $\frac{\mu_W}{\mu_W + \mu_B} = Q$. The quantity Q must therefore also be the probability with which the machine is working under the steady state induced by the mechanism. It then follows from definition of σ that $R(\theta) = Q \cdot \sigma(\theta, W)$ is the expected usage of an arriving agent of type θ while $P(\theta) = (1-Q) \cdot \sigma(\theta, B)$ is the expected contribution of an arriving agent of type θ.

8.9 Proof of Proposition 7

Assume first that the system process is stationary and let (R, P, Q) be its reduced form. The proof that (R, P, Q) satisfies the balance condition is very similar to the proof of Proposition 6 and is omitted. To prove the simplex bound, note that the non-anticipation assumption in (3) implies the so-called PASTA ("Poison arrivals see time average") property, under which the probability that an arriving agent of type θ encounters a working machine equals the steady-state uptime Q (Tijms 2003, Section 2.4). As a result, the simplex bounds follow from the fact that the agent cannot use or fix the machine more often than that action is available.

Assume now that (R, P, Q) satisfies the physical constraints, and consider the Markovian mechanism given by $\sigma(\theta, W) = R(\theta)/Q$ and $\sigma(\theta, B) = P(\theta)/(1-Q)$, with the convention that $\sigma(\theta, W) = 0$ if $Q = 0$ and $\sigma(\theta, B) = 0$ if $Q = 1$. The machine state process under this mechanism is an alternating renewal process with the same expectations as in the proof of Proposition 7. Therefore, by the same argument as in that proof the reduced form induced by this mechanism is (R, P, Q).

A   Monopoly sale problem with bounded transfers

Consider a generalized version of Myerson's single-good monopoly sale problem (Myerson 1981). Buyers are heterogeneous with types $\theta \in \Theta$. They have quasilinear payoffs over allocations and payments, with valuations $\nu : \Theta \to \mathbf{R}_+$ for the good. A mechanism is given by an allocation function $r : \Theta \to [0, 1]$ and payment rule

$p : \Theta \to \mathbf{R}_+$. (We rule out payments to buyers.) The seller's objective is a weighted average of buyer payoffs and payments of the form

$$\int \Big(r(\theta) \cdot \nu(\theta) - p(\theta) \Big) \, \mu_1(\mathrm{d}\theta) + \int p \, \mathrm{d}\mu_2 \tag{10}$$

for some finite measures μ_1, μ_2 over Θ.

A mechanism is incentive compatible (IC) if

$$r(\theta) \cdot \nu(\theta) - p(\theta) \geq r(\theta') \cdot \nu(\theta) - p(\theta')$$

for every $\theta, \theta' \in \Theta$. It is individually rational (IR) if $r(\theta) \cdot \nu(\theta) - p(\theta) \geq 0$ for all $\theta \in \Theta$. The seller chooses a mechanism to maximize (10) subject to IC and IR.

We first observe that, without loss, we may restrict attention to mechanisms satisfying $(r(\theta), p(\theta)) = (r(\theta'), p(\theta'))$ whenever $\nu(\theta) = \nu(\theta')$. Indeed, under any IC mechanism, all types with the same valuation receive the same utility. If such a mechanism does not offer all such types the same bundle, modify it to pool all types on the bundle involving the highest payment. This modification does not affect IC or IR and raises the seller's payoff. Given the restriction to such mechanisms, we may without loss relabel types so that the type space is the set of valuations $\mathbf{R}_+$.

As usual, the set of IC and IR mechanisms can be identified with the set of nondecreasing allocation functions. (Because the lowest type has valuation 0 for the good, the utility of the lowest type is not a free parameter in this environment.) We may further restrict attention to right-continuous allocation functions. Indeed, if $r(\nu) < r(\nu+)$ for some ν, then also $p(\nu) < p(\nu+)$. Additionally, IC implies that the bundle $(r(\nu+), p(\nu+))$ must deliver that type the same utility as $(r(\nu), p(\nu))$. Replacing that type's bundle with $(r(\nu+), p(\nu+))$ therefore preserves IC and IR, delivers that type the same utility, and raises the seller's revenue from that type. Going forward, we assume that r is both nondecreasing and right-continuous. The payment function and buyer payoffs corresponding to each such allocation function r are

$$\begin{aligned} p_r(\nu) &= \int_0^\nu v \, \mathrm{d}r(v), \\ U_r(\nu) &= \int_0^\nu r(v) \, dv. \end{aligned} \tag{11}$$

39

Standard theory says that the optimal mechanism is a posted price mechanism of the form $r = \mathbf{1}_{[\nu_1, \infty)}$ for some $\nu_1 \in \mathbf{R}_+$ (the price). The following lemma states an analog of this result for the case that payments are bounded. The proof follows a similar argument as in the corresponding proof for the case of unbounded payments, as in Börgers (2015, Section 2.5), with the twist that the set of extreme points of IC mechanisms is more complicated. We also develop the argument more generally, to allow for atoms in the type distributions. The argument is not new, but we do not know of a reference.

Proposition 8. *Consider the problem of maximizing objective* (10) *over all IC and IR mechanisms such that $0 \leq p(\nu) \leq 1$ for every $\nu \in \mathbf{R}_+$. The optimal allocation function is either of the form $r = \mathbf{1}_{[\nu_1, \infty)}$ for some $\nu_1 \leq 1$ or of the form $r = r_0 \mathbf{1}_{[\nu_0, \nu_1)} + \mathbf{1}_{[\nu_1, \infty)}$ for some $0 \leq \nu_0 \leq \nu_1$ and some $0 \leq r_0 \leq 1$ such that $r_0 \nu_0 + (1 - r_0)\nu_1 = 1$*

The first mechanism in the lemma is the familiar posted price with price ν_1. The second mechanism can be implemented as a menu with two options: pay $r_0 \cdot \nu_0$ and receive the good with probability r_0, or pay 1 and receive the good for sure.

Proof. The allocation functions $r : \mathbf{R}_+ \to [0, 1]$ of IR and IC mechanisms with payments bounded by 1 are nondecreasing and right-continuous, hence also u.s.c, and satisfy $\int_0^\infty v \, dr(v) \leq 1$. We can identify each such allocation with a finite measure over m over $\mathbf{R}_+$ such that

$$m(\mathbf{R}_+) \leq 1, \text{ and} \tag{12}$$

$$\int \nu \, m(d\nu) \leq 1, \tag{13}$$

so r is the cdf of m. The set $\mathcal{M}$ of all these measures is convex, and, by Helly's selection theorem, compact in the weak* topology.

By (11), for every buyer valuation ν, the utility $U_r(\nu)$ is linear and continuous in r and the payment $p_r(\nu)$ is linear and u.s.c. in r. Therefore, the objective function (10) is linear and u.s.c., which implies that the maximum is achieved at an extreme point. The set of extreme points of measures under a finite number of moment conditions is well-understood. (See, for example Karr (1983). The argument traces back to Karlin and Shapley (1953).)

The extreme points of $\mathcal{M}$ for which (12) is not saturated correspond to measures with a single atom and allocation function of the form $r = r_1 \mathbf{1}_{[\nu_1, \infty)}$ with $r_1 < 1$ and

$r_1\nu_1 \leq 1$, implementable by a mechanism that offers a price $r_1\nu_1$ in return for receiving the item with probability r_1. This mechanism cannot be optimal under our objective function, because both buyer utility and payments can be increased by offering the item with probability 1 for the same price.

The extreme points of $\mathcal{M}$ for which (12) is saturated are probability measures. If (13) is not saturated, then m is a probability measure with a single atom, corresponding to an allocation of the form $r = \mathbf{1}_{[\nu_1,\infty)}$ with and $\nu_1 \leq 1$, as in the proposition. If (13) is saturated, then m is a probability measure with at most two atoms, corresponding to an allocation of the form $r = r_0\mathbf{1}_{[\nu_0,\nu_1)} + \mathbf{1}_{[\nu_1,\infty)}$ for some $0 \leq \nu_0 \leq \nu_1$ and some $0 \leq r_0 \leq 1$ such that $r_0\nu_0 + (1-r_0)\nu_1 = 1$, as in the proposition. $\qquad\square$

Example 2 in Che and Gale (2000) presents an environment in which the bound on payments is binding and the optimal mechanism is a menu with two options. That example involves an environment in which the seller is uncertain of the buyer's budget as well as their valuation. However, the solution involves a menu with payments no higher than the lowest budget, and so it must also be the optimal mechanism in the problem where all buyers have the same (low) budget.

References

Admati, Anat R. and Motty Perry (1991). "Joint Projects without Commitment". *The Review of Economic Studies* 58 (2), pp. 259–276.

Akbarpour, Mohammad, Piotr Dworczak, and Scott Duke Kominers (2023). "Redistributive Allocation Mechanisms". Working paper.

Baccara, Mariagiovanna, SangMok Lee, and Leeat Yariv (2020). "Optimal dynamic matching". *Theoretical Economics* 15 (3), pp. 1221–1278.

Battaglini, Marco and Thomas R. Palfrey (2024a). "Dynamic Collective Action and the Power of Large Numbers". Working paper.

Battaglini, Marco and Thomas R. Palfrey (2024b). "Organizing for Collective Action: Olson Revisited". *Journal of Political Economy*, Forthcoming.

Börgers, Tilman (2015). *An introduction to the theory of mechanism design*. Oxford University Press, USA.

Borwein, Jonathan and Adrian Lewis (2006). *ConvexAnalysis and Nonlinear Optimization: Theory and Examples*. Springer.

Che, Yeon-Koo and Ian Gale (2000). "The optimal mechanism for selling to a budget-constrained buyer". *Journal of Economic theory* 92 (2), pp. 198–233.

Che, Yeon-Koo and Johannes Hörner (2018). "Recommender Systems as Mechanisms for Social Learning". *Quarterly Journal of Economics* 133 (2), pp. 871–925.

Che, Yeon-Koo and Olivier Tercieux (2021). "Optimal queue design". In: *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 312–313.

Compte, Olivier and Philippe Jehiel (2003). "Voluntary contributions to a joint project with asymmetric agents". *Journal of Economic Theory* 112 (2), pp. 334–342.

Condorelli, Daniele (2013). "Market and non-market mechanisms for the optimal allocation of scarce resources". *Games and Economic Behavior* 82, pp. 582–591.

Deb, Joyee, Aniko Oery, and Kevin R. Williams (2024). "Aiming for the Goal: Contribution Dynamics of Crowdfunding". *American Economic Review*, Forthcoming.

Doval, Laura and Vasiliki Skreta (2024). "Constrained information design". *Mathematics of Operations Research* 49 (1), pp. 78–106.

Dworczak, Piotr, Scott Duke Kominers, and Mohammad Akbarpour (2021). "Redistribution through markets". *Econometrica* 89 (4), pp. 1665–1698.

Fershtman, Chaim and Shmuel Nitzan (1991). "Dynamic voluntary provision of public goods". *European Economic Review* 35 (5), pp. 1057–1067.

Frazier, Peter, David Kempe, Jon Kleinberg, and Robert Kleinberg (2014). "Incentivizing Exploration". In: *Proceedings of the Fifteenth ACM Conference on Economics and Computation.* EC '14, pp. 5–22.

Guo, Yingni and Eran Shmaya (2019). "The interval structure of optimal disclosure". *Econometrica* 87 (2), pp. 653–675.

Hellwig, Martin F. (2003). "Public-Good Provision with Many Participants". *The Review of Economic Studies* 70 (3), pp. 589–614.

Immorlica, Nicole, Jieming Mao, Aleksandrs Slivkins, and Zhiwei Steven Wu (2019). "Bayesian Exploration with Heterogeneous Agents". In: *The World Wide Web Conference.* WWW '19, pp. 751–761.

Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian Persuasion". *American Economic Review* 101 (6), pp. 2590–2615.

Kang, Zi Yang (2023). "The Public Option and Optimal Redistribution". Working paper.

Karlin, Samuel and Lloyd S Shapley (1953). *Geometry of moment spaces.* Vol. 12. American Mathematical Soc.

Karr, Alan F (1983). "Extreme points of certain sets of probability measures, with applications". *Mathematics of Operations Research* 8 (1), pp. 74–85.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li (2017). "Persuasion of a privately informed receiver". *Econometrica* 85 (6), pp. 1949–1964.

Kremer, Ilan, Yishay Mansour, and Motty Perry (2014). "Implementing the "Wisdom of the Crowd"". *Journal of Political Economy* 122 (51), pp. 988–1012.

Le Treust, Maël and Tristan Tomala (2019). "Persuasion with limited communication capacity". *Journal of Economic Theory* 184, p. 104940.

Mansour, Yishay, Aleksandrs Slivkins, and Vasilis Syrgkanis (2015). "Bayesian Incentive-Compatible Bandit Exploration". In: *Proceedings of the Sixteenth ACM Conference on Economics and Computation.* EC '15, pp. 565–582.

Margaria, Chiara (2023). "Queueing to Learn". Working paper.

Marx, Leslie M. and Steven A. Matthews (2000). "Dynamic Voluntary Contribution to a Public Project". *The Review of Economic Studies* 67 (2), pp. 327–358.

Matthews, Steven A. (2013). "Achievable outcomes of dynamic contribution games". *Theoretical Economics* 8 (2), pp. 365–403.

Myerson, Roger B. (1981). "Optimal Auction Design". *Mathematics of Operations Research* 6 (1), pp. 58–73.

Norman, Peter (2004). "Efficient Mechanisms for Public Goods with Use Exclusions". *The Review of Economic Studies* 71 (4), pp. 1163–1188.

Pai, Mallesh M. and Philipp Strack (2024). "Taxing Externalities Without Hurting the Poor". Working paper.

Tijms, Henk C (2003). *A first course in stochastic models.* John Wiley and sons.

Topkis, Donald M (1998). *Supermodularity and complementarity.* Princeton university press.