Optimally Curating an Event

Author Name
Affiliation
email@example.com

Abstract

We study the optimal design of a self-financing event, a problem that requires balancing the recruitment of costly, positive-value participants with revenue-generating agents who may impose negative values on the event. We introduce a novel two-sided mechanism design framework with plus agents equipped with private costs and positive impact, and minus agents with private values and negative impact upon inclusion in the event, to maximize the overall quality of the event under a budget-balanced (BB) constraint so that the designer does not run a deficit.

We conduct a comprehensive study on the theory of optimal event curation on various utility functions. For an additive utility setting, we fully characterize the optimal incentive-compatible Bayesian mechanism under both the ex-ante and ex-post BB constraint. For submodular utility, we propose an ex-ante BB mechanism to achieve constant-factor approximation to the first-best outcome. We complement our results by investigating prior-free mechanism and conducting comparative statics.

1 Introduction

Mechanism design, especially auction design, has been one of the most foundational frameworks in recent decades, with numerous applications to online advertisements, spectrum auctions, and crowdsourcing platforms, to name a few—highlighted by the Nobel Prize in Economics in 2007 awarded to Myerson, Hurwicz, and Maskin for their foundation of mechanism design. In particular, two-sided mechanisms, also known as double auctions, serve as a primary tool to transfer goods between agents to maximize social efficiency. Most of the literature, however, has largely centered on two canonical objectives: maximizing revenue for the broker who governs the trade [Myerson and Satterthwaite, 1983; Kuang *et al.*, 2023], or maximizing social welfare [Deng *et al.*, 2022], typically defined as a function of agents' private valuations and payments. On the other hand, many modern platforms may realize arbitrary utilities depending on who is involved in the trades, independent of the agents' private valuations or transfers between them. Nevertheless, there have been only a few works studying such general objective functions, mostly in procurement auction frameworks—e.g., budget-feasible mechanism design [Singer, 2010; Anari *et al.*, 2014; Balkanski and Hartline, 2016] or all-pay auctions for crowdsourcing contests [Chawla *et al.*, 2019; DiPalantino and Vojnovic, 2009].

We discuss two representative examples to motivate our problem.[1]

Example 1: Conference curation. As an illustrative example, consider the challenge faced by the organizer of a prestigious event, such as an academic conference. The success of the event depends on how many attendees it attracts, or alternatively, how much welfare it provides to attendees. On one hand, the organizer seeks to attract renowned keynote speakers, popular artists, or flagship exhibitors. These participants generate significant positive value and draw more attendees, but they often require invitation fees, travel stipends, or other costs to secure their involvement. On the other hand, the organizer must fund these costs by securing corporate sponsorships

[1] Further applications can be found in Section 1.1.

or inviting advertisers. These participants provide essential revenue but may detract from the overall experience if their presence becomes overwhelming or overly commercial. The organizer's ultimate goal is to curate the optimal mix of these agents to maximize the attendees' utilities, ensuring that the organizer does not run a deficit by implementing the event.

Example 2: Double auction for LLMs. As another example, consider an LLM provider, such as Google with its Gemini model, which aims to deliver the most helpful and accurate responses to users. On one hand, to enhance the quality of its service, the provider can procure access to high-quality, up-to-date information from various *plus agents* such as news publishers, academic journals, or specialized data providers. Integrating these documents into a retrieval-augmented generation (RAG) framework allows the LLM to reduce hallucinations and provide more relevant answers, creating significant positive utility for the user. However, these content providers are strategic agents with private costs, and the LLM provider must design a mechanism to pay them for the right to use their data. On the other hand, the provider can fund these content acquisition costs by selling space within the LLM's output to *minus agents*, *i.e.,* advertisers.[2] These advertisers have private valuations for reaching the user and provide the necessary revenue, but their inclusion may detract from the user experience and the perceived objectivity of the answer. The LLM provider's challenge is to design a joint mechanism, balancing a procurement system for documents with an ad auction, to curate the optimal blend of retrieved knowledge and advertisements, thereby maximizing user satisfaction in a budget-balanced, self-financing ecosystem.

Our model. To formally capture this complex trade-off, we introduce a novel *two-sided mechanism design* problem for the optimal curation of agents with exogenous utilities, providing the first general framework for two-sided mechanisms where the designer's objective is neither revenue nor social welfare, but originates from an exogenous (possibly combinatorial) utility function.

Our model consists of two types of agents: *plus agents*, who contribute positively to the event's value but incur a private cost to participate, and *minus agents*, who have a private valuation for joining and are willing to pay, but may impose a negative externality on the event. The mechanism designer aims to maximize an exogenous utility function that maps an allocation to a nonnegative real value, which is monotone increasing in plus agents and monotone decreasing in minus agents, while satisfying (i) dominant-strategy incentive compatibility (DSIC) or Bayesian incentive compatibility (BIC), *i.e.,* agents report their true values; (ii) individual rationality (IR), *i.e.,* agents are not worse off by participating in the trade; and (iii) budget balance, *i.e.,* the mechanism does not run a deficit. We consider an additive utility function that can be represented as a simple sum of individual contributions, as well as a more complex combinatorial function capturing synergistic or substitution effects between agents. We analyze this curation problem in two standard settings of mechanism design: the Bayesian setting, where the designer has statistical knowledge of agent types, and the prior-free setting, which offers more robust performance guarantees without such distributional assumptions.

To highlight our main results:

1. For the case of additive utilities in the Bayesian setting, we provide a complete characterization of the optimal DSIC and IR mechanism under ex-ante BB constraint, and the optimal BIC and IR mechanism under ex-post BB constraint, both of which are efficiently computable.[3]

2. For separable submodular utility, we provide a posted pricing mechanism under regular distributions with polynomial running time that is a constant-factor approximation to the optimal Bayesian mechanism.[4]

In what follows, we overview our main results and techniques. Then, we introduce the formal problem setup, results, and their implications. The discussion on related works are deferred to the appendix due to page limit. In Appendix F, we also discuss several comparative statics on our mechanisms, revealing interesting spillover effects from each individual's statistical changes on the entire market.

[2] There has been increasing attention to integrating ads into LLM outputs, *e.g.,* see [Duetting *et al.*, 2024; Dubey *et al.*, 2024; Hajiaghayi *et al.*, 2024; Balseiro *et al.*, 2025].

[3] In Appendix C, we further discuss how our framework extends to the scenario with a set constraint.

[4] We also provide a simpler mechanism with weaker guarantees in Appendix E.

1.1 Further Applications

Notably, our framework subsumes broader applications beyond budget-balanced event curation.

Carbon market. Our framework also models the complex trade-offs in modern environmental policy. Consider a regulator tasked with reducing carbon emissions in a budget-neutral way. The success of their policy is measured by the net reduction in atmospheric carbon. On one hand, the regulator seeks to procure carbon abatement services from a market of plus agents. These agents, ranging from landowners paid to reforest their property to tech companies developing green energy, each face private costs to implement their solutions. While these initiatives provide direct positive environmental utility, they require significant public expenditure, compelling the regulator to design efficient procurement auctions to buy the most abatement for their budget. On the other hand, the regulator needs to fund these subsidies by regulating *minus agents*—the existing industrial polluters. By designing a carbon market, the regulator generates revenue from firms that have a high private value for emitting, but this activity imposes a negative environmental externality. The regulator's ultimate goal is to design this market to find the optimal carbon price, curating a portfolio of emissions allowances and abatement projects that maximizes environmental quality, ensuring the system is self-financing.

Urban planning. As another application, consider the challenge faced by a city planner aiming to foster a vibrant and livable urban environment. The success of a city is often measured by the quality of life it offers its residents, which depends on a careful balance of economic activity and public amenities. On one hand, a city seeks to enhance its appeal by procuring high-value public works and services from a diverse set of plus agents. These agents, such as construction firms contracted to build new parks or non-profit organizations that run cultural centers, each have private costs to complete their projects. While their contributions generate significant positive value for residents, they must be compensated via public funds, forcing the city to design procurement mechanisms to select the best projects without overspending. On the other hand, the city must fund these investments by fostering a strong economic base through *minus agents* such as industrial facilities or high-density commercial developments that provide tax revenue but may impose negative externalities. The planner's ultimate goal is to use tools like land auctions and procurement mechanisms to find the optimal mix of commercial and public projects, maximizing the city's overall livability while ensuring that public investments are sustainably funded by the tax revenue they help generate.

1.2 Our Results and Techniques

We briefly introduce our main results and the techniques used to obtain them. We begin with the additive utility setting in Section 4, where each plus agent imposes a fixed positive surplus $p_i \geq 0$ upon selection, and each minus agent induces a fixed negative impact $n_i \leq 0$ upon inclusion.[5] In this case, we provide a complete characterization of the optimal mechanism that maximizes the expected utility while remaining ex-ante BB, DSIC, and IR. We extend the classic Myersonian framework by characterizing the payment rule and derive the two-sided monotonicity of the allocation rule for any DSIC mechanism. This allows us to rewrite the budget-balancedness constraint as a function of the virtual costs and valuations of plus and minus agents, respectively. Then, we write the Lagrangian of the designer's resulting optimization problem. Finally, using the Karush-Kuhn-Tucker (KKT) conditions of the Lagrangian, we show that we can find the closed-form solution for the optimal Bayesian mechanism by determining the optimal Lagrange multiplier of the budget constraint—known as the shadow price—and deriving the resulting allocation function.[6] Then, by applying the interim transformation technique by [Brustle *et al.*, 2017], we show that our mechanism extends to the optimal BIC mechanism under ex-post BB constraint. Further discussions on the case of irregular distributions and ex-post BB DSIC mechanim are provided in Appendix B.

Second, we move beyond the additive utility to the submodular utility case and consider a utility function that is separable into the plus and minus agent sides. Namely, *including* a plus agent increases in a concave manner, while *excluding* a minus agent positively affects in a decreasing manner. [7] We first argue that the decision version of the offline problem setup (assuming truthful

[5] We refer to Section 3 for the detailed problem setup.

[6] In Appendix C, we also discuss how our framework extends to the scenario with a matroid constraint.

[7] In LLM auctions, one ad causes little distraction, but many ads collectively harm quality, motivating our supermodular negative-impact assumption. Since only a few ads are shown to preserve user experience, marginal degradation becomes significant.

bids) is still NP-hard to approximate within some constant factor, via a reduction from the submodular maximization problem with knapsack constraints. Accordingly, we provide a simple posted pricing mechanism under regular distributions that can be efficiently computed and achieves a constant-factor approximation to the optimal mechanism. Overall, our mechanism first generates a set of candidate budgets that can be used for plus agents. Then, we solve the plus and minus agent sides separately to exploit and procure the targeted budget—iteratively for each budget candidate—via the multilinear extension of the original optimization problems. Such multilinear extensions are guaranteed to approximate the optimal performance within a constant factor, thanks to the bounded correlation gap of submodular set functions shown by [Agrawal *et al.*, 2010]. Finally, we select the optimal budget that maximizes the expected utility among the candidates. Equipped with the fact that the expected utility under the multilinear extension is concave in the budget, we show that the error from using discretized budget candidates can be made small, implying that the resulting maximal mechanism is a good approximation to the optimal one. We complement this result by providing a mechanism with a weaker approximation guarantee but greater simplicity in Appendix E, and discuss how one could compute the multilinear extensions, as well as address irregular distributions and ex-post BB, in Appendix D.[8]

2 Related Works

Budget-feasible mechanism design. Our problem, in particular the plus agents' side market, has a close connection to the seminal budget-feasible mechanism design introduced by [Singer, 2010], where a buyer needs to procure a set of items from sellers given an exogenous budget constraint and a set function. [Singer, 2010] show that the standard VCG framework does not work, and provide several constant-factor approximate prior-free mechanisms for submodular utility. [Balkanski and Hartline, 2016] consider a Bayesian mechanism, where the benchmark is the optimal Bayesian mechanism rather than the offline benchmark. [Anari *et al.*, 2014] consider a budget-feasible mechanism with applications to the crowdsourcing market, justifying a *large market assumption* to derive stronger approximation guarantees. From then, there have been numerous works [Balkanski *et al.*, 2022; Jalaly and Tardos, 2021; Han *et al.*, 2023; Klumper and Schäfer, 2022] improving the approximation factors and generalizing the objective function.

Two-sided market. Plus and minus agents in our problem can be viewed as buyers and sellers in a two-sided market [McAfee, 1992; Myerson and Satterthwaite, 1983]. While seminal works in the literature [Myerson and Satterthwaite, 1983] focus on deriving optimal mechanisms for restricted cases or prove impossibility results, a recent work [Deng *et al.*, 2022] first proved that the gains-from-trade, a measure of social efficiency, can be approximated within a constant factor of the first-best outcome. Since then, there has been many works analyzing approximation factors or weakening the assumptions [Blumrosen and Dobzinski, 2014; Blumrosen and Dobzinski, 2021; Deng *et al.*, 2025]. Notably, similar to our problem where the mechanism designer aims to maximize a self-interested utility function, there has been a line of works studying the broker's mechanism—who governs the trade—to maximize its own profit [Kuang *et al.*, 2023; Eilat and Pauzner, 2021; Hajiaghayi *et al.*, 2025]. In stark contrast, in our setting, the designer has an exogenous utility function that can possibly depend (possibly combinatorially) on the allocation function.

3 Model

There are n *plus*-agents (*e.g.,* keynote speakers) and m *minus*-agents (*e.g.,* advertisers), denoted by $\mathcal{N} = \{1, \dots, n\}$ and $\mathcal{M} = \{n+1, \dots, n+m\}$, respectively. We consider a Bayesian setting where each plus agent i has a private cost c_i sampled from a distribution G_i supported on $[c_{\min}, c_{\max}]$ which represents how much it wants to be paid to participate in the event. Analogously, each minus agent j has a private valuation v_j sampled from a distribution F_j supported on $[v_{\min}, v_{\max}]$, denoting the willingness-to-pay to join the event. Let F and G be the product distributions of F_j's and G_i's. We will use boldface to denote vectors, *e.g.,* $\mathbf{c}$ and $\mathbf{v}$ for vectors of costs and values, respectively.

We consider a direct mechanism $(\mathbf{x}, \mathbf{t})$ that maps the reported bids $(\tilde{\mathbf{c}}, \tilde{\mathbf{v}})$ from the agents to an allocation vector $\mathbf{x} = (x_1, \dots, x_{n+m}) \in \{0, 1\}^{n+m}$ denoting whether each agent is selected to join the event and a payment

[8] In Appendix G, we also analyze prior-free mechanisms that do not require prior knowledge of the agents' distributions, and provide a constant-factor approximate mechanism against the first-best outcome under additive utility and the large market assumption.

vector $\mathbf{t} = (t_1, \dots, t_{n+m}) \in \mathbb{R}_{\geq 0}^{n+m}$ representing how much money each agent should pay or will be paid.[9] In particular, we will assume that $t_i > 0$ will be paid to plus agent i and $t_j > 0$ will be charged to minus agent j.

Given a direct mechanism $(\mathbf{x}, \mathbf{t})$, we write $\mathbf{b} = (b_1, \dots, b_{n+m})$ to denote the reported bids of the agents and b_{-i} to represent the vector of reports from all agents except agent i. We write $t_i(\mathbf{b}) = t_i(b_i, b_{-i})$ and $x_i(\mathbf{b}) = x_i(b_i, b_{-i})$ to denote the payment and allocation specific to agent i reporting b_i given others' bids b_{-i}.

Utilities. Given an allocation vector $\mathbf{x} \in \{0,1\}^{n+m}$, the utility of the mechanism designer is defined by a set function $u : \{0,1\}^{n+m} \to \mathbb{R}_{\geq 0}$. In particular, $u(\cdot)$ is *monotone increasing* over plus agents in a sense that for two allocation vectors $\mathbf{x}$ and $\mathbf{y}$ such that $\mathbf{x}_i \leq \mathbf{y}_i$ for $i \in \mathcal{N}$ and $\mathbf{x}_j = \mathbf{y}_j$ for $j \in \mathcal{M}$, we have $u(\mathbf{x}) \leq u(\mathbf{y})$, *i.e.,* allocating more plus agents do not decrease the utility. Similarly, $u(\cdot)$ is *monotone decreasing* over minus agents in a sense that for two allocation vectors $\mathbf{x}$ and $\mathbf{y}$ such that $\mathbf{x}_j \leq \mathbf{y}_j$ for $j \in \mathcal{M}$ and $\mathbf{x}_i = \mathbf{y}_i$ for $i \in \mathcal{N}$, we have $u(\mathbf{x}) \geq u(\mathbf{y})$, *i.e.,* allocating more minus agents do not increase the utility. In words, plus agents increase the utility but it requires cost to hire them, whereas minus agents decrease the utility but one can extract revenue from them. We write $\mathbf{1}_n$ to denote $(1, 1, \dots, 1) \in \mathbb{R}^n$, and similarly $\mathbf{0}_n$ to denote $(0, 0, \dots, 0) \in \mathbb{R}^n$. We assume that the utility function is always nonnegative, or equivalently $u(\mathbf{0}_n, \mathbf{1}_m) \geq 0$. We often abuse and write a set as the input, *e.g.,* $u(\mathcal{N})$ or $u(\mathcal{M})$, to denote the utility value given corresponding allocation vector.

A set function $f : 2^S \mapsto \mathbb{R}_{\geq 0}$ with finite set S is submodular if $f(X \cup \{x\}) - f(X) \geq f(Y \cup \{x\}) - f(Y)$ for any $X \subseteq Y \subseteq S$ and $x \in S \setminus Y$, and supermodular if the inverse of the inequality hols. We say f is monotone if $f(X) \leq f(Y)$ if $X \subseteq Y$. Note that these notations naturally extends to our utility function $u(\cdot)$.

Desiderata. Our objective is to design a mechanism that maximizes the utility while respecting several desirable properties. First, given others' bids b_{-i}, plus agent i's payoff from reporting $\tilde{c}_i$ when its true cost is c_i is:

$$U_i(\tilde{c}_i | c_i, b_{-i}) = t_i(\tilde{c}_i, b_{-i}) - c_i \cdot x_i(\tilde{c}_i, b_{-i})$$

Similarly, for fixed b_{-j}, minus agent j's utility from reporting $\tilde{v}_j$ when its true value is v_j is:

$$U_j(\tilde{v}_j | v_j, b_{-j}) = v_j \cdot x_j(\tilde{v}_j, b_{-j}) - t_j(\tilde{v}_j, b_{-j})$$

We say a mechanism is *dominant strategy incentive-compatible* (DSIC) if deviating from truthful reports do not increase the agents' payoffs. Formally, for any plus agent i, any true cost c_i, any lie c_i', and any reports b_{-i}:

$$t_i(c_i, b_{-i}) - c_i x_i(c_i, b_{-i}) \geq t_i(c_i', b_{-i}) - c_i x_i(c_i', b_{-i}), \tag{3.1}$$

and for any minus agent j, any true value v_j, any lie v_j', and any reports b_{-j}:

$$v_j x_j(v_j, b_{-j}) - t_j(v_j, b_{-j}) \geq v_j x_i(v_j', b_{-j}) - t_j(v_j', b_{-j}). \tag{3.2}$$

Further, we aim to design an *individually rational* mechanism such that no agent gets minus payoff from participating the mechanism. Formally, for plus agent i with cost c_i and minus agent j with value v_j, it should have the following respectively:

$$t_i(c_i, b_{-i}) - c_i \cdot x_i(c_i, b_{-i}) \geq 0$$
$$v_j \cdot x_j(v_j, b_{-j}) - t_j(v_j, b_{-j}) \geq 0$$

Our mechanism designer should not run a deficit by running the mechanism, *i.e.,* it should be *ex-ante (weakly) budget-balanced* (BB):

$$\int \sum_{j \in \mathcal{M}} t_j(\mathbf{v}, \mathbf{c}) \, dF(\mathbf{v}) dG(\mathbf{c}) \geq \int \sum_{i \in \mathcal{N}} t_i(\mathbf{v}, \mathbf{c}) \, dF(\mathbf{v}) dG(\mathbf{c}).$$

We often consider *ex-post budget-balancedness* (ex-post BB) where the budget constraint holds for every realization of valuations:

$$\sum_{j \in \mathcal{M}} t_j \geq \sum_{i \in \mathcal{N}} t_i.$$

We finally introduce virtual valuation and costs. Let f_j and g_i be the density functions of F_j and G_i, respectively.

Definition 1 (Virtual valuation and cost)**.** The virtual cost of plus agent i is defined as:

$$\psi_i(c_i) = c_i + \frac{G_i(c_i)}{g_i(c_i)}.$$

[9] One can easily check that it is without loss of generality to consider direct mechanisms via revelation principle.

The virtual valuation of minus agent j is defined as:

$$\phi_j(v_j) = v_j - \frac{1 - F_j(v_j)}{f_j(v_j)} \tag{3.3}$$

We say the agents' distributions are *regular* if $\psi_i(c_i)$ and $\phi_j(v_j)$ are monotone increasing for every i and j. Unless specified otherwise, we will restrict our focus to the regular distributions.

4 Optimal Mechanism for Additive Utility

We first consider a simplistic setting where the utility function is additive over each agent. Formally, the utility function is additive if it can be written as

$$u(\mathbf{x}) = \sum_{i \in \mathcal{N}} p_i x_i + \sum_{j \in \mathcal{M}} x_j n_j,$$

where $p_i \geq 0$ and $n_i \leq 0$.[10] That is, each plus agent i induces a fixed positive impact p_i upon its selection whereas each minus agent j incurs negative effect of n_j on inclusion to the event. Note that it is without loss of generality to consider the domain of allocation vector to be the intervals $[0, 1]^{n+m}$, due to the linearity of expectation given that the utility function is additive.

4.1 Optimal Ex-ante BB DSIC Mechanism

The following generalizes the Myerson's lemma:

Lemma 2. *For any DSIC mechanism, the allocation rule* $\mathbf{x}$ *is monotone decreasing in plus agent i's cost, and monotone increasing in minus agent j's value. Further, given others' bids b_{-i} and b_{-j}, the payment rule for plus agent i and minus agent j is uniquely characterized by:*

$$t_i(c_i, b_{-i}) = c_i \cdot x_i(c_i, b_{-i}) + \int_{c_i}^{c_{\max}} x_i(s, b_{-i})\, ds + U_{i,0}. \tag{4.1}$$

$$t_j(v_j, b_{-j}) = v_j \cdot x_j(v_j, b_{-j}) - \int_{v_{\min}}^{v_j} x_j(s, b_{-j})\, ds - U_{j,0}, \tag{4.2}$$

where $U_{i,0} = U_i(c_{\max}, b_{-i})$ and $U_{j,0} = U_j(v_{\min}, b_{-j})$.

The next crucial step is to analyze the *expected* payment from each agent from the designer's perspective. This allows us to replace payments with virtual costs and valuations.

Algorithm 1 Optimal Mechanism for Additive Utility

Input: Reported bids $\mathbf{c}$ and $\mathbf{v}$

Find the smallest shadow price $\lambda^* \geq 0$ such that the ex-ante budget balance constraint is met:

$$\mathbb{E}\left[\sum_{j \in \mathcal{M}} t_j(\lambda^*)\right] \geq \mathbb{E}\left[\sum_{i \in \mathcal{N}} t_i(\lambda^*)\right]$$

Define the final allocation rule $\mathbf{x}^* = \mathbf{x}(\lambda^*)$ based on the found shadow price:

$x_i^* \leftarrow \mathbb{I}[p_i - \lambda^* \psi_i(c_i) \geq 0]$ for all $i \in \mathcal{N}$.

$x_j^* \leftarrow \mathbb{I}[n_j + \lambda^* \phi_j(v_j) \geq 0]$ for all $j \in \mathcal{M}$.

Compute the final payments $\mathbf{t}^* = \mathbf{t}(\lambda^*)$ based on the allocation rule $\mathbf{x}^*$ using (4.1) and (4.2).

return $(\mathbf{x}^*, \mathbf{t}^*)$

Lemma 3. *The expected expenditure to plus agent i and expected revenue from minus agent j can be written as:*

$$\mathbb{E}_{c_i}\left[t_i(c_i)\right] = \mathbb{E}_{c_i}\left[x_i(c_i)\psi_i(c_i)\right]$$
$$\mathbb{E}_{v_j}\left[t_j(v_j)\right] = \mathbb{E}_{v_j}\left[x_j(v_j)\phi_j(v_j)\right]$$

We can now combine the characterization of implementable mechanisms with the concept of virtual valuation and cost to solve the designer's problem. The following theorem describes the optimal mechanism.

Theorem 4. *Under regular distributions, the mechanism described in Algorithm 1 is DSIC, IR, ex-ante BB, and maximizes the utility.*

The proof proceeds by first expressing the budget-balancedness constraint with virtual valuations and costs. Then, we write the Lagrangian of the designer's optimization problem, which is solely a function of allocation probabilities, p_i's and n_j's, and virtual valuations and costs. Then, we can regroup the terms by agent type with corresponding allocation probability. The stationarity condition of the KKT condition gives us the allocation rule described in the algorithm. Further, the complementary slackness guarantees that the budget constraint must bind or the Lagrangian multiplier should be zero, resulting in our choice of shadow price λ^*.

In Appendix C, we discuss how our framework extends to the scenario with a matroid constraint on the feasible allocation.

4.2 Optimal Ex-post BB BIC Mechanism

In this section we formalize the (now standard) payment transformation that converts any DSIC & ex-ante

[10]This is for sake of exposition, while our results easily generalize without such restriction.

BB mechanism into a BIC & *ex-post* BB (indeed, SBB) mechanism *without changing the allocation rule*. We then use it to argue that the DSIC & ex-ante BB optimal mechanism from Theorem 4 yields an *optimal* BIC & ex-post BB mechanism after the transformation.

Recall that ex-ante BB requires

$$\mathbb{E}\left[\sum_{j\in\mathcal{M}} t_j(\mathbf{v},\mathbf{c})\right] \;\geq\; \mathbb{E}\left[\sum_{i\in\mathcal{N}} t_i(\mathbf{v},\mathbf{c})\right].$$

Ex-post BB requires the above to hold pointwise for every realized profile; *strong* budget balance (SBB) requires equality ex post:

$$\sum_{j\in\mathcal{M}} t_j(\mathbf{v},\mathbf{c}) \;=\; \sum_{i\in\mathcal{N}} t_i(\mathbf{v},\mathbf{c}) \qquad \text{for all } (\mathbf{v},\mathbf{c}).$$

Interim Transformation. We state the transformation theorem first. It preserves the allocation rule and any objective that depends only on $\mathbf{x}$.

Lemma 5 (Payment transformation by [Brustle *et al.*, 2017])**.** *Let $(\mathbf{x},\mathbf{t})$ be a mechanism that is BIC, interim IR, and ex-ante BB with a nonnegative payment rule (i.e., $t_i(\cdot) \geq 0$ for $i \in \mathcal{N}$ and $t_j(\cdot) \geq 0$ for $j \in \mathcal{M}$). Then there exists another payment rule $\mathbf{t}'$ such that:*

1. Allocation preserved: $(\mathbf{x},\mathbf{t}')$ *uses the same allocation rule* $\mathbf{x}$.
2. BIC & interim IR preserved: $(\mathbf{x},\mathbf{t}')$ *remains BIC and interim IR, and each agent's interim expected payment is unchanged.*
3. Ex-post SBB: *For every profile* $(\mathbf{v},\mathbf{c})$*,*

$$\sum_{j\in\mathcal{M}} t'_j(\mathbf{v},\mathbf{c}) \;=\; \sum_{i\in\mathcal{N}} t'_i(\mathbf{v},\mathbf{c}).$$

Remark 6. Lemma 5 preserves the *ex-post* allocation rule $\mathbf{x}(\cdot)$ profile-by-profile; in particular, interim allocation probabilities remain identical. Since all of our objectives in this paper depend only on the allocation (transfers do not enter the objective), the designer's value is unchanged by the transformation.

Optimality. We now show that the DSIC & ex-ante BB optimum characterized in Section 4 yields an *optimal* BIC & ex-post BB mechanism after the transformation in Lemma 5.

Theorem 7 (Optimal BIC & ex-post BB)**.** *In the additive setting under regular distributions, let $(\mathbf{x}^*,\mathbf{t}^*)$ be the DSIC, IR, ex-ante BB mechanism from Theorem 4. Then there exists a payment rule $\widehat{\mathbf{t}}$ such that $(\mathbf{x}^*,\widehat{\mathbf{t}})$ is BIC, IR, and ex-post BB (indeed SBB), achieves the same expected utility as $(\mathbf{x}^*,\mathbf{t}^*)$, and is* optimal *among all BIC, IR, ex-post BB mechanisms.*

Corollary 8 (Equality of optimal values)**.** *In the additive setting with regular distributions,*

$$OPT_{\text{BIC, ex-ante BB}} \;=\; OPT_{\text{BIC, ex-post BB}} \;=\; \mathbb{E}\left[\,u(\mathbf{x}^*)\,\right],$$

where $\mathbf{x}^*$ is the allocation from Theorem 4.

5 Beyond Additive Utility

When the utility function $u(\cdot)$ is not additive but combinatorial, *e.g.,* submodular, the previous framework does not work. In particular, even the problem of solving the offline optimization problem, assuming truthful bids, is well-known to be NP-hard to solve for many class of functions. Thus, we aim to derive a mechanism that approximates the optimal mechanism.

Let S be the set of selected agents given an allocation vector $\mathbf{x}$. Accordingly, we consider the following class of separable functions

$$u(S) = v_p(S_p) - v_m(S_m),$$

where S_p and S_m are the set of *selected* plus and minus agents, respectively. Further, we assume v_p is monotone increasing submodular and v_m is monotone decreasing supermodular. This is practical since the marginal decrease in the audience experience of the event can be significantly exploding as we assign more advertisements, even though a few of them might be fine. From theoretical perspectives, this is reasonable since if we alternatively assume v_m is submodular, then the resulting function is a summation over submodular v_p and supermodular $-v_m$, which is admits no approximation algorithm.[11]

Note that maximizing u is equivalent to maximize

$$u(S) + v_m(\mathcal{M}) = v_p(S_p) + (v_m(\mathcal{M}) - v_m(S_m)).$$

In particular, we can write $u_m(S_m) = v_m(\mathcal{M}) - v_m(S_m)$ as a monotone increasing submodular function on *excluding* minus agents S_m. Thus, for notational convenience, we will consider the utility function $u(S) =$

[11] In general, maximizing a monotone supermodular always requires exponential query.

Algorithm 2 BEPP Mechanism

Input: Parameter $\varepsilon > 0$

Find $B_{\max}$ by solving single-dimensional revenue maximization on plus agents market.
Find minimal M such that $(1+\varepsilon)^M > B_{\max}$.
Define $\mathcal{B} \leftarrow \{(1+\varepsilon)^1, (1+\varepsilon)^2, \ldots, (1+\varepsilon)^M\}$.
for each budget target $B_k \in \mathcal{B}$ **do**
    Find $\mathbf{q}(B_k)$ from multilinear version of (5.1).
    Find $\mathbf{r}(B_k)$ from multilinear version of (5.2).
    $w(B_k) \leftarrow u_p(\mathbf{q}(B_k)) + u_m(\mathbf{r}(B_k))$.
end for
Find the best budget $B^* \leftarrow \arg\max_{B_k \in \mathcal{B}} W(B_k)$.
Set prices $\mathbf{t}^p \leftarrow G^{-1}(\mathbf{q}(B^*))$ and $\mathbf{t}^m \leftarrow F^{-1}(\mathbf{r}(B^*))$.
Offer posted prices $\mathbf{t}^p$ to pay plus agents.
Offer posted prices $\mathbf{t}^m$ to charge minus agents.

$u_p(S_p) + u_m(S_m)$ for monotone increasing nonnegative submodular functions u_p and u_m where S_p denotes the set of selected plus agents, and S_m denotes the set of excluded minus agents.

5.1 Budget-enumeration Posted Pricing

We first introduce our main mechanism, *budget-enumeration posted pricing* (BEPP) mechanism, presented in Algorithm 2. It retrieves a discretization parameter $\varepsilon > 0$ which controls the approximation factor as well as its time complexity. The following theorem formalizes the performance of the BEPP mechanism.

Theorem 9. *For any $\varepsilon > 0$,* BEPP *is DSIC, IR, and ex-ante BB, and achieves a $(1-1/e-O(\varepsilon))$-approximation to the optimal Bayesian mechanism for separable submodular utility and regular distributions.*

Proof overview. We now provide the overall proof steps to show that the BEPP mechanism is a constant-factor approximation of the optimal Bayesian mechanism for separable submodular utilities. The proof proceeds by defining a series of relaxations for the plus-agent and minus-agent subproblems, establishing an upper bound on the optimal welfare, and then showing that the expected welfare of our posted-price mechanism is within a constant factor of this bound.

We first formally define the two key relaxations of a set function, following the approaches used by [Balkanski and Hartline, 2016]. Let $u : 2^{\mathcal{O}} \to \mathbb{R}_{\geq 0}$ be a set function on a ground set $\mathcal{O}$. Let $\mathbf{p} = (p_k)_{k \in \mathcal{O}}$ be a vector of marginal probabilities. We first define the multilinear extension of u.

Definition 10 (Multilinear Extension)**.** The **multilinear extension** of u, denoted $U(\mathbf{p})$, is the expected value of the function when each element $k \in \mathcal{O}$ is included in a set S independently with probability p_k.

$$U(\mathbf{p}) = \mathbb{E}_{S \sim \mathbf{p}}[u(S)] = \sum_{S \subseteq \mathcal{O}} u(S) \prod_{k \in S} p_k \prod_{k \notin S} (1-p_k).$$

The key feature in analyzing approximation factors derived from multilinear extension is its relation to the following concave relaxation of u.

Definition 11 (Concave Closure)**.** The **concave closure** of u, denoted $U^+(\mathbf{p})$, is the maximum expected value of the function over any *correlated* distribution $\mathcal{D}$ on subsets of $\mathcal{O}$ with marginal probabilities $\mathbf{p}$.

$$U^+(\mathbf{p}) = \max_{\mathcal{D} \text{ with marginals } \mathbf{p}} \mathbb{E}_{S \sim \mathcal{D}}[u(S)]$$

Importantly, note here that the probability vectors can be correlated unlike the multilinear extension. The connection between these two relaxations is the following notion of correlation gap.

Theorem 12 ([Agrawal *et al.*, 2010])**.** *For any non-negative, monotone submodular function $u(\cdot)$, its multilinear extension and concave closure satisfy:*

$$\min_{\mathbf{p}} \frac{U(\mathbf{p})}{U^+(\mathbf{p})} \geq 1 - \frac{1}{e}$$

We structure the proof as a series of lemmas. First, we write concave closures of our original optimization problems for each side market. First, define $\mathbf{q} = (q_1, \ldots, q_n)$ as a vector that represents the marginal probability q_i plus agent i is selected, and similarly $\mathbf{r} = (r_1, \ldots, r_m)$ as a vector that represents the probability that minus agent j is excluded. In particular, we consider the following programming for the plus agent side for each B_k:

$$\text{(PLUS)} \quad \max_{\mathbf{q}} \quad u_p(\mathbf{q}) \quad \text{s.t.} \quad \sum_{i \in \mathcal{N}} q_i G_i^{-1}(q_i) \leq B_k, \tag{5.1}$$

and the following for the minus agent side:

$$\text{(MINUS)} \quad \max_{\mathbf{r}} \quad u_m(\mathbf{r}) \quad \text{s.t.} \quad \sum_{j \in \mathcal{M}} (1-r_j) F_j^{-1}(r_j) \geq B_k. \tag{5.2}$$

Multilinear versions of the above programmings require the probabilities to be independent, whereas concave closure allows the them to be correlated.

Given a budget B, we write $W_p^+(B)$ and $W_m^+(B)$ the optimum of the programming (PLUS) and (MINUS) in concave versions, $W_p(B)$ and $W_m(B)$ that in multilinear versions, and $\mathbf{q}(B)$ and $\mathbf{r}(B)$ be the optimal solutions of (PLUS) and (MINUS) in multilinear versions.

We will compare each optimization problem with the expected utilities obtainable from each side market. Formally, we will consider a class of mechanisms $\textsc{Mec}(B)$ parameterized by B such that its total expected payment to plus agents are at most B, and its total expected revenue from minus agents are at least B.

The performance of any mechanism given a budget, including the optimal one, will be upper-bounded by the concave closure as shown by the following lemma:

Lemma 13 (Concave relaxation)**.** *Given $B > 0$ and a mechanism $\textsc{Mec}(B)$ that is DSIC, IR and ex-ante BB, its expected utility u_p from plus agents is upper bounded by $W_p(B)$ and u_m (that from minus agent) is upper bounded by $W_m(B)$.*

Then, we consider a multilinear version of the programmings (PLUS) and (MINUS), and solving them gives us prices to offer as depicted in Algorithm 2. Thanks to Theorem 12, these prices guarantee constant approximation to the concave closure given B.

Lemma 14 (Posted pricing given a budget)**.** *For any budget B, offering prices $\mathbf{t}^p = G^{-1}(\mathbf{q}(B))$ to plus agents guarantees an expected utility of at least $(1 - 1/e) \cdot W_p(B)$ from plus agents, and offering $\mathbf{t}^m = F^{-1}(\mathbf{r}(B))$ to minus agents guarantees an expected utility of at least $(1 - 1/e) \cdot W_m(B)$ from minus agents.*

Therefore, one can conclude that if we know the expected budget used by the optimal mechanism, the corresponding posted pricing mechanism gives us a desired result. On the other hand, this is not known in advance, so we generate a finite list of candidate optimal budgets and take the maximal one. Since this list cannot always include the optimal budget due to the continuity of the entire budget space, we require the following concavity of the optimal expected utility with respect to the budget.

Lemma 15 (Concavity of the optimal utility)**.** *Under regular agent distributions, the plus-agent's optimal utility $W_p(B)$ and the minus-agent's optimal utility $W_m(B)$ given B are both concave functions of B.*

The above lemma results in the following lemma, implying that the error from the finiteness of the budget candidates can be made sufficiently small.

Lemma 16 (Error from budget discretization)**.** *Let $W(B) = W_p(B) + W_m(B)$ be the total utility function given B. Let B_k be the point from the discretized grid $\mathcal{B}$ such that $B_k \leq B^* \leq (1 + \varepsilon)B_k$. Then, we have $W(B_k) \geq (1 - \varepsilon)W(B^*)$.*

Combining, we get the desired result. In Appendix D, we discuss how to solve the multilinear programming given the budget constraint, and provide further discussions on ex-post BB and irregular distributions. We further provide a simpler mechanism with weaker guarantee in Appendix E, which uses a budget-feasible mechanism as a blackbox subroutine.

6 Conclusion

In this paper, we introduced and analyzed a novel two-sided mechanism design framework for self-financing events, where a designer's goal is to maximize an exogenous utility function by curating a set of participants. This problem requires balancing the recruitment of costly, high-impact plus agents with revenue-generating minus agents who may detract from the event's quality, all under a budget-balanced constraint.

Our analysis addresses this design problem primarily within the Bayesian setting. For the case of additive utilities, we fully characterize the optimal mechanisms that are incentive-compatible, IR, and BB. Extending our results to more general objectives, we then provide constant-factor approximation mechanisms for utility functions that are separable and submodular. Finally, to ensure performance without reliance on prior information, we also develop a constant-factor approximation for large markets in the prior-free setting.

Key avenues for future work include extending our analysis beyond submodular separable utility functions to model the complex interplay between participants. Broadening the scope of our prior-free results, by providing robust guarantees for submodular utilities or by relaxing the large market assumption, could also enhance the framework's practical applicability.

References

[Agrawal *et al.*, 2010] Shipra Agrawal, Yichuan Ding, Amin Saberi, and Yinyu Ye. Correlation robust stochastic optimization. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1087–1096. SIAM, 2010.

[Anari *et al.*, 2014] Nima Anari, Gagan Goel, and Afshin Nikzad. Mechanism design for crowdsourcing: An optimal 1-1/e competitive budget-feasible mechanism for large markets. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 266–275. IEEE, 2014.

[Balkanski and Hartline, 2016] Eric Balkanski and Jason D Hartline. Bayesian budget feasibility with posted pricing. In *Proceedings of the 25th International Conference on World Wide Web*, pages 189–203, 2016.

[Balkanski *et al.*, 2022] Eric Balkanski, Pranav Garimidi, Vasilis Gkatzelis, Daniel Schoepflin, and Xizhi Tan. Deterministic budget-feasible clock auctions. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2940–2963. SIAM, 2022.

[Balseiro *et al.*, 2025] Santiago Balseiro, Kshipra Bhawalkar, Yuan Deng, Zhe Feng, Jieming Mao, Aranyak Mehta, Vahab Mirrokni, Renato Paes Leme, Di Wang, and Song Zuo. Position auctions in ai-generated content. *arXiv preprint arXiv:2506.03309*, 2025.

[Blumrosen and Dobzinski, 2014] Liad Blumrosen and Shahar Dobzinski. Reallocation mechanisms. *arXiv preprint arXiv:1404.6786*, 2014.

[Blumrosen and Dobzinski, 2021] Liad Blumrosen and Shahar Dobzinski. (almost) efficient mechanisms for bilateral trading. *Games and Economic Behavior*, 130:369–383, 2021.

[Brustle *et al.*, 2017] Johannes Brustle, Yang Cai, Fa Wu, and Mingfei Zhao. Approximating gains from trade in two-sided markets via simple mechanisms. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 589–590, 2017.

[Chawla *et al.*, 2019] Shuchi Chawla, Jason D Hartline, and Balasubramanian Sivan. Optimal crowdsourcing contests. *Games and Economic Behavior*, 113:80–96, 2019.

[Chekuri *et al.*, 2011] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 783–792, 2011.

[Deng *et al.*, 2022] Yuan Deng, Jieming Mao, Balasubramanian Sivan, and Kangning Wang. Approximately efficient bilateral trade. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 718–721, 2022.

[Deng *et al.*, 2025] Yuan Deng, Jieming Mao, Balasubramanian Sivan, Kangning Wang, and Jinzhao Wu. Approximately efficient bilateral trade with samples. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 206–223, 2025.

[DiPalantino and Vojnovic, 2009] Dominic DiPalantino and Milan Vojnovic. Crowdsourcing and all-pay auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 119–128, 2009.

[Dubey *et al.*, 2024] Avinava Dubey, Zhe Feng, Rahul Kidambi, Aranyak Mehta, and Di Wang. Auctions with llm summaries. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 713–722, 2024.

[Duetting *et al.*, 2024] Paul Duetting, Vahab Mirrokni, Renato Paes Leme, Haifeng Xu, and Song Zuo. Mechanism design for large language models. In *Proceedings of the ACM on Web Conference 2024*, pages 144–155, 2024.

[Eilat and Pauzner, 2021] Ran Eilat and Ady Pauzner. Bilateral trade with a benevolent intermediary. *Theoretical Economics*, 16(4):1655–1714, 2021.

[Hajiaghayi *et al.*, 2024] MohammadTaghi Hajiaghayi, Sébastien Lahaie, Keivan Rezaei, and Suho Shin. Ad auctions for llms via retrieval augmented generation. *arXiv preprint arXiv:2406.09459*, 2024.

[Hajiaghayi *et al.*, 2025] Ilya Hajiaghayi, MohammadTaghi Hajiaghayi, Gary Peng, and Suho Shin. Gains-from-trade in bilateral trade with a broker. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4827–4860. SIAM, 2025.

[Han *et al.*, 2023] Kai Han, You Wu, He Huang, and Shuang Cui. Triple eagle: Simple, fast and practical budget-feasible mechanisms. *Advances in Neural Information Processing Systems*, 36:33894–33911, 2023.

[Jalaly and Tardos, 2021] Pooya Jalaly and Éva Tardos. Simple and efficient budget feasible mechanisms for monotone submodular valuations. *ACM Transactions on Economics and Computation (TEAC)*, 9(1):1–20, 2021.

[Klumper and Schäfer, 2022] Sophie Klumper and Guido Schäfer. Budget feasible mechanisms for procurement auctions with divisible agents. In *International Symposium on Algorithmic Game Theory*, pages 78–93. Springer, 2022.

[Kuang *et al.*, 2023] Zhonghong Kuang, Weiran Shen, and Fan Wu. Profit-maximizing mechanism in bilateral trade with interdependent valuations. *Available at SSRN 4474002*, 2023.

[McAfee, 1992] R Preston McAfee. A dominant strategy double auction. *Journal of economic Theory*, 56(2):434–450, 1992.

[Myerson and Satterthwaite, 1983] Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983.

[Myerson, 1981] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

[Singer, 2010] Yaron Singer. Budget feasible mechanisms. In *2010 IEEE 51st Annual Symposium on foundations of computer science*, pages 765–774. IEEE, 2010.

A Remaining Proofs in Section 4

A.1 Proof of Lemma 2

Proof of Lemma 2. The proof is quite straightforward extension of the standard Myerson's lemma, but we provide the proof here for completeness. We first prove the monotonicity condition for plus agents. The DSIC condition (3.1) must hold for any pair of reports, say c_1 and c_2.

$$t_i(c_1, b_{-i}) - c_1 \cdot x_i(c_1, b_{-i}) \geq t_i(c_2, b_{-i}) - c_1 \cdot x_i(c_2, b_{-i})$$
$$t_i(c_2, b_{-i}) - c_2 \cdot x_i(c_2, b_{-i}) \geq t_i(c_1, b_{-i}) - c_2 \cdot x_i(c_1, b_{-i})$$

Rearranging both inequalities to isolate the payment difference, $t_i(c_2, b_{-i}) - t_i(c_1, b_{-i})$, yields:

$$c_2[x_i(c_2, b_{-i}) - x_i(c_1, b_{-i})] \geq t_i(c_2, b_{-i}) - t_i(c_1, b_{-i})$$
$$c_1[x_i(c_2, b_{-i}) - x_i(c_1, b_{-i})] \leq t_i(c_2, b_{-i}) - t_i(c_1, b_{-i})$$

Combining these two:

$$c_1[x_i(c_2, b_{-i}) - x_i(c_1, b_{-i})] \leq c_2[x_i(c_2, b_{-i}) - x_i(c_1, b_{-i})],$$

which is equivalent to

$$(c_2 - c_1)[x_i(c_2, b_{-i}) - x_i(c_1, b_{-i})] \leq 0.$$

Assuming $c_2 > c_1$, then $x_i(c_2, b_{-i}) \leq x_i(c_1, b_{-i})$, *i.e.,* it is monotone decreasing on the plus agent's cost.

The argument for minus agents is symmetric. The DSIC condition from Equation (3.2) must hold for any two reported valuations, say v_1 and v_2.

$$v_1 \cdot x_j(v_1, b_{-j}) - t_j(v_1, b_{-j}) \geq v_1 \cdot x_j(v_2, b_{-j}) - t_j(v_2, b_{-j})$$
$$v_2 \cdot x_j(v_2, b_{-j}) - t_j(v_2, b_{-j}) \geq v_2 \cdot x_j(v_1, b_{-j}) - t_j(v_1, b_{-j})$$

Rearranging both inequalities to isolate the payment difference $t_j(v_2, b_{-j}) - t_j(v_1, b_{-j})$:

$$v_1[x_j(v_2, b_{-j}) - x_j(v_1, b_{-j})] \leq t_j(v_2, b_{-j}) - t_j(v_1, b_{-j})$$
$$v_2[x_j(v_2, b_{-j}) - x_j(v_1, b_{-j})] \geq t_j(v_2, b_{-j}) - t_j(v_1, b_{-j})$$

Combining these two inequalities yields:

$$v_1[x_j(v_2, b_{-j}) - x_j(v_1, b_{-j})] \leq v_2[x_j(v_2, b_{-j}) - x_j(v_1, b_{-j})]$$

which is equivalent to

$$(v_2 - v_1)[x_j(v_2, b_{-j}) - x_j(v_1, b_{-j})] \geq 0.$$

Assuming $v_2 > v_1$, this implies that $x_j(v_2, b_{-j}) \geq x_j(v_1, b_{-j})$, i.e., the allocation rule is monotone increasing in the minus agent's valuation.

Payment rule. Recall that

$$U_i(\tilde{c}_i | c_i, b_{-i}) = t_i(\tilde{c}_i, b_{-i}) - c_i \cdot x_i(\tilde{c}_i, b_{-i}).$$

Let $U_i(c_i, b_{-i})$ be the plus agent i's equilibrium utility. By the Envelope Theorem:

$$\frac{dU_i(c_i, b_{-i})}{dc_i} = -x_i(c_i, b_{-i})$$

Integrating from c_i to $c_{\max}$ gives the agent's utility:

$$U_i(c_i, b_{-i}) = \int_{c_i}^{c_{\max}} x_i(s, b_{-i})\, ds + U_i(c_{\max}, b_{-i})$$

Recalling that

$$U_i(c_i, b_{-i}) = t_i(c_i, b_{-i}) - c_i \cdot x_i(c_i, b_{-i}),$$

we finally have

$$t_i(c_i, b_{-i}) = c_i \cdot x_i(c_i, b_{-i}) + \int_{c_i}^{c_{\max}} x_i(s, b_{-i})\, ds + U_i(c_{\max}, b_{-i}).$$

Let $U_j(v_j, b_{-j})$ be the minus agent j's equilibrium utility given their true valuation v_j and others' bids b_{-j}. By the Envelope Theorem, the derivative of their utility with respect to their type is:

$$\frac{dU_j(v_j, b_{-j})}{dv_j} = x_j(v_j, b_{-j})$$

Integrating from $v_{\min}$ to v_j gives the agent's total utility:

$$U_j(v_j, b_{-j}) = \int_{v_{\min}}^{v_j} x_j(s, b_{-j})\, ds + U_j(v_{\min}, b_{-j})$$

Recalling the definition of utility for a minus agent,

$$U_j(v_j, b_{-j}) = v_j \cdot x_j(v_j, b_{-j}) - t_j(v_j, b_{-j}),$$

we can solve for the payment rule t_j by substitution:

$$t_j(v_j, b_{-j}) = v_j \cdot x_j(v_j, b_{-j}) - \left(\int_{v_{\min}}^{v_j} x_j(s, b_{-j})\, ds + U_j(v_{\min}, b_{-j}) \right).$$

This completes the characterization of the payment rule for minus agents. □

A.2 Proof of Lemma 3

Proof of Lemma 3. First, if $c_i = c_{\max}$ or $v_j = 0$, it is obvious that an optimal mechanism should set the allocation probability to be zero, thus without loss of generality we assume $U_{i,0} = U_{j,0} = 0$.

From Equation (4.2), the payment is:

$$t_j(v_j) = v_j \cdot x_j(v_j) - \int_{v_{\min}}^{v_j} x_j(s)\, ds$$

The designer's expected revenue from agent j is:

$$\mathbb{E}_{v_j}[t_j(v_j)] = \mathbb{E}_{v_j}[v_j x_j(v_j)] - \mathbb{E}_{v_j}\left[\int_{v_{\min}}^{v_j} x_j(s)\, ds\right]$$

Again, we use integration by parts on the second term. Note that this is equivalent to

$$\int_{v_{\min}}^{v_{\max}} \left(\int_{v_{\min}}^{v_j} x_j(s)\, ds\right) f_j(v_j)\, dv_j.$$

Using integrating by parts:

$$\begin{aligned}
&= \left[\left(\int_{v_{\min}}^{v_j} x_j(s, b_{-j})\, ds\right) \cdot (-(1 - F_j(v_j)))\right]_{v_{\min}}^{v_{\max}} - \int_{v_{\min}}^{v_{\max}} (-(1 - F_j(v_j)))\, x_j(v_j, b_{-j})\, dv_j \\
&= (0 - 0) + \int_{v_{\min}}^{v_{\max}} x_j(v_j, b_{-j})(1 - F_j(v_j))\, dv_j \\
&= \mathbb{E}_{v_j}\left[x_j(v_j)\frac{1 - F_j(v_j)}{f_j(v_j)}\right]
\end{aligned}$$

Substituting this back:

$$\begin{aligned}
\mathbb{E}_{v_j}[t_j(v_j)] &= \mathbb{E}_{v_j}[v_j x_j(v_j)] - \mathbb{E}_{v_j}\left[x_j(v_j)\frac{1 - F_j(v_j)}{f_j(v_j)}\right] \\
&= \mathbb{E}_{v_j}\left[x_j(v_j)\left(v_j - \frac{1 - F_j(v_j)}{f_j(v_j)}\right)\right]
\end{aligned}$$

The derivation for the plus agents' expected payment is symmetric, showing that their expected cost equals their expected virtual cost. From Equation (4.1), assuming $U_{i,0} = 0$, the payment to agent i is:

$$t_i(c_i) = c_i \cdot x_i(c_i) + \int_{c_i}^{c_{\max}} x_i(s)\, ds$$

The designer's expected cost for agent i is therefore:

$$\mathbb{E}_{c_i}[t_i(c_i)] = \mathbb{E}_{c_i}[c_i x_i(c_i)] + \mathbb{E}_{c_i}\left[\int_{c_i}^{c_{\max}} x_i(s)\, ds\right]$$

We use integration by parts on the second term. Let $g_i(\cdot)$ and $G_i(\cdot)$ be the PDF and CDF of costs for agent i.

$$\mathbb{E}_{c_i}\left[\int_{c_i}^{c_{\max}} x_i(s)\, ds\right] = \int_{c_{\min}}^{c_{\max}} \left(\int_{c_i}^{c_{\max}} x_i(s)\, ds\right) g_i(c_i)\, dc_i$$

Using integration by parts with $u = \int_{c_i}^{c_{\max}} x_i(s)\, ds$ and $dv = g_i(c_i)\, dc_i$, we get $du = -x_i(c_i)\, dc_i$ and $v = G_i(c_i)$.

This yields:

$$
\begin{aligned}
&= \left[\left(\int_{c_i}^{c_{\max}} x_i(s)\,ds\right) G_i(c_i)\right]_{c_{\min}}^{c_{\max}} - \int_{c_{\min}}^{c_{\max}} G_i(c_i)(-x_i(c_i))\,dc_i \\
&= (0-0) + \int_{c_{\min}}^{c_{\max}} x_i(c_i)G_i(c_i)\,dc_i \\
&= \mathbb{E}_{c_i}\left[x_i(c_i)\frac{G_i(c_i)}{g_i(c_i)}\right]
\end{aligned}
$$

Substituting this result back into the expected cost expression:

$$
\begin{aligned}
\mathbb{E}_{c_i}[t_i(c_i)] &= \mathbb{E}_{c_i}[c_i x_i(c_i)] + \mathbb{E}_{c_i}\left[x_i(c_i)\frac{G_i(c_i)}{g_i(c_i)}\right] \\
&= \mathbb{E}_{c_i}\left[x_i(c_i)\left(c_i + \frac{G_i(c_i)}{g_i(c_i)}\right)\right]
\end{aligned}
$$

This completes the derivation, showing that the expected payment to a plus agent is their expected virtual cost, $\psi_i(c_i) = c_i + G_i(c_i)/g_i(c_i)$, conditional on being selected. □

A.3 Proof of Theorem 4

Proof of Theorem 4. Recall that The designer's objective is to choose an allocation rule x to maximize the expected utility, subject to the constraints that the mechanism is incentive compatible (IC) and budget balanced (BB).

$$
\begin{aligned}
\max_{x} \quad & \mathbb{E}\left[\sum_i x_i(c_i)p_i + \sum_j x_j(v_j)n_j\right] \\
\text{subject to} \quad & (IC, IR) \quad x \text{ is implementable (monotone) and IR} \\
& (BB) \quad \mathbb{E}\left[\sum_j t_j(v_j) - \sum_i t_i(c_i)\right] \geq 0
\end{aligned}
$$

Note that the budget-balancedness constraint can be written as a function over the allocation rule:

$$
\mathbb{E}\left[\sum_j x_j(v_j)\phi_j(v_j) - \sum_i x_i(c_i)\psi_i(c_i)\right] \geq 0
$$

This is a constrained optimization problem. We introduce a Lagrange multiplier $\lambda \geq 0$ for the budget-balancedness constraint. The Lagrangian function L is:

$$
\begin{aligned}
L(\mathbf{x}, \lambda) &= \mathbb{E}\left[\sum_i x_i p_i + \sum_j x_j n_j\right] + \lambda\left(\mathbb{E}\left[\sum_j x_j\phi_j - \sum_i x_i\psi_i\right]\right) \\
&= \mathbb{E}\left[\sum_i x_i p_i + \sum_j x_j n_j + \lambda\left(\sum_j x_j\phi_j - \sum_i x_i\psi_i\right)\right]
\end{aligned}
$$

For simplicity, let us write $L(\mathbf{x}, \lambda) = W(\mathbf{x}) + \lambda B(\mathbf{x})$. We now regroup the terms by agent type:

$$L(\mathbf{x}, \lambda) = \mathbb{E}\left[\sum_i x_i(c_i) \underbrace{(p_i - \lambda\psi_i(c_i))}_{\text{virtual surplus of plus agent}} + \sum_j x_j(v_j) \underbrace{(n_j + \lambda\phi_j(v_j))}_{\text{virtual surplus of minus agent}}\right] \quad \text{(A.1)}$$

For a solution $(\mathbf{x}^*, \lambda^*)$ to be optimal, it must satisfy the following KKT conditions:

1. (Stationarity) The Lagrangian is maximized at $\mathbf{x}^*$. For our problem, this means $\mathbf{x}^*$ must be the allocation rule that maximizes $L(\mathbf{x}, \lambda^*)$. As can be seen from (A.1), this is the rule where each agent is selected if and only if their virtual surplus is non-negative.

$$x_i^*(c_i) = 1 \iff p_i - \lambda^*\psi_i(c_i) \geq 0$$
$$x_j^*(v_j) = 1 \iff n_j + \lambda^*\phi_j(v_j) \geq 0$$

2. (Primal Feasibility) The solution $\mathbf{x}^*$ must be feasible for the original problem. This simply means the budget constraint must be satisfied.

$$B(\mathbf{x}^*) \geq 0$$

3. (Dual Feasibility) The Lagrange multiplier for a $\geq$ constraint must be non-negative.

$$\lambda^* \geq 0$$

4. (Complementary Slackness) This is the most insightful condition. It states that the multiplier must be zero if the constraint is not binding (i.e., if there is a surplus).

$$\lambda^* \cdot B(\mathbf{x}^*) = 0$$

The KKT conditions, particularly Complementary Slackness, give us a concrete algorithm for finding the optimal λ^*. Let $B(\lambda)$ denote the expected budget surplus when the allocation rule is determined by a given λ, *i.e.,* $B(\lambda) = B(\mathbf{x}^*(\lambda))$.

Suppose we set $\lambda = 0$ and the budget constraint binds, *i.e.,* $B(\lambda) \geq 0$. Then this is the optimal allocation rule. For instance, if $p_i \geq 0$ and $n_j \leq 0$, the optimal allocation rule $\mathbf{x}^*(0)$ is to recruit every plus agent and no minus agent, according to the stationarity condition.

Now suppose we test $\lambda = 0$ and find that $B(0) < 0$. This violates the Primal Feasibility condition, so $(\lambda = 0)$ cannot be the optimal solution. From Dual Feasibility, we know $\lambda^* \geq 0$. Since $\lambda^* \neq 0$, we must have $\lambda^* > 0$. Now we turn to Complementary Slackness: $\lambda^* \cdot B(\mathbf{x}^*(\lambda^*)) = 0$. Since we know $\lambda^* > 0$, we must have $B(\mathbf{x}^*(\lambda^*)) = 0$. Note that $B(\lambda)$ is a non-decreasing function of λ, and $u(\mathbf{x}(\lambda))$ is a non-increasing function of λ since by the stationarity condition, increasing λ only recruits less plus agent and more minus agents. Thus, the optimal mechanism is to find the minimal λ^* that binds $B(\lambda^*) = 0$, and compute corresponding $\mathbf{x}^*(\lambda^*)$.

Finally, we must confirm that the derived allocation rule is implementable.

- For a plus agent i, the allocation rule is $x_i = 1 \iff p_i \geq \lambda\psi_i(c_i)$. Since $\psi_i(c_i)$ is assumed to be an increasing function of c_i, a higher cost c_i makes this condition harder to satisfy. Thus, $x_i(c_i)$ is non-increasing in c_i.
- For a minus agent j, the rule is $x_j = 1 \iff n_j + \lambda\phi_j(v_j) \geq 0$. Since $\phi_j(v_j)$ is an increasing function of v_j, a higher valuation v_j makes this condition easier to satisfy. Thus, $x_j(v_j)$ is non-decreasing in v_j.

Since the allocation rules are monotone, they are implementable. This completes the proof that the mechanism described in Theorem 4 is indeed the optimal one. □

A.4 Proof of Theorem 7

Proof. Proof of Theorem 7 $(\mathbf{x}^*, \mathbf{t}^*)$ *is optimal for BIC & ex-ante BB.* We write the BIC & ex-ante BB program explicitly, eliminate payments using the payment/virtual identities (Lemmas 2 and 3), form the Lagrangian, and solve it via KKT. This yields the same threshold allocation as in Theorem 4.

BIC & ex-ante BB as an allocation-only program. Under BIC and interim IR in a single-parameter environment, the interim selection probabilities are monotone in own type and payments satisfy the envelope identities. In particular (Lemma 3),

$$\mathbb{E}[\,t_i\,] \;=\; \mathbb{E}[\,x_i\,\psi_i(c_i)\,] \quad \text{and} \quad \mathbb{E}[\,t_j\,] \;=\; \mathbb{E}[\,x_j\,\phi_j(v_j)\,].$$

Therefore the designer's BIC problem with ex-ante BB can be written purely over allocation rules:

$$\begin{aligned}
\max_{\mathbf{x}} \quad & \mathbb{E}\left[\sum_{i\in\mathcal{N}} p_i\,x_i \;+\; \sum_{j\in\mathcal{M}} n_j\,x_j\right] \\
\text{s.t.} \quad & \mathbb{E}\left[\sum_{j\in\mathcal{M}} x_j\,\phi_j(v_j) \;-\; \sum_{i\in\mathcal{N}} x_i\,\psi_i(c_i)\right] \;\geq\; 0 \quad \text{(ex-ante BB)} \\
& \mathbf{x} \text{ is BIC-implementable and interim IR.}
\end{aligned}$$

BIC-implementability here is equivalent to (interim) monotonicity of each agent's selection probability in own type; interim IR fixes the integration constants in the envelope formula. Feasible payments then exist once $\mathbf{x}$ is fixed.

Lagrangian and separation. Introduce a multiplier $\lambda \geq 0$ for the ex-ante BB constraint. The Lagrangian is

$$\begin{aligned}
\mathcal{L}(\mathbf{x}, \lambda) &= \mathbb{E}\left[\sum_{i\in\mathcal{N}} p_i\,x_i + \sum_{j\in\mathcal{M}} n_j\,x_j\right] + \lambda\,\mathbb{E}\left[\sum_{j\in\mathcal{M}} x_j\,\phi_j(v_j) - \sum_{i\in\mathcal{N}} x_i\,\psi_i(c_i)\right] \\
&= \mathbb{E}\left[\sum_{i\in\mathcal{N}} x_i\,\big(p_i - \lambda\,\psi_i(c_i)\big) \;+\; \sum_{j\in\mathcal{M}} x_j\,\big(n_j + \lambda\,\phi_j(v_j)\big)\right].
\end{aligned}$$

For fixed λ, the integrand is linear in $x_k \in [0, 1]$; absent implementability constraints, pointwise maximization sets $x_k = 1$ whenever its coefficient is nonnegative and $x_k = 0$ otherwise.

Stationarity and implementability (threshold structure). Define the virtual-surplus coefficients

$$\Gamma_i^+(c_i; \lambda) \equiv p_i - \lambda\,\psi_i(c_i), \qquad \Gamma_j^-(v_j; \lambda) \equiv n_j + \lambda\,\phi_j(v_j).$$

Under regularity, $\psi_i(\cdot)$ and $\phi_j(\cdot)$ are (weakly) increasing, so the pointwise maximizers

$$x_i^\lambda(c_i) \;=\; \mathbb{I}\big[\Gamma_i^+(c_i;\lambda) \geq 0\big], \qquad x_j^\lambda(v_j) \;=\; \mathbb{I}\big[\Gamma_j^-(v_j;\lambda) \geq 0\big]$$

are nonincreasing in c_i and nondecreasing in v_j, respectively. Hence their interim selection probabilities are monotone in own type and are BIC-implementable (with interim IR payments given by the envelope integrals of Lemma 2). In fact, because these rules depend only on own type, they are ex-post monotone as well.

Choosing λ^*: primal/dual feasibility and complementary slackness. Let

$$B(\lambda) \;\equiv\; \mathbb{E}\left[\sum_{j\in\mathcal{M}} x_j^\lambda\,\phi_j(v_j) - \sum_{i\in\mathcal{N}} x_i^\lambda\,\psi_i(c_i)\right]$$

be the induced (expected) budget surplus. As λ increases, the plus-side thresholds become stricter and the minus-side thresholds weaker, so $B(\lambda)$ is weakly increasing and right-continuous. Let

$$\lambda^* \;=\; \min\{\lambda \geq 0 :\; B(\lambda) \geq 0\}.$$

By construction,

$$\text{(primal feasibility)} \quad B(\lambda^*) \geq 0, \qquad \text{(dual feasibility)} \quad \lambda^* \geq 0, \qquad \text{(compl. slackness)} \quad \lambda^* B(\lambda^*) = 0,$$

and $\mathbf{x}^* \equiv x^{\lambda^*}$ maximizes $\mathcal{L}(\cdot, \lambda^*)$ among BIC-implementable allocations (stationarity). Therefore $\mathbf{x}^*$ solves the BIC & ex-ante BB program.

Equality with the DSIC allocation. The same Lagrangian, threshold stationarity, and KKT logic underlie Theorem 4 for DSIC & ex-ante BB. Hence the optimizer here is exactly the same threshold allocation

$$x_i^*(c_i) \;=\; \mathbb{I}\big[p_i - \lambda^* \, \psi_i(c_i) \geq 0\big], \qquad x_j^*(v_j) \;=\; \mathbb{I}\big[n_j + \lambda^* \, \phi_j(v_j) \geq 0\big],$$

with payments given by the (BIC/DSIC) envelope integrals. This proves that $(\mathbf{x}^*, \mathbf{t}^*)$ attains the BIC & ex-ante BB optimum.

Transformation and Optimality Now we apply Lemma 5 to $(\mathbf{x}^*, \mathbf{t}^*)$. We obtain $(\mathbf{x}^*, \widehat{\mathbf{t}})$ that is BIC, interim IR, and ex-post SBB, with the same allocation rule and the same interim payments. Since the designer's objective depends only on $\mathbf{x}$, the expected utility of $(\mathbf{x}^*, \widehat{\mathbf{t}})$ equals that of $(\mathbf{x}^*, \mathbf{t}^*)$.

By Lemma 5 (and its converse), under BIC & interim IR the ex-ante BB and ex-post SBB feasible regions are linearly related via payment transformations while preserving $\mathbf{x}$. Therefore,

$$OPT_{\text{BIC, ex-ante BB}} \;=\; OPT_{\text{BIC, ex-post BB}}.$$

Since $(\mathbf{x}^*, \mathbf{t}^*)$ attains $OPT_{\text{BIC, ex-ante BB}}$ and $(\mathbf{x}^*, \widehat{\mathbf{t}})$ is feasible for BIC & ex-post BB with the same value, $(\mathbf{x}^*, \widehat{\mathbf{t}})$ is optimal for BIC & ex-post BB. □

B Additive Utility: Ex-post BB and Irreuglar Distributions

B.1 Ex-post BB DSIC Mechanism

While the optimal Bayesian mechanism from Theorem 4 is budget-balanced in expectation (ex-ante), for many applications, a stronger ex-post guarantee is desirable. Following the approach [Balkanski and Hartline, 2016], one can easily transform an ex-ante budget-balanced mechanism into an ex-post budget-balanced mechanism with small sacrifice in the utility, under a proper assumption. To this end, we first introduce the following *large market* assumption.

Definition 17 (Two-Sided Large Market Assumption)**.** A market is considered k**-large** with respect to an ex-ante policy $(\mathbf{x}, \mathbf{t})$ with an expected budget transfer of $B = \mathbb{E}\left[\sum_{i \in \mathcal{N}} t_i\right] = \mathbb{E}\left[\sum_{j \in \mathcal{M}} t_j\right]$, if for any realization of costs and valuations, the payment to any single agent is bounded:

- For every plus agent $i \in \mathcal{N}$, $t_i \leq B/k$.
- For every minus agent $j \in \mathcal{M}$, $t_j \leq B/k$.

This ensures that the outcome is not dependent on any single whale agent, allowing for the application of concentration bounds.

To extend our ex-ante budget-balanced mechanism to one that is ex-post budget-balanced with high probability, we rely on standard contention resolution scheme for sums of independent bounded random variables, as formulated by [Chekuri *et al.*, 2011].

First, we state the bound for the plus-agent side, which is a direct application of their result.

Lemma 18 ([Chekuri *et al.*, 2011])**.** *Given a budget B and a set of independent random variables $\{t_i\}_{i\in\mathcal{N}}$ for the payments to plus agents satisfying:*

- $\mathbb{E}\left[\sum_i t_i\right] \le (1-\varepsilon)B$,
- $t_i \in [0, B/k]$ *for all* i,
- $k > 2/\varepsilon$,

then the probability that the realized total cost exceeds the budget is small. Specifically, for $\varepsilon \in (0, 1/2)$*:*

$$\Pr\left[\sum_i t_i \le (1-1/k)B\right] \ge 1 - e^{-\varepsilon^2(1-\varepsilon)k/12}.$$

Next, we require an analogous concentration bound for the minus-agent side. The challenge is that we need a *lower tail* bound on revenue (i.e., the probability that revenue is unexpectedly low), whereas the above lemma is an *upper tail* bound on costs. We formally state the required lemma and then prove that it follows from the first.

Lemma 19. *Given a revenue target B and a set of independent random variables $\{t_j\}_{j\in\mathcal{M}}$ for the revenues from minus agents satisfying:*

- $\mathbb{E}\left[\sum_{j\in\mathcal{M}} t_j\right] \ge (1+\varepsilon)B$,
- $t_j \in [0, B/k]$ *for all* j,
- $k > 2/\varepsilon$,

then the probability that the realized total revenue falls below the target is small. Specifically, for $\varepsilon \in (0, 1/2)$*:*

$$Pr(\sum_{j\in\mathcal{M}} t_j \ge (1+1/k)B) \ge 1 - e^{-\varepsilon^2(1-\varepsilon)k/12}.$$

We omit the proof as it is a direct extension Lemma 18 (see Lemma 4.15 by [Chekuri *et al.*, 2011]) given the two-sided nature of Chernoff-Hoeffding bound.

Then, one can immediately reuse the machinery by [Balkanski and Hartline, 2016] to obtain the following result:

Theorem 20. *For $\varepsilon \in (0, 1/2)$, suppose the posted prices obtained by the multilinear extension programs for PLUS with budget $(1-\varepsilon)B$ and MINUS with budget $(1+\varepsilon)B$ satisfy $2/\varepsilon \le k \le B/t_i$ for any $i \in \mathcal{N}$ and $2/\varepsilon \le k \le B/t_j$ for any $j \in \mathcal{M}$. Then, this posted pricing mechanism is $(1-\varepsilon)(1-e^{-\varepsilon^2(1-\varepsilon)k/12})$ approximation to the optimal Bayesian mechanism under regular distributions.*

B.2 Irregular Distributions

If the agents' distributions are not regular, one can apply the standard ironing technique by [Myerson, 1981]. Essentially, for the plus agent, we can define the cost curve of agent i by $C_i(q_i) = q_i G_i^{-1}(q_i)$. The standard ironing technique constructs a convex hull $\bar{C}_i(q_i)$ of the cost curve, and define ironed virtual costs $\bar{\phi}_i(c_i) = \bar{C}'_i(q_i)$ be the derivative of it, which becomes monotone given the ironed cost curve. Then, [Myerson, 1981] show that the expected payment (under any incentive compatible) to any agent is at least $\bar{C}_i(q_i)$ if agent i's allocation probability is q_i, and this is further achievable by a simple randomized posted pricing mechanism. For minus agent side, we can analogously define a revenue curve to be $R_j(p_j) = p_j F_j^{-1}(1-p_j)$, and follow the same arguments, *e.g.,* see the two-sided market counterpart by [Myerson and Satterthwaite, 1983].

C Additive Utility under Matroid Constraint

In many practical applications, feasible allocations are restricted by structural constraints. For example, in conference scheduling, only one paper per session or topic can be selected; in sponsorship allocation, certain sponsors may not

Algorithm 3 Optimal Mechanism for Additive Utility with Matroid Feasibility

Input: Reported bids $\mathbf{c}$ and $\mathbf{v}$, matroid $(E, \mathcal{I})$ over agents $E = \mathcal{N} \cup \mathcal{M}$

Find the smallest shadow price $\lambda^* \geq 0$ such that the ex-ante budget balance constraint is met:

$$\mathbb{E}\left[\sum_{j\in\mathcal{M}} t_j(\lambda^*)\right] \geq \mathbb{E}\left[\sum_{i\in\mathcal{N}} t_i(\lambda^*)\right]$$

Define the final allocation rule $\mathbf{x}^* = \mathbf{x}(\lambda^*)$ based on the found shadow price:
  Compute virtual weights $w_i(\lambda^*) = p_i - \lambda^* \psi_i(c_i)$ for all $i \in \mathcal{N}$.
  Compute virtual weights $w_j(\lambda^*) = n_j + \lambda^* \phi_j(v_j)$ for all $j \in \mathcal{M}$.
  Sort all agents $k \in E$ by nonincreasing $w_k(\lambda^*)$.
  Initialize $S \leftarrow \emptyset$.
  for each agent k in the above order:
    if $w_k(\lambda^*) \leq 0$ **then break**
    if $S \cup \{k\} \in \mathcal{I}$ **then** $S \leftarrow S \cup \{k\}$.
  Set $x_k^* \leftarrow 1$ if $k \in S$, and $x_k^* \leftarrow 0$ otherwise.
Compute the final payments $\mathbf{t}^* = \mathbf{t}(\lambda^*)$ based on the allocation rule $\mathbf{x}^*$ using the integral formulas in Equations (4.1) and (4.2).
return $(\mathbf{x}^*, \mathbf{t}^*)$

appear together; and in team formation, skill overlaps or conflicts may restrict which participants can coexist. These situations can often be modeled by *matroid constraints* on the set of chosen agents.

Let $(E, \mathcal{I})$ be a matroid where $E = \mathcal{N} \cup \mathcal{M}$ is the ground set consisting of all positive and minus agents, and $\mathcal{I} \subseteq 2^E$ is the collection of feasible subsets satisfying the hereditary and augmentation properties. As before, each plus agent $i \in \mathcal{N}$ has utility p_i and cost c_i, and each minus agent $j \in \mathcal{M}$ has utility n_j and valuation v_j. The mechanism must choose a feasible set $S \in \mathcal{I}$ of agents to maximize the utility while maintaining budget balance.

We show that a modified version of Algorithm 1 provides the optimal utility given matroid constraints under regular distributions. For any fixed Lagrange multiplier $\lambda \geq 0$, define the *virtual weights*

$$w_i^\lambda(c_i) = p_i - \lambda\psi_i(c_i) \quad \text{for } i \in \mathcal{N}, \qquad w_j^\lambda(v_j) = n_j + \lambda\phi_j(v_j) \quad \text{for } j \in \mathcal{M}.$$

Then, recalling the proof of theorem 4, the goal of the allocation for a given λ would be to maximize

$$\sum_{i\in\mathcal{N}} x_i^\lambda(c_i) w_i^\lambda(c_i) + \sum_{j\in\mathcal{M}} x_j^\lambda(v_j) w_j^\lambda(v_j)$$

with x_i^λ and x_j^λ values now needing to satisfy the matroid independence constraints. Thus, we can define an allocation rule for λ that greedily selects agents in decreasing order of $w_i^\lambda(c_i)$ and $w_j^\lambda(v_j)$ weights, subject to feasibility under $\mathcal{I}$. Since greedy selection is optimal under matroid constraints for additive utility, this results in the allocation rule that optimizes our objective value for a given value of λ. Finally, to satisfy complementary slackness and find the overall optimum, we find the smallest value of λ that can satisfy budget-balance conditions, similar to Algorithm 1. This allocation rule is described in Algorithm 3.

The main adjustment needed from the proof of Theorem 4 to complete the proof here is to show that

1. The function $B(\lambda)$ calculating the budget surplus given λ is non-decreasing, and
2. The allocation rule x_i^λ and x_j^λ are implementable.

Starting with the second point, we can first consider any plus agent. Then, for a given value of λ, whether it is chosen or not depends only on how it is placed in the ordering using $w_i^\lambda(c_i) = p_i - \lambda\psi_i(c_i)$. As $\psi_i(c_i)$ is assumed to be an increasing function of c_i, increasing c_i can only decrease the weight of agent i, and push it further down the greedy ordering. So, x_i^λ will be non-increasing in c_i. Similarly, x_j^λ for $j \in \mathcal{M}$ will be non-decreasing in v_j, and therefore the allocation rule is implementable.

Discussion 21. Note that without any set constraint, the interim payment for each agent can be easily calculated by Lemma 3 since the optimal allocation simply allocates to every agent with nonnegative virtual valuation, *i.e.,* $x_i(c_i) = 1$ iff $\psi_i(c_i) = 1$ and $x_j(v_j) = 1$ iff $\phi_j(v_j)$ = 1. On the other hand, under the matroid constraint, we cannot simply rely on Lemma 3 since the allocation could vary depending on the set constraint. A simple way to calculate the expected payment for each agent is to directly calculate the expected value of Equations (4.1) and (4.2) by considering every possible realization of b_{-i} and b_{-j}. Note, however, that this requires exponential time. We leave it as a major open problem to whether it is possible to make it more efficient.

C.1 Proof of Lemma 13

Proof of Lemma 13. The proof is analogous to [Balkanski and Hartline, 2016], extending single-sided argument to two-sided one.

We first consider plus-agent side. The optimal mechanism for the plus-agent subproblem, which achieves welfare $W_p(B)$, solves the following program:

$$\begin{aligned} \max_{\mathbf{x}_p(\cdot),\mathbf{t}_p(\cdot)} \quad & \mathbb{E}\left[\, u_p(S_p) \,\right] \\ \text{s.t.} \quad & \mathbb{E}\left[\sum_{i\in\mathcal{N}} t_i \right] \leq B \\ & (\mathbf{x}_p, \mathbf{t}_p) \text{ is DSIC and IR.} \end{aligned}$$

We relax this program in three steps to find an upper bound on $W_p(B)$. First, by Myerson's lemma, for any IC mechanism, the expected payment to a plus agent (seller) equals their expected virtual cost. The budget constraint becomes $\mathbb{E}\left[\sum_{i\in\mathcal{N}} x_i\psi_i(c_i)\right] \leq B$.

Next, an allocation rule $\mathbf{x}_p(\cdot)$ induces a distribution $\mathcal{D}_p$ over selected sets S_p, with marginal probabilities $q_i = \mathbb{E}\left[\, x_i \,\right]$. By the definition of the concave closure, the expected utility is upper-bounded by the value of the concave closure on the marginals: $\mathbb{E}_{S_p\sim\mathcal{D}_p}\left[u_p(S_p)\right] \leq U_p^+(q)$.

Finally, for a given marginal selection probability q_i, the minimum possible expected virtual cost is achieved by a posted-price mechanism that offers price $\hat{c}_i = G_i^{-1}(q_i)$ (where G_i is the CDF of costs). The expected payment in this case is $q_iG_i^{-1}(q_i)$. Since the true mechanism might be less efficient, we have $\mathbb{E}\left[\, x_i\psi_i(c_i) \,\right] \geq q_iG_i^{-1}(q_i)$. To relax the budget constraint (i.e., to make the feasible set larger), we replace the expected virtual cost with this lower bound.

Combining these yields a final relaxed program, the concave closure program for plus agents, whose value is an upper bound on $W_p(B)$:

$$\begin{aligned} \max_{q\in[0,1]^n} \quad & U_p^+(q) \\ \text{s.t.} \quad & \sum_{i\in\mathcal{N}} q_iG_i^{-1}(q_i) \leq B \end{aligned}$$

Now we provide analogous argument for the minus-agent side. The optimal mechanism for the minus-agent subproblem, which achieves welfare $W_m(B)$, solves:

$$\begin{aligned} \max_{\mathbf{x}_m(\cdot),\mathbf{t}_m(\cdot)} \quad & \mathbb{E}\left[\, u_m(E_m) \,\right] \\ \text{s.t.} \quad & \mathbb{E}\left[\sum_{j\in\mathcal{M}} t_j \right] \geq B \\ & (\mathbf{x}_m, \mathbf{t}_m) \text{ is DSIC and IR.} \end{aligned}$$

Here, the decision is which agents to exclude, E_m. The revenue is generated by the admitted agents $S_m = \mathcal{M} \setminus E_m$. Let $r_j = \mathbb{E}[x_j]$ be the marginal probability that agent j is *excluded*. We apply a symmetric sequence of relaxations.

First, by Myerson's lemma, the expected revenue from minus agents equals their expected virtual valuations. The constraint becomes $\mathbb{E}\left[\sum_{j\in S_m} \phi_j(v_j)\right] \geq B$.

The allocation rule over exclusions induces a distribution $\mathcal{D}_m$ on sets E_m, with marginal exclusion probabilities r_j. The expected utility is upper-bounded by the concave closure: $\mathbb{E}_{E_m\sim\mathcal{D}_m}[u_m(E_m)] \leq U_m^+(r)$.

Finally, the constraint is $\mathbb{E}\left[\sum_{j\in\mathcal{M}\setminus E_m} \phi_j(v_j)\right] \geq B$. The probability that agent j is admitted is $(1 - r_j)$. To relax this constraint (make the feasible set of r vectors larger), we must consider the *maximum* possible expected revenue for a given admission probability $(1 - r_j)$. This is achieved by admitting agents with the highest virtual valuations. A posted admission price q_j such that $\Pr[v_j \geq q_j] = 1 - r_j$ (i.e., $F_j(q_j) = r_j$) would generate expected revenue of $q_j(1-r_j) = F_j^{-1}(r_j)(1-r_j)$. The true mechanism's revenue may be lower, so we replace the expected revenue with this upper bound to relax the constraint.

This yields the concave closure program for minus agents, whose value is an upper bound on $W_m(B)$:

$$\begin{aligned} \max_{r\in[0,1]^m} \quad & U_m^+(r) \\ \text{s.t.} \quad & \sum_{j\in\mathcal{M}} (1-r_j)F_j^{-1}(r_j) \geq B \end{aligned}$$

Thus, we have established that the value of the optimal mechanism for each subproblem is upper-bounded by the value of its corresponding concave closure program. □

C.2 Proof of Lemma 14

Proof of Lemma 14. We prove the result for each subproblem separately.

We first consider plus-agent side. Let q_{cc}^* be the optimal solution to the concave closure program for the plus-agent side, whose value $U_p^+(q_{cc}^*)$ is an upper bound on $W_p(B)$ by Lemma 13. Let q_{ml}^* be the optimal solution to the corresponding multilinear extension program. The expected utility achieved by our mechanism $M_p(B)$ is, by definition, $U_p(q_{ml}^*)$. We can lower-bound this value as follows:

$$\begin{aligned} U_p(q_{ml}^*) &\geq U_p(q_{cc}^*) \\ &\geq \left(1 - \frac{1}{e}\right) U_p^+(q_{cc}^*) && \text{(By Theorem 12)} \\ &\geq \left(1 - \frac{1}{e}\right) W_p(B) && \text{(By Lemma 13)} \end{aligned}$$

Thus, the mechanism $M_p(B)$ achieves at least a $(1 - 1/e)$-approximation of $W_p(B)$.

The argument for minus-agent side is symmetric. Let r_{cc}^* be the optimal solution to the concave closure program for the minus-agent side (maximizing $U_m^+(r)$ subject to the revenue constraint). We know its value $U_m^+(r_{cc}^*)$ is an upper bound on $W_m(B)$. Let r_{ml}^* be the optimal solution to the corresponding multilinear extension program. The utility achieved by our mechanism $M_m(B)$ is $U_m(r_{ml}^*)$. We can lower-bound this value with the same chain of reasoning:

$$\begin{aligned} U_m(r_{ml}^*) &\geq U_m(r_{cc}^*) \\ &\geq \left(1 - \frac{1}{e}\right) U_m^+(r_{cc}^*) && \text{(By Theorem 12)} \\ &\geq \left(1 - \frac{1}{e}\right) W_m(B) && \text{(By Lemma 13)} \end{aligned}$$

Thus, the mechanism $M_m(B)$ also achieves at least a $(1-1/e)$-approximation of its respective optimal value, $W_m(B)$.

□

C.3 Proof of Lemma 15

Proof of Lemma 15. The sum of concave functions is concave, so the total optimal utility function $W(B) = W_p(B) + W_m(B)$ will be concave if its components are. We provide the full proof for both $W_p(B)$ and $W_m(B)$.

First, consider the plus-agent side. The optimal utility $W_p(B)$ is defined as the value of the multilinear extension program:

$$\begin{aligned} W_p(B) = \max_{q \in [0,1]^n} \quad & U_p(q) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}} q_i G_i^{-1}(q_i) \leq B \end{aligned}$$

where $U_p(q)$ is the multilinear extension of the submodular utility function $u_p(\cdot)$.

This value function $W_p(B)$ is the result of maximizing a concave function over a convex set. First, the objective $U_p(q)$ is a concave function of the probability vector q, as it is the multilinear extension of a submodular function [Balkanski and Hartline, 2016]. Second, the expected cost for agent i, $C_i(q_i) = q_i G_i^{-1}(q_i)$, is a convex function for regular distributions. The total cost $C(q) = \sum_i C_i(q_i)$ is therefore also convex, making the feasible set $\mathcal{Q}_B = \{q \mid C(q) \leq B\}$ a convex set. Standard results in convex optimization show that the value function of a program maximizing a concave objective over an expanding convex set is concave in the constraint parameter. Thus, $W_p(B)$ is concave in B.

The proof for the minus-agent side is analogous. The optimal utility $W_m(B)$ is the value of the program:

$$\begin{aligned} \max_{\mathbf{r}} \quad & U_m(r) \\ \text{s.t.} \quad & \sum_{j \in \mathcal{M}} (1 - r_j) F_j^{-1}(r_j) \geq B \end{aligned}$$

where r_j is the probability of *excluding* agent j, and $U_m(r)$ is the multilinear extension of the submodular utility $u_m(\cdot)$ defined over excluded agents.

To demonstrate concavity, we re-cast this as an equivalent procurement problem, which is structurally identical to the plus-agent case. Let $R_{max} = \sum_{j \in \mathcal{M}} \mathbb{E}_{v_j}[\max_{q_j}\{q_j \mid v_j \geq q_j\}]$ be the maximum possible expected revenue from all minus agents. The revenue constraint can be rewritten in terms of *forgone revenue*:

$$\begin{aligned} \mathbb{E}\left[\text{Revenue from admitted agents}\right] &\geq B \\ R_{max} - \mathbb{E}\left[\text{Forgone revenue from excluded agents}\right] &\geq B \\ \mathbb{E}\left[\text{Forgone revenue from excluded agents}\right] &\leq R_{max} - B \end{aligned}$$

The cost of excluding agent j is the expected revenue we forego. Let this be $C_j^E(r_j)$. The minus-agent problem is now equivalent to:

$$\begin{aligned} \max_{r \in [0,1]^m} \quad & U_m(r) \\ \text{s.t.} \quad & \sum_{j \in \mathcal{M}} C_j^E(r_j) \leq R_{max} - B \end{aligned}$$

This is a procurement problem where we "buy" exclusions to gain utility u_m, and our budget for these costs is $B' = R_{max} - B$. From the analysis of the plus-agent side, we know that the optimal value of this program, let's call it $W_m^{procure}(B')$, is a concave function of its budget, B'.

Our original value function is $W_m(B) = W_m^{procure}(R_{max} - B)$. This is the composition of a concave function, $W_m^{procure}(\cdot)$, with an affine transformation of B. The composition of a concave function with an affine function is concave. Therefore, $W_m(B)$ is concave in B.

Since both $W_p(B)$ and $W_m(B)$ are concave, their sum $W(B)$ is also concave. □

C.4 Proof of Lemma 16

Proof of Lemma 16. The function $W(B)$ is the sum of two concave functions, and is therefore concave. The point B_k can be written as a convex combination of 0 and B^*: $B_k = \frac{B_k}{B^*}B^* + (1 - \frac{B_k}{B^*}) \cdot 0$. By the definition of concavity:

$$W(B_k) \geq \frac{B_k}{B^*}W(B^*) + \left(1 - \frac{B_k}{B^*}\right) W(0)$$

Since $W(0) \geq 0$, we can drop the second term. From the bracketing property, we have $\frac{B_k}{B^*} \geq \frac{1}{1+\varepsilon} \geq 1-\varepsilon$. Therefore:

$$W(B_k) \geq (1-\varepsilon)W(B^*) = (1-\varepsilon)W_{OPT}$$

□

C.5 Proof of Theorem 9

Proof of Theorem 9. The proof for DSIC was established previously and relies on the fact that the final prices are computed ex-ante and are fixed from the agents' perspective. We focus on the approximation guarantee.

Let W_{OPT} be the optimal welfare achieved by the (computationally intractable) optimal Bayesian mechanism, and let B^* be its corresponding optimal expected budget transfer. Let W_{ALG} be the expected welfare of the BEPP mechanism.

The algorithm's utility W_{ALG} is the maximum value found over the grid $\mathcal{B}$, so it is at least the value found at the specific grid point B_k that brackets the true optimum, i.e., where $B_k \leq B^* \leq (1+\varepsilon)B_k$.

$$\begin{aligned}
W_{ALG} &= \max_{B_j \in \mathcal{B}} \{W_p^{ALG}(B_j) + W_m^{ALG}(B_j)\} \\
&\geq W_p^{ALG}(B_k) + W_m^{ALG}(B_k) \\
&\geq (1-1/e) \cdot (W_p(B_k) + W_m(B_k)) && \text{(by Lemma 14)} \\
&= (1-1/e) \cdot W(B_k) \\
&\geq (1-1/e) \cdot (1-\varepsilon) \cdot W(B^*) && \text{(by Lemma 16 and concavity of } W(B)) \\
&= (1-1/e)(1-\varepsilon) \cdot W_{OPT}.
\end{aligned}$$

The polynomial runtime follows because the number of budget points to check, K, is polynomial in the input parameters and $1/\varepsilon$, and the greedy algorithm to solve the multilinear program for each subproblem is also polynomial in the number of agents and the discretization factor. This completes the proof. □

D Further Discussions on Section 5

D.1 Computing Prices

Note that the multilinear extension of (PLUS) can be solved by reducing it to the submodular maximization problem with cardinality constraint, as shown in [Balkanski and Hartline, 2016], which is known to be approximately solvable in polynomial time with multiplicative factor of $1 - 1/e$ via greedy algorithm. We will omit the detail as the approach is analogous to [Balkanski and Hartline, 2016]. To solve the multilinear extension of (MINUS), we can essentially transform the budget constraint into the one with a upper bound on the *foregone* budget instead of acquired budget. Formally, one can first compute the maximal revenue obtainable from each minus agent j by finding

$r_j = \max_{q_j} q_j(1 - F_j(q_j))$. Then, we can set $r_{\max} = \sum_{j \in \mathcal{M}} r_j$, and write the budget constraint to be

$$r_{\max} - \sum_{j \in \mathcal{M}} \left((1 - r_j)F_j^{-1}(r_j)\right) \leq R_{max} - B_k.$$

Ex-post BB and Irregular Distributions. Analogous to Appendix B, one can construct an ex-post budget-balanced mechanism with slightly weaker guarantees. Especially, we can slightly relax the budget constraint for each plus and minus agent side problem, and solve the corresponding ex-ante programming. This would imply good approximate posted pricing mechanism with high probability.

For irregular distributions, however, it is not clear how one can generalize the argument, especially since we directly use the concavity of the utility function with respect to the budget in Lemma 15. If this does not hold, there might be a case that the resulting utility not covered by our discretized grid can possibly induce a significantly large utility than maximum over our candidates. Extending the analysis to irregular distributions remain a major open problem.

That being said, we note that our simple mechanism provided shortly in Appendix E does not require any assumption on the types of distributions, providing more robust guarantees at weaker approximation factors with a simpler mechanism.

E Simple Mechanism for Submodular Utility

In this section, we show that a simple randomized mechanism can achieve a constant-factor approximation of the optimal mechanism in this case.

For any budget B, let $W_p(B)$ be the optimal utility obtainable from plus agents given an expected budget of B, and $W_n(B)$ be the optimal utility obtainable from minus agents while receiving a total payment of B. Note that W_p is a non-decreasing function, while W_n is a non-increasing function. Then, for some value B^* used by the optimal policy, $W(B^*) = W_p(B^*) + W_n(B^*)$ is the optimal obtainable utility given an ex-ante budget balanced, DSIC, and IR mechanism.

Next, let $B = \mathbb{E}\left[\sum_j \phi_j(v_j)\right]$ be the maximum expected revenue obtainable from the minus agents using a truthful auction, achievable using a posted-price auction when the distribution F_j for each v_j is regular. We consider the following mechanism. With probability $1/2$, the designer chooses no agents on either side, achieving a utility of

$$u_n(\mathcal{M}) = W_n(0) \geq W_n(B^*).$$

Otherwise, with probability $1/2$, the designer first runs a posted-price auction on the negative side to obtain a budget, with an expected value of B. Then, on the plus agent side, the designer runs a budget-feasible mechanism with budget B, achieving an approximation factor of c for some constant $0 < c \leq 1$. This achieves a utility of at least

$$cW_p(B) \geq cW_p(B^*)$$

since $B \geq B^*$ is the maximum obtainable budget and W_p is non-decreasing. Overall, this randomized mechanism achieves a utility of at least[12]

$$\frac{1}{2}(W_n(B^*) + cW_p(B^*) \geq \frac{c}{2}W(B^*).$$

F Comparative Statics

A key strength of the BEPP framework is that we can formally analyze how its final, posted-price policy changes in response to shifts in the economic environment. These comparative statics results are not just properties of a theoretical optimum, but of the actual prices computed by our polynomial-time mechanism. The foundational result that enables this analysis is the monotonicity of the subproblem pricing policies.

[12]Note that one can derandomize the mechanism by simplying taking a maximal mechanism over two.

Lemma 22 (Price Monotonicity of Subproblem Policies)**.** *In the subproblems solved by the* BEPP *mechanism for a given budget parameter B, the computed posted prices are monotone in B. Specifically, for regular distributions:*

1. *The procurement price for a plus agent, $p_i(B)$, is non-decreasing in B.*
2. *The admission price for a minus agent, $q_j(B)$, is non-increasing in B.*

Proof. The proof follows from the Karush-Kuhn-Tucker (KKT) conditions of the respective multilinear extension programs that the BEPP algorithm solves.

First, consider the plus-agent side. The optimal selection probabilities $\mathbf{q}^*$ for the plus-agent subproblem with budget B satisfy the stationarity condition $\frac{\partial U_p}{\partial q_i} = \lambda_p^* \cdot C_i'(q_i)$, where λ_p^* is the shadow price of the budget. The optimal utility function of the subproblem, $W_p(B)$, is concave in B, which implies that its derivative, the shadow price λ_p^*, is non-increasing in B. When B increases, λ_p^* weakly decreases. To restore the KKT equality, the selection probability q_i must weakly increase. Since the posted price is $p_i(B) = G_i^{-1}(q_i^*(B))$ and G_i^{-1} is monotone, an increase in q_i^* implies a non-decreasing price $p_i(B)$.

For the minus-agent side, the optimal exclusion probabilities $\mathbf{r}^*$ for the minus-agent subproblem with revenue target B satisfy a similar KKT condition on its Lagrangian. A higher revenue target B is a tighter constraint, which implies a higher shadow price on that constraint. This, in turn, forces a decrease in the exclusion probability r_j (i.e., an increase in the admission probability). Since the admission price is $q_j(B) = F_j^{-1}(r_j^*(B))$, a decrease in r_j^* implies a non-increasing admission price $q_j(B)$. □

This lemma is the tool we use to analyze the final output of the full BEPP mechanism.

The Effect of Better Agent Populations. We first investigate the mechanism's response to an improvement in the quality of an agent. We formalize this notion using the economic concept of first-order stochastic dominance (FOSD) and analyze its effect on the equilibrium budget and prices computed by the BEPP mechanism.

Theorem 23 (Monotonicity in agent quality)**.** *Consider the* BEPP *mechanism operating in a market with a profile of plus-agent cost distributions $(G_1, \dots, G_n)$. Let its final output policy be $(\mathbf{p}^*, \mathbf{q}^*)$, chosen based on an internal budget evaluation of B^*. Now, consider a new market where the cost distribution for a single plus agent $k \in \mathcal{N}$, G_k, is replaced by a new distribution G_k' that FOSD G_k. Let the new output of the* BEPP *mechanism be $(\mathbf{p'}^*, \mathbf{q'}^*)$, based on a new internal budget evaluation B'^*. Then: $B'^* \geq B^*$.*

Proof. Let $W_p^{ALG}(B)$ and $W_p'^{ALG}(B)$ be the expected utilities computed by the plus-agent sub-mechanism for a given budget B before and after the change to agent k's distribution, respectively. An FOSD improvement from G_k to G_k' means agent k is stochastically cheaper. This implies that the expected cost curve for agent k, $C_k(q_k) = q_k G_k^{-1}(q_k)$, is now lower for any given selection probability q_k.

This change enlarges the feasible set $\mathcal{Q}_B = \{\mathbf{q} \mid \sum_i C_i(q_i) \leq B\}$ for the underlying multilinear program for any given B. The mechanism is maximizing the same objective function $U_p(\mathbf{q})$ over a strictly larger set of feasible probability vectors. This can only yield a better value, so $W_p'^{ALG}(B) \geq W_p^{ALG}(B)$ for all B.

The total utility curve that BEPP computes, $W'^{ALG}(B) = W_p'^{ALG}(B) + W_m^{ALG}(B)$, is therefore pointwise higher than or equal to the original curve. The mechanism selects the budget that maximizes this curve. Since the curve has shifted up due to the increased productivity of the plus-agent side (specifically from agent k becoming more efficient), the new selected budget, B'^*, will be greater than or equal to the original selected budget, B^*. □

Corollary 24 (Two-sided spillover from agent amprovement)**.** *Under the conditions of the preceding theorem, an FOSD improvement in a single plus agent's cost distribution leads to weakly more generous offers for all plus agents and weakly more lenient admission criteria for all minus agents. Specifically, for the final posted prices:*

1. *The prices offered to all plus agents weakly increase: $\mathbf{p'}^* \geq \mathbf{p}^*$.*
2. *The admission prices for all minus agents weakly decrease: $\mathbf{q'}^* \leq \mathbf{q}^*$.*

Proof. The proof is a direct application of Lemma 22 to the result of the preceding theorem. The theorem establishes that the new equilibrium budget B'^* chosen by the BEPP mechanism is weakly greater than the original budget B^*. Applying the price monotonicity property to this higher budget choice directly implies that the final plus-agent prices must weakly increase and the final minus-agent admission prices must weakly decrease. □

This result shows that an efficiency gain from a single agent on one side of the market creates positive spillovers for all participants. When one plus agent becomes stochastically cheaper, the designer finds it optimal to increase the overall scale of the project (a higher B^*). This larger scale is funded by making it easier for minus agents to be admitted (a lower q_j^*), and the increased budget is spent on being more generous to the entire set of plus agents, not just the one who improved (a higher p_i^*).

The Effect of Individual Externalities. In addition to population-wide shifts, it is important to understand how the mechanism adapts to changes in a single agent's contribution. Here, we analyze the effect of an increase in the positive externality provided by a single plus agent, focusing on the additive utility case where this effect is clearest.

Theorem 25 (Monotonicity in individual externalities)**.** *Consider the* BEPP *mechanism for an additive utility function. Let the output policy be* $(\mathbf{p}^*, \mathbf{q}^*, B^*)$ *for a given set of externalities. If the externality for a single agent improves (i.e., for a plus agent* $k \in \mathcal{N}$*,* p_k *increases, or for a minus agent* $k \in \mathcal{M}$*,* n_k *increases), the new policy* $(\mathbf{p}'^*, \mathbf{q}'^*, B'^*)$ *computed by the mechanism will satisfy* $B'^* \geq B^*$ *and* ${W'}^*_{ALG} \geq W^*_{ALG}$.

Proof. The proof follows by analyzing the shift in the value curve $W^{ALG}(B)$ that the BEPP algorithm computes. We consider the two cases for an improved externality.

Case 1: A plus agent's externality p_k increases. In the additive case, the plus-agent sub-mechanism for a budget B solves $\max_q \sum_i p_i q_i$ subject to the budget constraint. When p_k increases, the objective function's value increases for any feasible allocation with $q_k > 0$. Therefore, the computed value $W_p'^{ALG}(B) \geq W_p^{ALG}(B)$ for all B.

Case 2: A minus agent's externality n_k increases (becomes less negative). The minus-agent sub-mechanism for a revenue target B solves $\max_q \sum_j n_j r_j$ subject to a revenue constraint. When n_k increases, the objective's value increases for any allocation where agent k is admitted. This leads to a higher computed value, $W_m'^{ALG}(B) \geq W_m^{ALG}(B)$.

In both cases, the total utility curve that BEPP evaluates, $W'^{ALG}(B)$, is pointwise greater than or equal to the original curve. The mechanism selects the budget that maximizes this curve, leading to a new selected budget $B'^* \geq B^*$ and a higher total achieved utility. □

Corollary 26 (Universal Positive Spillover)**.** *Under the conditions of the preceding theorem, an improvement in any single agent's externality (plus or minus) leads to weakly more generous offers for all plus agents and weakly more lenient admission criteria for all minus agents. Specifically:* $p_i'^* \geq p_i^*$ *for all* $i \in \mathcal{N}$ *and* $q_j'^* \leq q_j^*$ *for all* $j \in \mathcal{M}$.

Proof. The proof is a direct application of Lemma 22 to the result of the preceding theorem. The theorem establishes that an improvement in any single agent's externality, regardless of type, leads the BEPP mechanism to select an equilibrium budget transfer B'^* such that $B'^* \geq B^*$. The Price Monotonicity Lemma dictates how the computed prices for the subproblems change as a function of the budget parameter B. Since the final policy is based on a new, higher budget parameter B'^*, the plus-agent prices must weakly increase and the minus-agent admission prices must weakly decrease. □

This result demonstrates a powerful *rising tide* effect. Making any single agent more valuable to the event—either by increasing their positive contribution or decreasing their negative one—does not just benefit that agent. It incentivizes the designer to increase the overall scale and investment in the project (a higher B^*). This larger project scale, in turn, translates into more generous offers and more lenient criteria for all other participants, creating a market-wide positive spillover.

The Effect of Market Expansion. Finally, we analyze how the mechanism adapts to an increase in competition, demonstrating a powerful complementarity where the addition of a new agent to either side of the market can create positive spillovers for incumbents.

Theorem 27 (Complementarity from Market Expansion)**.** *Let* $(\mathbf{p}^*, \mathbf{q}^*, B^*)$ *be the policy computed by* BEPP *for a given market with* n *plus agents and* m *minus agents. If a new agent is introduced (either plus agent* $n+1$ *or minus agent* $m+1$*), the new policy* $(\mathbf{p'}^*, \mathbf{q'}^*, B'^*)$ *computed by* BEPP *will satisfy* $B'^* \geq B^*$*.*

Proof. The proof follows by analyzing the effect of the new agent on the total utility curve, $W^{ALG}(B)$, that the BEPP algorithm evaluates. We consider the two cases.

Case 1: A new plus agent is introduced. This expands the set of choices available to the plus-agent sub-mechanism, $M_p(B)$. For any budget B, the optimization is now over a higher-dimensional space of probabilities. Since any solution for the n-agent problem is also a feasible solution for the $(n+1)$-agent problem (by setting the new agent's probability to 0), the computed value $W_p'^{ALG}(B) \geq W_p^{ALG}(B)$.

Case 2: A new minus agent is introduced. This expands the set of choices available to the minus-agent sub-mechanism, $M_m(B)$. For any revenue target B, the designer has an additional source of revenue, which can only increase the maximal utility achievable for that target. Therefore, the new value function $W_m'(B)$ is pointwise greater than or equal to the original, $W_m'(B) \geq W_m(B)$.

In either case, the total utility curve that BEPP evaluates is raised or remains the same. The mechanism selects the budget that maximizes this curve, and thus the new selected budget B'^* will be greater than or equal to the original selected budget B^*. □

Corollary 28 (Universal Spillovers from Market Expansion)**.** *Under the conditions of the preceding theorem, the introduction of a new agent (of either type) leads to weakly more generous offers for all plus agents and weakly more lenient admission criteria for all incumbent minus agents. Specifically:*

1. *The prices offered to all plus agents weakly increase:* $p_i'^* \geq p_i^*$*.*
2. *The admission prices for all incumbent minus agents weakly decrease:* $q_j'^* \leq q_j^*$*.*

Proof. The proof is a direct application of Lemma 22 to the result of the preceding theorem. The theorem establishes that the introduction of any new agent, plus or minus, leads to a weakly higher equilibrium budget B'^* being selected by the BEPP mechanism. Applying the price monotonicity property to this new, higher budget choice directly implies that all plus-agent prices must weakly increase and all incumbent minus-agent admission prices must weakly decrease. □

This result formalizes a powerful and non-obvious insight about complementarity. In many settings, new entrants are viewed as substitutes, and their arrival would be expected to decrease the value or offers made to incumbents. Our result shows that in this curation context, the opposite is true. A new *star* agent (either a high-value plus agent or an efficient revenue-generating minus agent) makes the entire project so much more productive that the designer finds it optimal to increase the total investment and scale (a higher B^*). This increase translates into more generous offers for all other participants, who benefit from the positive spillover. The agents act as complements in driving the value and scale of the entire curated event.

F.1 Comparative Statics for the Optimal Mechanism

While the preceding theorems were stated for the output of our polynomial-time BEPP mechanism, the underlying arguments about how the equilibrium policy shifts are fundamental properties of the optimal Bayesian mechanism itself. The optimal mechanism is not a simple posted-price scheme, and its payments are complex functions of all agent reports. However, we can analyze its behavior by studying its effective *selection thresholds*.

The key to the comparative statics is a foundational monotonicity property of the optimal policy with respect to the scale of the project, which we now formalize.

Lemma 29 (Monotonicity of Optimal Selection Thresholds)**.** *Consider the optimal Bayesian mechanisms for the separated subproblems.*

- *For the plus-agent problem* $\max \mathbb{E}\left[u_p(S_p)\right]$ *s.t.* $\mathbb{E}\left[\textit{cost}\right] \leq B$*, the effective cost-acceptance threshold implemented by the optimal mechanism is non-decreasing in the budget* B*.*
- *For the minus-agent problem* $\max \mathbb{E}\left[u_m(E_m)\right]$ *s.t.* $\mathbb{E}\left[\textit{revenue}\right] \geq B$*, the effective valuation-acceptance threshold is non-increasing in the revenue target* B*.*

Proof. We provide the proof for the plus-agent subproblem. The proof for the minus-agent side is symmetric.

Let $W_p(B)$ be the optimal utility for the plus-agent problem with budget B, and let $x^*(B)$ be the optimal allocation rule that achieves it. An allocation rule x can be characterized by the set of agent types it accepts. Let us denote the set of acceptable cost-types for agent i as $\mathcal{C}_i(B) = \{c_i \mid x_i^*(B)(c_i, \cdot) > 0 \text{ for some reports from others}\}$. A more lenient policy corresponds to a larger set $\mathcal{C}_i(B)$.

Consider two budget levels $B' > B$. Since the feasible set of mechanisms for budget B' contains all mechanisms that are feasible for budget B, the optimal value must be non-decreasing: $W_p(B') \geq W_p(B)$.

Assume for contradiction that the policy for budget B' is strictly more restrictive for some agent k than the policy for budget B, meaning $\mathcal{C}_k(B') \subset \mathcal{C}_k(B)$. This implies that to achieve the optimal welfare $W_p(B')$, the mechanism relies less on agent k. To achieve a welfare of at least $W_p(B)$, the mechanism must compensate for the reduced participation of agent k by increasing the participation of other agents. This, however, would require a larger expenditure from the budget on those other agents. This creates a tension: a more restrictive threshold for one agent necessitates a more lenient (and costly) threshold for others just to achieve the same utility level, making it an inefficient way to spend an even larger budget B'. The optimal policy for the larger budget B' will instead weakly relax the acceptance thresholds for all agents to best leverage the additional resources. Therefore, the effective cost-acceptance threshold must be non-decreasing in B. □

With this foundational lemma about the optimal policy's structure, we can formally state the comparative statics results. We present the market expansion theorem as a representative example.

Theorem 30 (Complementarity from Market Expansion in the Optimal Mechanism)**.** *Let* (B^*, W^*) *be the equilibrium budget and welfare for the optimal Bayesian mechanism in a market with* n *plus agents. If a new plus agent,* $n+1$*, is introduced, the new optimal policy will have an equilibrium budget* $B'^* \geq B^*$ *and its selection thresholds for all incumbent agents will be weakly more lenient.*

Proof. Let $W(B) = W_p(B) + W_m(B)$ be the true optimal utility function for the original market. The introduction of a new plus agent creates a new plus-agent value function, $W_p'(B)$, where $W_p'(B) \geq W_p(B)$ for all B, because the designer's optimization problem for any given budget is over a larger set of choices.

This implies that the new total utility function, $W'(B) = W_p'(B) + W_m(B)$, is pointwise greater than or equal to the original. The optimal Bayesian mechanism finds the maximum of this function. Since the function's value has increased due to the higher productivity of the plus-agent side, the new equilibrium will feature weakly higher welfare, $W'^* \geq W^*$, and will be achieved at a weakly higher budget transfer, $B'^* \geq B^*$.

The final claim follows from Lemma 29. Since the equilibrium budget transfer of the optimal policy has increased from B^* to B'^*, the lemma dictates that the optimal mechanism's effective selection thresholds must become more lenient for all participants. This means higher cost-acceptance thresholds for plus agents and lower valuation-acceptance thresholds for minus agents. □

The other comparative statics results regarding improved agent populations (FOSD) and increased individual externalities follow the same formal structure. In each case, the change to the environment raises the total utility curve $W(B)$, which leads to a higher optimal budget choice B^*, which in turn implies more lenient selection thresholds via Lemma 29. This confirms that the `BEPP` mechanism's qualitative behavior correctly mirrors the fundamental properties of the true optimal policy.

G Prior-free Mechanisms

In this section, we consider the problem in a prior-free setting, where the distributions of agents' valuations and costs are not known to the designer. In this scenario, comparing performance against the second-best benchmark can be less useful, as the lack of information can severely limit the performance of any mechanism. Instead, we aim to find a mechanism that is competitive with the offline optimal utility, which is stronger than the second-best benchmark. To achieve such a guarantee, we adopt the standard large market assumption, which is commonly imposed in the literature [Anari *et al.*, 2014; Balkanski and Hartline, 2016]. Applied to our setting, this assumption means that the number of agents, both on the minus and plus side, is large enough that the effect of any singular agent on the overall outcome is negligible.

We adopt the following large market assumption:

Assumption 31 (Large Market Assumption). *Let $B = \sum_j v_j$ be the maximum attainable budget. Then, for each plus agent i and minus agent j, the following holds:*

$$c_i \leq \frac{M}{n}B, v_j \leq \frac{M}{m}B,$$

for constant $M > 0$, respectively.

It is easy to construct a counterexample that without any assumption, no mechanism can approximate the first-best outcome, by considering a single high-cost and high-utility plus agent as no mechanism could recruit this agent under incentive-compatibility.

We prove the following result under this assumption.

Theorem 32. *There exists a prior-free mechanism that provides a constant approximation to the offline optimal under the large market assumption.*

To prove this, we first show that it is possible to guarantee a constant fraction of the maximum budget obtained from the minus agents, while ignoring its effect on the utility. Next, we use the large market assumption to show that a fraction of the optimal budget is enough to obtain a constant factor of the optimal utility on the plus agent side. Finally, we randomize over two mechanisms with probabilities parameterized by the large market assumption constant, which gives us a desired result.

To prove this, we require two technical lemmas. The first is to show that it is possible to guarantee a constant fraction of the maximum budget obtained from minus agents.

Lemma 33. *There exists a truthful mechanism that achieves a revenue of at least*

$$(\frac{1}{4M} - \frac{M}{m})B$$

from the minus agents given the large market assumption with constant M and maximum possible budget B.

Proof. For any value of $1 \leq k \leq m-1$, we can consider a truthful auction that accepts the largest k bids, receiving payments equal to the $k+1$st largest bid. Now, we want to optimize k to achieve the maximum revenue. For any value of k, it follows from the large market assumption that the $k+1$-st largest bid is at least

$$\frac{\sum_j v_j - k\frac{M}{m}B}{m-k} \geq \frac{B - k\frac{M}{m}B}{m-k} = B\frac{m-Mk}{m(m-k)}.$$

Therefore, the total budget received is at least

$$\frac{k(m-Mk)}{m(m-k)}B.$$

To simplify the following calculations, we ignore the constant B and focus on the value of the coefficient $\frac{k(m-Mk)}{m(m-k)}$. Next, assume that $k = \alpha m$ for some value $0 < \alpha < 1$. Then, the expression can be simplified to

$$\alpha m \frac{m - \alpha m M}{m(m - \alpha m)} = \alpha \frac{1 - \alpha M}{1 - \alpha}.$$

Next, we find the value of α maximizing this expression, and choose $k = \lfloor \alpha m \rfloor$. Note that given the large market assumption, the budget attained by choosing the top k bids is at most $\frac{M}{m}B$ less than the value attained by maximizing α. This term is maximized when when $\alpha = 1 - \frac{\sqrt{M^2-M}}{M}$, resulting in a value of

$$(2M - 2\sqrt{M(M-1)} - 1).$$

Therefore, choosing $k = \lfloor 1 - \frac{\sqrt{M^2-M}}{M} \rfloor$ can achieve a budget of at least

$$(2M - 2\sqrt{M(M-1)} - 1 - \frac{M}{m})B \geq (\frac{1}{4M} - \frac{M}{m})B.$$

□

Next, we use the large market assumption to show that a fraction of the optimal budget is enough to obtain a constant factor of the optimal utility on the plus agent side.

Lemma 34. *For any budget B', let $W_p(B')$ be the maximum utility that can be obtained on the plus agent side given budget B'. Then, for any constant $\delta > 0$,*

$$W_p(\delta B) \geq (\delta - \frac{M}{n})W_p(B).$$

Proof. Consider the selection of agents achieving the maximum utility given budget B. Then, greedily choose agents based on the ratio of their utility to their cost, stopping when adding an agent would require a budget greater than δB. By the large market assumption, each agent has cost at most $\frac{M}{n}B$, so at least $(\delta - \frac{M}{n})B$ of the budget will be used. Furthermore, since we have chosen the agents with the highest ratio of marginal utility to cost, the utility of the selected set will be at least a $(\delta - \frac{M}{n})$ fraction of the utility achievable given budget B. □

Finally, we can combine the budget obtained from the minus agents with a budget-feasible mechanism for the plus agents to obtain the following.

Lemma 35. *There exists a truthful mechanism achieving a utility of at least*

$$(\frac{1}{4M} - \frac{M}{m} - \frac{M}{n})(1 - \frac{1}{e})W_p(B),$$

where the $W_p(B)$ denotes the maximum attainable utility on the plus agent side using total budget B.

Proof. By Lemma 33, we can obtain a budget of $\frac{1}{4M} - \frac{M}{m}$ from the minus agents. Then, we can use budget-feasible mechanism introduced [Anari *et al.*, 2014] for large markets to achive a utility of at least

$$(1 - \frac{1}{e})W_p((\frac{1}{4M} - \frac{M}{m})B).$$

Finally, by Lemma 34, we have

$$W_p((\frac{1}{4M} - \frac{M}{m})B) \geq (\frac{1}{4M} - \frac{M}{m} - \frac{M}{n})W_p(B).$$

Noting that the utility from the minus agents is non-negative, this completes the proof. □

Proof of Theorem 32. Finally, to complete our proof of Theorem 32, we note that the utility achieved by the offline optimal can be written as

$$W_p(B^*) + W_m(B^*)$$

for some budget B^*, where given the budget, W_p and W_m denote the maximum achievable utility on the plus and minus side, respectively. We randomize between the mechanism obtained in Lemma 35 and a mechanism choosing no agents, choosing the first with probability $\frac{1}{1+(1-1/e)(1/4M-M/n-M/m)}$. This achieves a utility of

$$\frac{(1-1/e)(1/4M-M/n-M/m)}{1+(1-1/e)(1/4M-M/n-M/m)}(W_p(B)+W_m(0).$$

Given $W_p(B) \geq W_p(B^*)$ and $W_m(0) \geq W_m(B^*)$, this implies this mechanism achieves a

$$\frac{(1-1/e)(1/4M-M/n-M/m)}{1+(1-1/e)(1/4M-M/n-M/m)}$$

approximation of the offline optimal. □

G.1 Monte Carlo estimation of the ex-ante budget surplus

Fix a shadow price $\lambda \geq 0$ and let $\mathcal{M}_\lambda$ denote the DSIC mechanism whose allocation rule $\mathbf{x}^\lambda(\mathbf{c}, \mathbf{v})$ is computed by greedy maximization of virtual weights under the (separable) matroid feasibility constraint, and whose payments are the corresponding DSIC threshold payments.

Define the realized *budget surplus* under $\mathcal{M}_\lambda$ as the random variable

$$Z(\lambda; \mathbf{c}, \mathbf{v}) \;:=\; \sum_{j \in \mathcal{M}} t_j^\lambda(\mathbf{c}, \mathbf{v}) \;-\; \sum_{i \in \mathcal{N}} t_i^\lambda(\mathbf{c}, \mathbf{v}), \tag{G.1}$$

and the ex-ante surplus as

$$B(\lambda) \;:=\; \mathbb{E}_{\mathbf{c} \sim G, \mathbf{v} \sim F}\big[Z(\lambda; \mathbf{c}, \mathbf{v})\big]. \tag{G.2}$$

Boundedness.. Since costs and values are supported on $[c_{\min}, c_{\max}]$ and $[v_{\min}, v_{\max}]$, and $\mathcal{M}_\lambda$ is DSIC and IR, each realized payment satisfies $0 \leq t_i^\lambda(\mathbf{c}, \mathbf{v}) \leq c_{\max}$ for $i \in \mathcal{N}$ and $0 \leq t_j^\lambda(\mathbf{c}, \mathbf{v}) \leq v_{\max}$ for $j \in \mathcal{M}$. Hence,

$$Z(\lambda; \mathbf{c}, \mathbf{v}) \in [-H, H] \quad \text{for} \quad H := n \cdot c_{\max} + m \cdot v_{\max}. \tag{G.3}$$

Lemma 36 (Monte Carlo estimation)**.** *Fix $\lambda \geq 0$. Let $\{(\mathbf{c}^{(s)}, \mathbf{v}^{(s)})\}_{s=1}^T$ be i.i.d. samples from $G \times F$, and define the Monte Carlo estimator*

$$\widehat{B}_T(\lambda) \;:=\; \frac{1}{T} \sum_{s=1}^{T} Z\Big(\lambda; \mathbf{c}^{(s)}, \mathbf{v}^{(s)}\Big). \tag{G.4}$$

Then for any $\eta > 0$,

$$\Pr\left[\,\right]\left[\left|\widehat{B}_T(\lambda) - B(\lambda)\right| \geq \eta\right] \leq 2\exp\left(-\frac{T\eta^2}{2H^2}\right), \tag{G.5}$$

where H is as in (G.3). In particular, if

$$T \;\geq\; \frac{2H^2}{\eta^2} \ln \frac{2}{\delta}, \tag{G.6}$$

then $\left|\widehat{B}_T(\lambda) - B(\lambda)\right| \leq \eta$ holds with probability at least $1 - \delta$.

Proof. Fix λ. By (G.3), each sample $Z(\lambda; \mathbf{c}^{(s)}, \mathbf{v}^{(s)})$ lies in an interval of length $2H$. Since the samples are i.i.d., Hoeffding's inequality gives

$$\Pr\left[\,\right]\left[\left|\widehat{B}_T(\lambda) - \mathbb{E}[Z(\lambda; \mathbf{c}, \mathbf{v})]\right| \geq \eta\right] \leq 2\exp\left(-\frac{2T\eta^2}{(2H)^2}\right) = 2\exp\left(-\frac{T\eta^2}{2H^2}\right),$$

which is the claimed bound. Solving $2\exp(-T\eta^2/(2H^2)) \leq \delta$ yields the sample complexity. □

Monotonicity in λ.. Let $U(\mathbf{x})$ denote the designer's expected *additive* utility under allocation rule $\mathbf{x}$, and let $B(\mathbf{x})$ denote the ex-ante budget surplus induced by $\mathbf{x}$ (i.e., expected revenue minus expected cost). Consider the Lagrangian

$$\mathcal{L}(\mathbf{x}, \lambda) \;:=\; U(\mathbf{x}) + \lambda \cdot B(\mathbf{x}), \tag{G.7}$$

Algorithm 4 Sampled bisection for an approximately budget-balanced shadow price

Input: accuracy $\eta > 0$, failure probability $\delta \in (0, 1)$, bracket $[0, \Lambda]$

Set $L \leftarrow \lceil \log_2(\Lambda/\eta) \rceil$.
Set per-iteration confidence $\delta' \leftarrow \delta/L$.
Set sample size $T \leftarrow \left\lceil \frac{2H^2}{(\eta/4)^2} \ln \frac{2}{\delta'} \right\rceil$ (using H from (G.3)).
Initialize $\lambda_{\mathrm{lo}} \leftarrow 0$, $\lambda_{\mathrm{hi}} \leftarrow \Lambda$.
for $\ell = 1, 2, \dots, L$ **do**
  $\lambda \leftarrow (\lambda_{\mathrm{lo}} + \lambda_{\mathrm{hi}})/2$.
  Draw T i.i.d. samples $(\mathbf{c}^{(s)}, \mathbf{v}^{(s)}) \sim G \times F$.
  Compute $\widehat{B}_T(\lambda) \leftarrow \frac{1}{T} \sum_{s=1}^{T} Z(\lambda; \mathbf{c}^{(s)}, \mathbf{v}^{(s)})$.
  if $\widehat{B}_T(\lambda) \geq \eta/2$ **then**
    $\lambda_{\mathrm{hi}} \leftarrow \lambda$.
  else if $\widehat{B}_T(\lambda) \leq -\eta/2$ **then**
    $\lambda_{\mathrm{lo}} \leftarrow \lambda$.
  else
    break
  end if
end for
return $\widehat{\lambda} \leftarrow \lambda_{\mathrm{hi}}$.

with $\lambda \geq 0$. For each λ, let $\mathbf{x}^\lambda$ be any maximizer of $\mathcal{L}(\cdot, \lambda)$ over the feasible allocation rules (including the matroid feasibility and implementability/monotonicity constraints).

Lemma 37 (Surplus monotonicity)**.** *For any $\lambda_1 < \lambda_2$, we have $B(\mathbf{x}^{\lambda_1}) \leq B(\mathbf{x}^{\lambda_2})$. Equivalently, the function $B(\lambda) := B(\mathbf{x}^\lambda)$ is (weakly) nondecreasing in λ.*

Proof. Let $\mathbf{x}^1 := \mathbf{x}^{\lambda_1}$ and $\mathbf{x}^2 := \mathbf{x}^{\lambda_2}$. By optimality of $\mathbf{x}^1$ at λ_1,

$$U(\mathbf{x}^1) + \lambda_1 B(\mathbf{x}^1) \;\geq\; U(\mathbf{x}^2) + \lambda_1 B(\mathbf{x}^2). \tag{G.8}$$

By optimality of $\mathbf{x}^2$ at λ_2,

$$U(\mathbf{x}^2) + \lambda_2 B(\mathbf{x}^2) \;\geq\; U(\mathbf{x}^1) + \lambda_2 B(\mathbf{x}^1). \tag{G.9}$$

Adding the two inequalities cancels the $U(\cdot)$ terms and yields

$$(\lambda_2 - \lambda_1)\, B(\mathbf{x}^2) \;\geq\; (\lambda_2 - \lambda_1)\, B(\mathbf{x}^1).$$

Since $\lambda_2 - \lambda_1 > 0$, we conclude $B(\mathbf{x}^2) \geq B(\mathbf{x}^1)$. □

Theorem 38 (Approximate ex-ante budget balance via sampling)**.** *Assume $B(\lambda)$ is nondecreasing (Lemma 37) and that $B(0) \leq 0 \leq B(\Lambda)$ for the bracket $[0, \Lambda]$. Then Algorithm 4 returns $\widehat{\lambda}$ such that, with probability at least $1 - \delta$,*

$$B(\widehat{\lambda}) \;\geq\; -\eta. \tag{G.10}$$

*Equivalently, the resulting mechanism $\mathcal{M}_{\widehat{\lambda}}$ is η-*approximately ex-ante budget-balanced*:*

$$\mathbb{E}\Big[\sum_{j \in \mathcal{M}} t_j^{\widehat{\lambda}}\Big] \;\geq\; \mathbb{E}\Big[\sum_{i \in \mathcal{N}} t_i^{\widehat{\lambda}}\Big] - \eta.$$

Moreover, the total number of samples used is $T \cdot O(\log(\Lambda/\eta))$.

Proof. Let $\mathcal{E}$ be the event that for every λ queried by the algorithm, $\left|\widehat{B}_T(\lambda) - B(\lambda)\right| \leq \eta/4$. By Lemma 36 and our choice of T with confidence $\delta' = \delta/L$, each query violates this bound with probability at most δ'. There are at most L queries, so by a union bound, $\Pr\left[[\,]\mathcal{E}\right] \geq 1 - \delta$.

Condition on $\mathcal{E}$. If the algorithm terminates in the "middle" case $|\widehat{B}_T(\lambda)| < \eta/2$, then $|B(\lambda)| \leq |\widehat{B}_T(\lambda)| + \eta/4 < 3\eta/4$, so in particular $B(\lambda) \geq -\eta$.

Otherwise, the algorithm maintains the invariant $B(\lambda_{\mathrm{lo}}) \leq -\eta/4$ and $B(\lambda_{\mathrm{hi}}) \geq \eta/4$ (because the update decisions are correct under $\mathcal{E}$ and $B(\cdot)$ is nondecreasing). At the end, $\widehat{\lambda} = \lambda_{\mathrm{hi}}$ satisfies $B(\widehat{\lambda}) \geq \eta/4 \geq -\eta$. This proves the claimed approximate ex-ante budget balance. □

Lemma 39 (High-probability ex-ante BB implies unconditional approximate ex-ante BB)**.** *Let $\widehat{\lambda}$ be a random output of a sampling-based procedure, and let*

$$B(\lambda) := \mathbb{E}\Big[\sum_{j \in \mathcal{M}} t_j^{\lambda} - \sum_{i \in \mathcal{N}} t_i^{\lambda}\Big]$$

denote the ex-ante budget surplus of the induced mechanism $\mathcal{M}_\lambda$ (expectation over types and the mechanism's internal randomness after *λ is fixed). Assume (i) $|B(\lambda)| \leq H$ for all λ, and (ii)*

$$\Pr\left[\, B(\widehat{\lambda}) \geq -\eta \,\right] \geq 1 - \delta.$$

Then the randomized *mechanism that first samples $\widehat{\lambda}$ and then runs $\mathcal{M}_{\widehat{\lambda}}$ satisfies*

$$\mathbb{E}\big[B(\widehat{\lambda})\big] \geq -(\eta + \delta H),$$

where the outer expectation is over the procedure's randomness that generates $\widehat{\lambda}$.

Proof. Let $\mathcal{E}$ be the event $B(\widehat{\lambda}) \geq -\eta$. Then

$$\mathbb{E}[B(\widehat{\lambda})] = \mathbb{E}[B(\widehat{\lambda}) \mid \mathcal{E}]\Pr\left[[\,]\mathcal{E}\right] + \mathbb{E}[B(\widehat{\lambda}) \mid \neg\mathcal{E}]\Pr\left[[\,]\neg\mathcal{E}\right].$$

On $\mathcal{E}$ we have $B(\widehat{\lambda}) \geq -\eta$, hence $\mathbb{E}[B(\widehat{\lambda}) \mid \mathcal{E}] \geq -\eta$. On $\neg\mathcal{E}$ we use the uniform bound $B(\widehat{\lambda}) \geq -H$. Therefore

$$\mathbb{E}[B(\widehat{\lambda})] \geq (-\eta)(1-\delta) + (-H)\delta \geq -\eta - \delta H.$$

□

G.2 BEPP with separable matroid feasibility

Let $\mathcal{M} = (\mathcal{N}, \mathcal{I}_p)$ be a matroid over plus agents and $\mathcal{M}_m = (\mathcal{M}, \mathcal{I}_m)$ a matroid over minus agents. A feasible outcome selects $S_p \in \mathcal{I}_p$ plus agents and admits $S_m \in \mathcal{I}_m$ minus agents. We assume the utility is separable:

$$u(S_p, S_m) = u_p(S_p) + u_m(\mathcal{M} \setminus S_m),$$

where u_p is nonnegative monotone submodular on $\mathcal{N}$ and u_m is nonnegative monotone submodular on excluded minus agents.

For a matroid $\mathcal{M} = (E, \mathcal{I})$, let $P(\mathcal{M}) \subseteq [0, 1]^E$ denote its matroid polytope.

G.3 Estimating multilinear extensions

Let $f : 2^E \to \mathbb{R}_{\geq 0}$ be a value-oracle monotone submodular function on ground set E. For $\mathbf{x} \in [0, 1]^E$, let $R(\mathbf{x}) \subseteq E$ be a random set where each $e \in E$ is included independently with probability x_e. The multilinear extension is

$$F(\mathbf{x}) := \mathbb{E}\big[f(R(\mathbf{x}))\big].$$

Unbiased estimators.. Given $\mathbf{x}$, an unbiased estimator of $F(\mathbf{x})$ is obtained by sampling $R(\mathbf{x})$ and querying $f(R(\mathbf{x}))$. Moreover, the partial derivative has the representation

$$\frac{\partial F(\mathbf{x})}{\partial x_e} = \mathbb{E}\big[f(R(\mathbf{x}) \cup \{e\}) - f(R(\mathbf{x}) \setminus \{e\})\big] = \mathbb{E}_{R \sim \mathbf{x}_{-e}}\big[f(R \cup \{e\}) - f(R)\big],$$

where $R \sim \mathbf{x}_{-e}$ means sampling each $e' \neq e$ independently with probability $x_{e'}$.

Lemma 40 (Monte Carlo estimation of $F(\mathbf{x})$ and $\nabla F(\mathbf{x})$)**.** *Assume* $0 \leq f(S) \leq f(E)$ *for all* $S \subseteq E$*. Fix* $\mathbf{x} \in [0, 1]^E$ *and accuracy* $\gamma > 0$*, failure probability* $\delta \in (0, 1)$*. There is a Monte Carlo procedure that, using*

$$T = O\left(\frac{f(E)^2}{\gamma^2} \log \frac{|E|}{\delta}\right)$$

independent samples and value-oracle calls, returns estimates $\widehat{F}(\mathbf{x})$ *and* $\widehat{\nabla F}(\mathbf{x})$ *such that with probability at least* $1 - \delta$*,*

$$\big|\widehat{F}(\mathbf{x}) - F(\mathbf{x})\big| \leq \gamma, \qquad \|\widehat{\nabla F}(\mathbf{x}) - \nabla F(\mathbf{x})\|_\infty \leq \gamma.$$

Proof. Each sample $f(R(\mathbf{x}))$ lies in $[0, f(E)]$, so Hoeffding's inequality gives concentration for $\widehat{F}(\mathbf{x})$. For each coordinate e, the random variable $f(R \cup \{e\}) - f(R)$ lies in $[0, f(E)]$ by monotonicity and nonnegativity, so the same concentration bound applies. A union bound over $|E|$ coordinates yields the stated sample complexity. ☐

G.4 Optimizing multilinear extensions over a matroid polytope

Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid and let $P(\mathcal{M}) \subseteq [0, 1]^E$ be its matroid polytope. We consider the relaxation

$$\max_{\mathbf{x} \in P(\mathcal{M})} \quad F(\mathbf{x}), \tag{G.11}$$

where F is the multilinear extension of a monotone submodular f.

Linear optimization oracle.. Linear optimization over $P(\mathcal{M})$ is solvable by the greedy algorithm: for any weight vector $\mathbf{w} \in \mathbb{R}^E$, one can compute $\arg\max_{\mathbf{x} \in P(\mathcal{M})} \langle \mathbf{w}, \mathbf{x} \rangle$ by sorting E in decreasing w_e and greedily adding elements while maintaining independence.

Algorithm 5 Stochastic Continuous Greedy (sketch)

Input: value oracle for f, matroid $\mathcal{M} = (E, \mathcal{I})$, step size $\Delta \in (0, 1)$, accuracy γ, failure δ

```
Initialize x^(0) ← 0, K ← ⌈1/Δ⌉.
for k = 0, 1, ..., K − 1 do
    Estimate ĝ^(k) ≈ ∇F(x^(k)) using Lemma 40 with per-iteration failure δ/K.
    Compute y^(k) ∈ arg max_{y∈P(M)} ⟨ĝ^(k), y⟩ via matroid greedy.
    Update x^(k+1) ← x^(k) + Δy^(k).
end for
return x^(K).
```

Theorem 41 (Standard guarantee)**.** *There exists a choice of parameters* (Δ, γ, T) *polynomial in* $(|E|, 1/\varepsilon, \log(1/\delta))$ *such that Algorithm 5 returns* $\widehat{\mathbf{x}} \in P(\mathcal{M})$ *satisfying*

$$F(\widehat{\mathbf{x}}) \geq \left(1 - \frac{1}{e} - \varepsilon\right) \cdot \max_{\mathbf{x} \in P(\mathcal{M})} F(\mathbf{x})$$

with probability at least $1 - \delta$.

G.5 BEPP under separable matroid feasibility

Let $\mathcal{M}_p = (\mathcal{N}, \mathcal{I}_p)$ and $\mathcal{M}_m = (\mathcal{M}, \mathcal{I}_m)$ be matroids. A feasible outcome selects (S_p, S_m) with $S_p \in \mathcal{I}_p$ and $S_m \in \mathcal{I}_m$. Assume separable utility

$$u(S_p, S_m) = u_p(S_p) + u_m(S_m),$$

where u_p and u_m are nonnegative monotone submodular functions.

For a target budget B, define the *plus* and *minus* multilinear relaxations:

$$\text{(PLUS-M)} \quad \max_{\mathbf{q} \in P(\mathcal{M}_p)} \quad U_p(\mathbf{q}) \qquad \text{s.t.} \quad \sum_{i \in \mathcal{N}} q_i G_i^{-1}(q_i) \leq B, \tag{G.12}$$

$$\text{(MINUS-M)} \quad \max_{\mathbf{r} \in P(\mathcal{M}_m)} \quad U_m(\mathbf{r}) \qquad \text{s.t.} \quad \sum_{j \in \mathcal{M}} (1 - r_j) F_j^{-1}(r_j) \geq B. \tag{G.13}$$

Let $\mathbf{q}(B)$ and $\mathbf{r}(B)$ be (approximate) solutions returned by stochastic continuous greedy (and the standard extension to include the one budget constraint).

Implementation via posted pricing + matroid filtering.. Set posted prices $\tau_i := G_i^{-1}(q_i(B))$ for plus agents and $\pi_j := F_j^{-1}(r_j(B))$ for minus agents. Process agents in a fixed order; when an agent accepts its posted price, add it to the chosen set if it maintains matroid independence, otherwise skip it. (This is a standard contention-resolution implementation for matroids.)

Theorem 42 (Constant-factor approximation under separable matroids)**.** *For any* $\varepsilon > 0$*, the above BEPP-with-matroids mechanism is DSIC, IR, and ex-ante BB. Moreover, it achieves a constant-factor approximation (up to* $O(\varepsilon)$ *discretization loss) to the optimal Bayesian mechanism under separable submodular utility and separable matroid feasibility.*

Proof sketch. Fix a budget B.

Step 1 (Upper bound via concave closure). As in Lemma 13 of the main text, any DSIC, IR mechanism induces marginal vectors $\mathbf{q} \in P(\mathcal{M}_p)$ and $\mathbf{r} \in P(\mathcal{M}_m)$, and its expected utilities are upper bounded by the corresponding concave-closure programs with the same constraints. (We omit the identical envelope and relaxation steps.)

Step 2 (From concave closure to multilinear optimum). By the correlation-gap bound for monotone submodular functions (Theorem 12 in the main text), for any marginals, the multilinear extension value is at least a $(1 - 1/e)$ fraction of the concave closure. Therefore the optima of (G.12) and (G.13) are constant-factor approximations to the corresponding concave-closure upper bounds.

Step 3 (From fractional marginals to a feasible posted-price outcome). The posted prices realize the desired acceptance probabilities. The matroid feasibility filter is a contention resolution step that outputs a feasible set. Standard contention-resolution analysis for matroids implies that the expected submodular value of the feasible outcome is within a constant factor of the multilinear-extension value at $(\mathbf{q}(B), \mathbf{r}(B))$.

Step 4 (Budget enumeration). As in Lemmas 15–16 in the main text, discretizing the budget space introduces only $O(\varepsilon)$ loss, and selecting the best grid budget yields the claimed approximation.

DSIC and IR follow because prices are posted ex ante and acceptance is a dominant strategy. Ex-ante BB holds by construction since the minus side is parameterized to raise at least B in expectation and the plus side to spend at most B in expectation. □