On the Structure of Informationally Robust Optimal Mechanisms*

Benjamin Brooks Songzi Du

May 21, 2024

Abstract

We study the design of optimal mechanisms when the designer is uncertain both about the form of information held by the agents and also about which equilibrium will be played. The *guarantee* of a mechanism is its worst performance across all information structures and equilibria. The *potential* of an information structure is its best performance across all mechanisms and equilibria. We formulate a pair of linear programs, one of which is a lower bound on the maximum guarantee across all mechanisms, and the other of which is an upper bound the minimum potential across all information structures. In applications to public expenditure, bilateral trade, and optimal auctions, we use the bounding programs to characterize guarantee-maximizing mechanisms and potential-minimizing information structures and show that the max guarantee is equal to the min potential.

Keywords: Mechanism design, information design, public expenditure, optimal auctions, max-min, Bayes correlated equilibrium, robustness.

JEL Classification: C72, D44, D82, D83.

*Brooks: Department of Economics, University of Chicago, babrooks@uchicago.edu; Du: Department of Economics, University of California San Diego, sodu@ucsd.edu. This research has been supported by NSF #2215259 and #2215475. It has benefitted from discussions with Dirk Bergemann, Gabriel Carroll, Stephen Morris, Roger Myerson, Andy Postlewaite, Doron Ravid, and Phil Reny, and numerous seminar and conference audiences. We are also very grateful for helpful feedback from the editor and four anonymous referees. Karen Wu and Joey Feffer provided valuable research assistance.

1 Introduction

1.1 Motivation

In the standard model of Bayesian mechanism design, the designer is assumed to know the precise form of the agents' private information about payoff relevant states of the world, specified as an information structure. As is well known, the predictions of the model may depend on fine details of the information structure.[1] There are settings in which the information structure corresponds to objects in the world that an analyst could conceivably observe and measure. But more often, the information structure is an abstract "as-if" representation of agents' thought processes and preferences. This representation is conceptually appealing and also disciplines our modeling of behavior under incomplete information. But the abstract and artificial nature of the information structure is problematic, insofar as it is not something that we should expect a real-world mechanism designer to know with any confidence, and the dependence of the theory on the particulars of the information structure limit its practical usefulness.

A distinct issue is that many theories of Bayesian mechanism design assume that the designer can coordinate the agents on the designer's preferred equilibrium. In some cases, the mechanisms suggested by the theory have equilibria that are both normatively desirable and compelling as a positive prediction. But in other cases, the theory leads to mechanisms that are vulnerable to the agents coordinating on equilibria that are bad for the designer (such as low revenue "bidding ring" equilibria of second-price auctions). Existing methodologies for ruling out such undesirable equilibria often involve theoretically valid but impracticable constructions, such as integer games.

In this paper, we propose a new framework for informationally-robust mechanism design that does not depend on the precise structure of agents' private information or on which equilibrium will be played. The designer only specifies the distribution of the underlying payoff relevant states. The *guarantee* of a mechanism is its lowest performance across all equilibria and all common prior information structures for which the marginal on states matches the designer's prior. We characterize mechanisms that maximize the guarantee. Such mechanisms provide the best possible lower bound on performance, given these minimal assumptions about information and behavior.

In parallel, we also study information structures that are especially challenging for the designer: The *potential* of an information structure is maximum performance across all mechanisms and equilibria. We characterize information structures that minimize the potential. Such information structures can be used to certify that a mechanism maximizes the guarantee. In particular, given a mechanism and information structure, if the associated guarantee and potential are equal, then they are also equal to the max guarantee and the min potential. A further reason for analyzing potential-minimizing information structures is that they represent the environments that guarantee-maximizing mechanisms are guarding against. The plausibility of these environments is important for assessing the value of this particular kind of robustness.

[1]For example, whether or not there is correlation in agents' signals can dramatically affect what outcomes can be implemented (Myerson, 1981; Crémer and McLean, 1988; McAfee and Reny, 1992).

1.2 Results

Our first main result, Theorem 1, describes a pair of *bounding linear programs*, one of which lower bounds the max guarantee, and the other of which upper bounds the min potential. The programs are parameterized by a finite number of actions in the mechanism or signals in the information structure. The bounds are obtained from the max guarantee and min potential programs by fixing an arbitrary order on actions or signals, dropping equilibrium constraints that are non-local with respect to that order, and choosing units for actions or signals so that the Lagrange multiplier on local constraints is normalized to one. The interest in these programs stems primarily from the fact that for a number of applications, the bounds turn out to be tight, in the sense that difference between the optimal value of the bounding programs goes to zero as the number of actions and signals grows large. *A fortiori*, for these applications, max guarantee is equal to min potential. Moreover, whenever the bounds are tight, the solutions to the bounding programs are approximate guarantee maximizers and approximate potential minimizers.

The structure of the bounding programs also sheds light on the essential properties of guarantee-maximizing mechanisms and potential-minimizing information structures. Given an information structure, we define a new object associated with each signal profile and outcome, which we term the *informational virtual objective*. This is the designer's objective less information rents accruing to the agents' from the ability to mimic nearby (lower) types. In the special case of revenue maximization from the sale of private goods, the informational virtual objective coincides with the "virtual value" familiar from the theory of optimal auctions (Myerson, 1981). The upper bounding program is simply choosing the information structure to minimize the expectation (across signals) of the highest (across outcomes) informational virtual objective.

Analogously, given a mechanism, an action profile, and a payoff relevant state, the *strategic virtual objective* is the designer's objective plus a strategic adjustment term coming from the agents' ability to deviate to nearby (higher) actions. The lower bounding program is simply choosing the mechanism to maximize the expectation (across states) of the lowest (across action profiles) strategic virtual objective.

The upshot is that when the bounding programs are tight, what makes a mechanism robust in terms of the guarantee is that it achieves a favorable expected lowest strategic virtual objective, and what makes an information structure unfavorable in terms of the potential is that it depresses the expected highest informational virtual objective.

An issue that is of fundamental importance to this theory is whether or not the bounds are tight. We show that the bounding programs are "almost" a dual pair, in the sense that the dual of the upper bound program has the same form as the lower bound program, but with a subtly modified virtual objective. The key differences are that in the lower bound, the relevant equilibrium constraints point away from the action with the relevant participation constraint, whereas in the upper bound, the relevant equilibrium constraints point towards the type with the relevant participation constraint. Also, there is an important difference in how we model participation: in the lower bound, we impose a novel condition on mechanisms that we call *participation security*—each agent must have an action that guarantees them a payoff greater than their outside option, analogous to bidding zero in an auction—whereas for the upper bound, we impose the usual constraint that interim utility is greater than the outside option. Thus, whether or not the bounds are tight is related to whether the solutions are sufficiently smooth—so that the direction of local equilibrium constraints is immaterial—and whether or not the two forms of participation constraint are equivalent.

We apply our framework to two classic problems in mechanism design: public goods provision and optimal auctions. The application to public goods is fully developed in Brooks and Du (2023); in the present paper we simply outline the results, as they relate to the bounding programs. In the public goods problem, the mechanism determines expenditure on a public good, subject to budget balance and participation constraints. The designer's goal is to maximize social surplus. (The two agent version of this model can be reinterpreted as a model of bilateral trade, thus showing that our methodology can be fruitfully applied to that problem as well.) We use the bounding programs to construct a saddle point consisting of a guarantee-maximizing mechanism and a potential-minimizing information structure.

We then turn attention to revenue maximization in multi-good auctions. Our main result for that section, Theorem 2, shows that for this class of problems, the bounds are always tight. Theorem 2 also reveals additional structure of the bounding programs, in particular why it is suboptimal for the designer to use mechanisms that are discontinuous in the limit infinitely many actions. Theorem 2 is a kind of strong duality theorem: Even though the argument is non-constructive, it gives us assurance that the bounds are tight, thereby motivating us to solve for solutions to the bounding programs. We also apply the bounding programs to characterize revenue guarantee-maximizing mechanisms for a new class of environments, where the designer knows the empirical distribution of agents' values, but does not know which agent has which value. Collectively, these applications and our prior work demonstrate the utility of the bounding programs for solving informationally-robust mechanism design problems.

1.3 Related literature

This paper builds on prior work on revenue maximization from the sale a single good (Bergemann, Brooks, and Morris, 2016; Du, 2018; Brooks and Du, 2021a,b). These papers used similar bounding methodology to solve for guarantee-maximizing mechanisms and potential-minimizing information structures. The contribution of this paper is to explicitly describe and to generalize the bounding methodology, to show that it can be fruitfully applied in diverse applications, and to prove non-constructively that the bounds are tight for optimal auction design with multiple goods and interdependent values.[2]

We also contribute to the large literature on robust mechanism design. Much of this literature has attempted to relax the assumption of common knowledge of the information

[2]He and Li (2022) also study robust revenue maximization in private value auctions, but look for robustness with respect to the correlation between agents' values rather than information per se. In contrast, the application of our framework to auctions allows for values to be interdependent, and we hold the joint distribution of values fixed throughout (although it is straightforward to extend our theory to one where the guarantee is over a set of value distributions).

structure on the part of the agents by adopting stronger implementation concepts, most prominently ex post implementation. When restricting to direct mechanisms, this requires truthful reporting to be optimal regardless of agents' beliefs.[3] Ex post implementation does not address our primary concern, which is that the standard model requires the designer to have an implausibly detailed description of the informational environment. In particular, to provide ex post incentives, the designer still has to know the possible signals of the agents and how they are related to the agents' and designer's preferences.[4] Moreover, ex post implementation entails the restrictive assumption that the outcome of the mechanism cannot vary with agents' beliefs, even when such variation might be acceptable or even desirable.[5]

Our work is also related to the literature on full implementation, pioneered by Maskin (1999). Serrano and Vohra (2010) provide necessary and sufficient conditions for full implementation in mixed-strategy Bayes Nash equilibrium, meaning that there exists a mechanism for which a given social choice set is precisely the set of social choice functions that are induced in some equilibrium.[6] This approach may be contrasted with partial implementation, described in our opening paragraphs, where the designer only asks for a given social choice function to be implemented in one equilibrium. The mechanisms that we construct are, of course, implementing *some* social choice set, which must satisfy the conditions of Serrano and Vohra (2010). Moreover, the mechanism's guarantee is the minimum expected payoff over all elements of the social choice set and common priors with the correct marginal over states.

Our approach represents a middle ground between the literatures on full implementation and partial implementation. We share the concern that is central in the full implementation literature about how performance varies across environments and equilibria. But like the

[3]Ex post implementation is equivalent to dominant strategy implementation when values are private (Dasgupta et al., 1979; Chung and Ely, 2007; Yamashita, 2016; Chen and Li, 2018; Che, 2020; Bachrach et al., 2022).

[4]For example, in their classic paper on ex post implementation, Bergemann and Morris (2005) restrict attention to a class of information structures parameterized by a collection of "payoff types." They assume that agents know their own payoff types, which collectively capture everything about the environment that is payoff relevant to the agents and the designer.

[5]Chung and Ely (2007) and Bergemann and Morris (2005) give conditions under which a designer would not benefit from having implemented outcomes depend on agents' beliefs. Our view is that these conditions are quite demanding, and they suggest that the range of applications to which ex post implementation can be fruitfully applied may be quite limited. In the auction context, Chung and Ely (2007) require a generalized form of regularity á la Myerson (1981). Yamashita and Zhu (2018) and Chen and Li (2018) provide analogous conditions for more general environments. Bergemann and Morris's (2005) result relies on a "separability" condition: the designer is flexible only with respect to agent-specific dimensions of the outcome (e.g., transfers), there are no joint feasibility restrictions across agents, and each agent cares only about their own dimension. Relatedly, Jehiel et al. (2006) show that in generic environments with multidimensional types, only constant mechanisms are ex post implementable.

[6]A superficial difference between our model and the Bayesian full implementation literature is that that the latter typically considers a single information structure, with a fixed set of signals. Only interim beliefs are specified, and the prior over the whole information structure plays no role in the question of which social choice sets can be fully implemented. Setting significant technical details aside, we can view this single "grand" information structure as the disjoint union of the information structures considered in our model.

partial implementation literature, we derive the implemented social choice set from primitive preferences of the designer, under the retained assumption of a common prior that is shared between the designer and the agents. Also, the guarantee-maximizing mechanisms end up being tailored to the potential-minimizing information structures, which often have a great deal of structure that is inherited by the mechanism, and vice versa. In deriving mechanisms that are optimal at the potential minimizer, we obtain relatively natural looking mechanisms, with meaningfully ordered actions, reminiscent of classical results in partial implementation, such as Myerson (1981). However, the mechanisms that we derive are not tied to an exogenously given language for types, nor do they require the agents to explicitly report their higher-order beliefs or the information structure itself (even though this is allowed in our model). This is a desirable and emergent feature of our theory.[7] But in spite of the focus on potential minimizers, we can still partially characterize the performance of these mechanisms in other environments and equilibria.

The rest of this paper proceeds as follows. Section 2 describes our model. Section 3 presents our main results on the bounding linear programs. Section 4 is an application to the public expenditure problem and bilateral trade. Section 5 develops our tightness results for revenue maximization with multiple private goods. Section 6 solves a special case of revenue maximization with a single good where the empirical distribution of values is known. Section 7 concludes the paper with a discussion of our assumptions and directions for future research. Appendix A contains additional theoretical results and omitted proofs, and Online Appendix B contains further results and numerical examples.

2 Model

There is a mechanism designer and a finite group of agents indexed by $i \in \{1, \ldots, N\}$. The designer controls an outcome $\omega \in \Omega$, where Ω is finite. The designer and the agents have expected utility preferences over outcomes. In particular, the preferences of agent $i = 1, \ldots, N$ over outcomes and states are represented by the utility index $u_i(\omega, \theta)$, which depends on a payoff-relevant state of the world $\theta \in \Theta$, where Θ is also finite. The designer's preferences are similarly represented by the utility index $w(\omega, \theta)$. The designer has a prior belief about θ, denoted $\mu \in \Delta(\Theta)$, which is held fixed throughout our analysis.[8]

Each agent could choose not to participate in the designer's mechanism and receive a certain state-dependent payoff. We normalize this outside option to zero and interpret agent i's utility as their payoff net of the outside option.

[7]The focus on potential-minimizing information structures also means that the guarantee-maximizing mechanisms that we describe look very different from the mechanisms used in the full implementation literature, which generally require the agents to report their signals in a grand information structure, in addition to sending auxiliary messages that are used to kill off undesirable mixed equilibria, e.g., integer games. The reporting of signals in a grand information structure is antithetical to our goal of building a theory of mechanism design that is not dependent on a complex and artificial language for private information.

[8]We do not assume that $\mu(\theta) > 0$ for all $\theta \in \Theta$, although this assumption will later be imposed for Theorem 2. The distinction between Θ and the support of μ allows the participation security condition to be formulated independently of the prior. In Brooks and Du (2023), we allow for the designer to have a set of priors, which need not all have the same support.

The agents' private information about θ is described by an *information structure*, which consists of: a finite product set of signal profiles $S = \prod_i S_i$,[9] where S_i is agent i's set of signals, and a joint distribution $\sigma \in \Delta(S \times \Theta)$ for which the marginal on Θ is μ. An information structure is denoted $I = (S, \sigma)$, and $\mathcal{I}$ is the set of information structures.[10]

The designer commits to a *mechanism*, which consists of: a finite product set of action profiles $A = \prod_i A_i$, where A_i is agent i's set of actions, and an outcome function $m : A \to \Delta(\Omega)$ that maps action profiles to lotteries over outcomes. An action $a_i \in A_i$ is *participation secure* if $\sum_\omega u_i(\omega, \theta) m(\omega | a_i, a_{-i}) \geqslant 0$ for all a_{-i} and θ. A mechanism is *participation secure* if every agent has an action that is participation secure. We will restrict the mechanism designer to use only mechanisms that are participation secure. This ensures that, regardless of the information structure and other agents' strategies, no agent will have a strict incentive to exit the mechanism, since they can always play a participation secure action and receive a weakly higher payoff than their outside option. A mechanism is denoted by $M = (A, m)$, the set of all mechanisms is $\mathcal{M}$, and the set of participation secure mechanisms are $\mathcal{M}^*$. We assume that a participation secure mechanism exists.

A mechanism and an information structure (M, I) together define a *Bayesian game*, in which a *(behavioral) strategy* for agent i is a mapping $b_i : S_i \to \Delta(A_i)$. A strategy profile $b = (b_1, \ldots, b_N)$ is identified with the function from S to $\Delta(A)$ defined by $b(a|s) = \prod_i b_i(a_i | s_i)$. Expected utility for agent i is

$$U_i(M, I, b) = \sum_{\theta, s, a, \omega} u_i(\omega, \theta) m(\omega | a) b(a|s) \sigma(s, \theta),$$

and the designer's welfare is

$$W(M, I, b) = \sum_{\theta, s, a, \omega} w(\omega, \theta) m(\omega | a) b(a|s) \sigma(s, \theta).$$

A strategy profile b is a *(Bayes Nash) equilibrium* of (M, I) if $U_i(M, I, b) \geqslant U_i(M, I, b_i', b_{-i})$ for all $i = 1, \ldots, N$ and b_i'. The set of equilibria is $\mathcal{E}(M, I)$, which we note is always non-empty, since the mechanism and information structure are both finite.

The *guarantee* of a mechanism M is

$$G(M) = \inf_{I \in \mathcal{I}} \inf_{b \in \mathcal{E}(M,I)} W(M, I, b),$$

that is, the infimum welfare of the designer across all information structures and equilibria. The *potential* of an information structure I is

$$P(I) = \sup_{M \in \mathcal{M}^*} \sup_{b \in \mathcal{E}(M,I)} W(M, I, b),$$

that is, the supremum welfare of the designer across all participation-secure mechanisms and equilibria. It is immediate that for any $M \in \mathcal{M}^*$ and $I \in \mathcal{I}$, $G(M) \leqslant P(I)$. The purpose of

[9] Throughout our exposition, a sum or a product with respect to a variable without qualification means that the operation should be applied for all values of the variable. In this case, the product is over all i, that is, $i = 1, \ldots, N$.

[10] The set of (finite) information structures is defined by identifying finite sets of signals with finite subsets of the natural numbers. Likewise for the set of (finite) mechanisms.

this paper is to study mechanisms that maximize the guarantee and information structures that minimize the potential, i.e., solutions to the programs

$$\sup_{M \in \mathcal{M}^*} G(M) = \sup_{M \in \mathcal{M}^*} \inf_{I \in \mathcal{I}} \inf_{b \in \mathcal{E}(M,I)} W(M, I, b) \tag{MAX-G}$$

$$\inf_{I \in \mathcal{I}} P(I) = \inf_{I \in \mathcal{I}} \sup_{M \in \mathcal{M}^*} \sup_{b \in \mathcal{E}(M,I)} W(M, I, b) \tag{MIN-P}$$

We illustrate these definitions with an application to the *public expenditure problem*: Designer chooses the total expenditure $E \in [0, 1]$ on a public good. The designer must balance the budget, and therefore chooses the contribution $e_i \geqslant 0$ of each agent i, with $\sum_i e_i = E$. Agent i's value from a unit of expenditure is $\theta_i \geqslant 0$, and the utility index of agent i is $u_i = \theta_i E - e_i$. The designer's objective is to maximize utilitarian welfare: $w = (\sum_i \theta_i - 1)E$. The designer can choose any participation secure mechanism, which maps actions to individual contributions. We can embed this problem in our framework by setting $\Theta \subset \mathbb{R}^N_+$ and $\Omega = \{0, 1, \ldots, N\}$, where $\omega = 0$ is the outcome that $E = 0$, and $\omega \neq 0$ is the outcome that $e_\omega = E = 1$, i.e., agent ω contributes the full expenditure. Thus, we can interpret $E(a) = 1 - m(0|a)$ as the (expected) total expenditure implemented by the designer, and $e_i(a) = m(i|a)$ is agent i's (expected) contribution. We assume that for every i there exists a $\theta \in \Theta$ such that $\theta_i = 0$, so that participation security is equivalent to the existence of an action 0 for which $e_i(0, a_{-i}) = 0$ for all a_{-i}, that is, the good may be produced but agent i refuses to bear any part of the cost. We return to this application in Section 4, where we informally derive guarantee-maximizing mechanisms and potential-minimizing information structures. (The problem is treated rigorously in Brooks and Du (2023).)

Another application is the *optimal auctions problem*: There are L goods for sale, indexed by $l = 1, \ldots, L$. The set of value profiles is given by a finite set $\Theta \subset \mathbb{R}^{NL}_+$, with $\theta_{i,l}$ being agent i's value for good l. We further assume that $\mu(\theta) > 0$ for all $\theta \in \Theta$, and for every i, there is a $\theta \in \Theta$ with $\theta_{i,l} = 0$ for all l. The outcome consists of an allocation of each good to one of the agents (or withholding the good) and also a transfer that each agent i pays to the mechanism. We write $q_{i,l}$ for the likelihood that agent i is allocated good l and t_i for agent i's transfer. The allocation satisfies $q_{i,l} \geqslant 0$ for all i, l and $\sum_i q_{i,l} \leqslant 1$ for all l, and the transfers are unrestricted. Each agent i has quasilinear-additive utility $u_i = \sum_l \theta_{i,l} q_{i,l} - t_i$, and the designer's payoff is revenue $w = \sum_i t_i$. In the special case where $L = 1$, we will drop the l subscript on the values and allocations. Because of our assumption that values could be zero for all goods, participation security is equivalent to requiring that each agent i has an action 0 such that $t_i(0, a_{-i}) \leqslant 0$ for all a_{-i}.[11]

The allocation q can be mapped into in our formalism in a similar manner as with the public expenditure problem, where the designer chooses between the finitely many alternatives of whether to withhold the good or to allocate to one of the agents. We can also embed the transfers by fixing a large maximum transfer $\bar{t}$ and interpreting t_i as the expectation of a lottery on $\{-\bar{t}, \bar{t}\}$. Our analysis in Section 5 will in fact work with the cleaner limit model where t_i is unrestricted, to which our main results readily generalize (and which we prove formally in Online Appendix B.2).

[11] In Section 5, we will actually work with an even stronger form of participation security that requires $t_i(0, a_{-i}) = 0$ for all a_{-i}, which makes our tightness result Theorem 2 even stronger as well.

The single-good version of the optimal auctions problem was studied by Bergemann, Brooks, and Morris (2016), Du (2018), and Brooks and Du (2021b) in the case of common values (where $\theta_{1,1} = \cdots = \theta_{N,1}$) and by Brooks and Du (2021a) where each agent has a known expected value for the good. In Section 6, we solve the single good problem when the empirical distribution of the agents' values is known, but it is unknown which agent has which value.

3 The Bounding programs

We now derive the bounding linear programs and state our first result, Theorem 1, which asserts that these programs do indeed provide a lower bound on the max guarantee and an upper bound on the min potential. A discussion follows.

3.1 Deriving the bounding programs

3.1.1 Preliminaries

The bounding programs are parametrized by a set which will represent actions in a mechanism for the lower bound and signals in an information structure for the upper bound. Specifically, for each i and $k \in \mathbb{N}$, the set of actions/signals of a given agent is

$$X_i(k) = \left\{ \frac{l}{k} \middle| 0 \leqslant l \leqslant k^2, l \in \mathbb{N} \right\},$$

and $X(k) = \prod_i X_i(k)$ is the set of action/signal profiles. Note that $X_i(k)$ has $k^2 + 1$ elements. As k goes to infinity, the number of actions and signals becomes arbitrarily large, and "fills in" the non-negative real line.

At this point, the labels for actions/signals are completely arbitrary. But they will acquire a meaning when we use the natural order on $X_i(k)$ to construct a particular lower bound on the guarantee for a mechanism and a particular upper bound on the potential for the information structure.

3.1.2 The lower bound

We first describe how this works for the lower bound. To start, we will lower bound the max guarantee by constraining the designer to only use mechanisms for which $X(k)$ is the message space and for which the lowest action $0 \in X_i(k)$ is participation secure. Let $\mathcal{M}_k$ be the set of mechanisms defined on the action space $X(k)$, and let $\mathcal{M}_k^0$ be the subset of $\mathcal{M}_k$ that satisfy

$$\sum_{\omega} u_i(\omega, \theta) m(\omega | 0, x_{-i}) \geqslant 0 \ \forall i, \theta, x_{-i}.$$

In words, $\mathcal{M}_k^0$ is the set of mechanisms defined on $X(k)$ for which the action $0 \in X_i(k)$ is participation secure for each agent i. With a slight abuse of the notation, we identify $\mathcal{M}_k$ with the associated set of outcome functions $m : X(k) \to \Delta(\Omega)$, and and likewise for $\mathcal{M}_k^0$.

In the lower bound program, the designer can only use mechanisms in $\mathcal{M}_k^0$. By itself, this is only a substantive restriction in that it bounds the cardinality of the action space. Now, the guarantee of such a mechanism is the minimum welfare of the designer across all information structures and equilibria. As is well known, in computing the optimal value of this information design problem, it is without loss to restrict attention to *Bayes correlated equilibria* (Bergemann and Morris, 2016): These are "direct recommendation" information structures, in which each agent's signal is a recommended action, and the joint distribution of actions and states is such that playing the recommended actions is an equilibrium. Let $\mathcal{I}_k$ be the set of information structures on $X(k)$, which (again, slightly abusing notation) we identify with the subset of $\Delta(X(k) \times \Theta)$ such that the marginal on Θ is the prior μ. To compute the guarantee, we minimize over $\sigma \in \mathcal{I}_k$, subject to *obedience constraints*: For every agent i and "recommended" action $x_i \in X_i(k)$, x_i must be a best response to the conditional distribution of (x_{-i}, θ).

Next, to obtain an even more permissive lower bound, we will make the problem of minimizing the designer's welfare easier by dropping all obedience constraints except for those that are associated with deviating from an action x_i to the next higher action $x_i + 1/k$ (as long as $x_i < k$).[12] In fact, we will go one step further: rather than imposing local-upward equilibrium constraints, we minimize a Lagrangian formed by adding the slack in the obedience constraints into the objective, weighted by a particular choice of Lagrange multipliers. The following multipliers may seem arbitrary, but as we elaborate on in Section 3.3.2 below, they are essentially a normalization of the units for actions in the mechanism (given our focus on local upward obedience constraints).

To be more precise, given a function $f : X(k) \to \mathbb{R}$, we define the discrete upward partial derivative $\nabla_i^+ f(x)$ by[13]

$$\nabla_i^+ f(x) = \begin{cases} (k-1)\left[f(x_i + 1/k, x_{-i}) - f(x)\right] & \text{if } x_i < k; \\ 0 & \text{if } x_i = k. \end{cases} \tag{1}$$

In this notation, the obedience constraint that agent i not benefit by deviating to the next higher action is equivalent to, for all x_i,

$$\sum_{\omega,\theta,x_{-i}} \sigma(x_i, x_{-i}, \theta) u_i(\omega, \theta) \nabla_i^+ m(\omega | x_i, x_{-i}) \leqslant 0. \tag{2}$$

Adding these constraints to the designer's objective yields the Lagrangian

$$\sum_{x,\theta} \sigma(x,\theta) \sum_{\omega} \left[w(\omega,\theta) m(\omega|x) + \sum_i u_i(\omega,\theta) \nabla_i^+ m(\omega|x) \right]. \tag{3}$$

[12] In the interest of arriving at Theorem 1 sooner, we have not yet given any intuition for why this particular relaxation should give a tight lower bound. Such an intuition is given below in Section 3.3.2.

[13] Given that the increment between elements in $X_i(k)$ is $1/k$, a seemingly more natural definition of a discrete derivative would have a factor k rather than $k-1$. Of course, these definitions are equivalent in the limit as k tends to infinity, and by using $k-1$ rather than k, we simplify the arguments for Theorem 2. See Footnote 28 below.

Thus, for any $m \in \mathcal{M}_k^0$, the minimum of (3) across all $\sigma \in \mathcal{I}_k$ is a lower bound on the guarantee of $(X(k), m)$. But notice that the only remaining restriction on the joint distribution σ is that the marginal over θ has to be μ. Hence, the σ that minimizes (3) will, for each θ, put probability one on an action profile x that minimizes the inner sum over ω. We refer to this minimand as the *strategic virtual objective*:

$$\sum_{\omega}\left[w(\omega,\theta)m(\omega|x)+\sum_i u_i(\omega,\theta)\nabla_i^+ m(\omega|x)\right] \tag{4}$$

This is the welfare of the designer plus the sum of the agents' gains from local upward deviations. The lower bounding program is simply maximizing the minimum value of (3) across all $m \in \mathcal{M}_k^0$:

$$\max_{m\in\mathcal{M}_k^0}\sum_{\theta}\mu(\theta)\min_x\sum_{\omega}\left[w(\omega,\theta)m(\omega|x)+\sum_i u_i(\omega,\theta)\nabla_i^+ m(\omega|x)\right]. \tag{LB-G-k}$$

In words, the lower bounding program is to maximize (over $m \in \mathcal{M}_k^0$) the expected (over θ) lowest (over x) strategic virtual objective. In effect, the concern is that information could coordinate the agents on actions with a low strategic virtual objective (which may be dependent on the state θ), and the designer chooses the mechanism in order to guarantee that this kind of coordination will not be too harmful.

3.1.3 The upper bound

A parallel approach leads us to the upper bounding program. We first restrict the set of information structures over which we minimize to those of the form $(X(k), \sigma)$ for $\sigma \in \mathcal{I}_k$. In addition, we obtain an upper bound on the potential by relaxing constraints on the mechanism designer. In particular, we first relax participation security by only requiring that interim expected utilities in equilibrium are non-negative, that is, interim individual rationality. Then, by the revelation principle, it is without loss for the designer to restrict attention to direct mechanisms, in which each agent's action is a report of their signal, truthful reporting is an equilibrium, and interim utilities are non-negative. We then obtain even more permissive upper bound by dropping all individual rationality constraints except for the lowest signal $x_i = 0$, and by dropping all truth-telling constraints except for those associated with a type $x_i > 0$ mimicking the next lower type $x_i - 1/k$.

The remaining individual rationality and truth-telling constraints can be represented concisely by introducing a discrete downward derivative: for $f : X(k) \to \mathbb{R}$, we define

$$\nabla_i^- f(x) = \begin{cases} f(k, x_{-i}) - f(k - 1/k, x_{-i}) & \text{if } x_i = k; \\ k[f(x_i, x_{-i}) - f(x_i - 1/k, x_{-i})] & \text{if } 0 < x_i < k; \\ kf(0, x_{-i}) & \text{if } x_i = 0. \end{cases}$$

Then individual rationality for the lowest type and local downward truth-telling constraints are equivalent to, for all i and x_i,

$$\sum_{\omega,\theta,x_{-i}}\sigma(x_i,x_{-i},\theta)u_i(\omega,\theta)\nabla_i^- m(\omega|x_i,x_{-i}) \geqslant 0. \tag{5}$$

As a final step, we relax the mechanism designer's problem even further by adding these constraints to the objective and letting the designer maximize a Lagrangian:

$$\sum_{x,\theta} \sigma(x,\theta) \sum_{\omega} \left[w(\omega,\theta) m(\omega|x) + \sum_i u_i(\omega,\theta) \nabla_i^- m(\omega|x) \right]. \tag{6}$$

Again, implicit in the definition of ∇_i^- is a particular choice of Lagrange multipliers, which essentially fixes the units for signals in the information structure. Thus, an upper bound on the potential of an information structure of the form $(X(k), \sigma)$ is the maximum of (6) over all $m \in \mathcal{M}_k$.

To see this maximum a bit more clearly, it is helpful to sum (6) by parts, move the discrete downward derivative off of m, and replace it with a discrete upward derivative on σ (the only other term that involves x). As Lemma 2 in Appendix A.1 shows, the correct discrete upward derivative is not given by ∇_i^+, but rather has a slightly different definition, primarily regarding the boundary cases:

$$\tilde{\nabla}_i^+ f(x) = \begin{cases} -f(k, x_{-i}) & \text{if } x_i = k \\ f(k, x_{-i}) - k f(k - 1/k, x_{-i}) & \text{if } x_i = k - 1/k \\ k[f(x_i + 1/k, x_{-i}) - f(x)] & \text{otherwise.} \end{cases} \tag{7}$$

Applying the summation-by-parts formula of Lemma 2, we rewrite Lagrangian (6) as

$$\sum_{x,\omega} m(\omega|x) \sum_{\theta} \left[w(\omega,\theta)\sigma(x,\theta) - \sum_i u_i(\omega,\theta) \tilde{\nabla}_i^+ \sigma(x,\theta) \right].$$

But notice that the mechanism can depend arbitrarily on x, so that the optimum will, for each x, put probability one on an outcome ω that maximizes inner sum over θ. We refer to the inner maximand as the *informational virtual objective*:

$$\sum_{\theta} \left[w(\omega,\theta)\sigma(x,\theta) - \sum_i u_i(\omega,\theta) \tilde{\nabla}_i^+ \sigma(x,\theta) \right] \tag{8}$$

This is welfare of the designer plus the sum of the agents' gains from local downward misreports, as well as the payoffs for the lowest type. The upper bounding program is simply minimizing the maximum value of (6) over all $\sigma \in \mathcal{I}_k$:

$$\min_{\sigma \in \mathcal{I}_k} \sum_x \max_{\omega} \sum_{\theta} \left[w(\omega,\theta)\sigma(x,\theta) - \sum_i u_i(\omega,\theta) \tilde{\nabla}_i^+ \sigma(x,\theta) \right]. \tag{UB-P-k}$$

In words, the upper bounding program is to minimize (over $\sigma \in \mathcal{I}_k$) the expectation (over x) of the highest (over ω) informational virtual objective. In effect, the designer uses the agents' information to select the outcome that maximizes the informational virtual objective, and the information structure is chosen in order to limit the potential benefits from this selection.[14]

[14]In the optimal auctions problem, the informational virtual objective reduces to an interdependent values analogue of the virtual value of Myerson (1981). See Remark 5 below.

3.2 Main result

We are now ready to state the main result of this section. Given an optimization program P, let $W(P)$ denote its optimal value.

Theorem 1. *For all $k \in \mathbb{N}$, we have*

$$W(\text{UB-P-}k) \geqslant W(\text{MIN-P}) \geqslant W(\text{MAX-G}) \geqslant W(\text{LB-G-}k).$$

Moreover,

- *If m solves* (LB-G-k)*, then $G(X(k), m) \geqslant W(\text{LB-G-}k)$.*
- *If σ solves* (UB-P-k)*, then $P(X(k), \sigma) \leqslant W(\text{UB-P-}k)$.*

The formal proof of Theorem 1 is in Appendix A.1. The steps are the same as in the preceding derivation, but we fully write out the programs that are referenced along the way, and we are more explicit in our invocations of duality.

Remark 1. The programs (LB-G-k) and (UB-P-k) are presented as saddle point problems, but they are easily converted into linear programs by introducing auxiliary variables. In particular, (LB-G-k) is equivalent to the linear program:

$$\begin{aligned} &\max_{\substack{m:X(k)\times\Omega\to\mathbb{R}_+,\\ \lambda:\Theta\to\mathbb{R}}} \sum_\theta \mu(\theta)\lambda(\theta) \\ &\text{s.t. } \lambda(\theta) \leqslant \sum_\omega \left[w(\omega,\theta)m(\omega|x) + \sum_i u_i(\omega,\theta)\nabla_i^+ m(\omega|x) \right] \ \forall \theta, x \\ &\sum_\omega u_i(\omega,\theta) m(\omega|0, x_{-i}) \geqslant 0 \ \forall i, x_{-i}, \theta \\ &\sum_\omega m(\omega|x) = 1 \ \forall x \end{aligned} \tag{9}$$

Similarly, (UB-P-k) is equivalent to:

$$\begin{aligned} &\min_{\substack{\sigma:X(k)\times\Theta\to\mathbb{R}_+,\\ \gamma:X(k)\to\mathbb{R}}} \sum_x \gamma(x) \\ &\text{s.t. } \gamma(x) \geqslant \sum_\theta \left[w(\omega,\theta)\sigma(x,\theta) - \sum_i u_i(\omega,\theta)\tilde{\nabla}_i^+ \sigma(x,\theta) \right] \ \forall x, \omega \\ &\sum_x \sigma(x,\theta) = \mu(\theta) \ \forall \theta \end{aligned} \tag{10}$$

Remark 2. The prior μ captures all of the designer's uncertainty about the economy. In our view, asking the designer to specify μ is much more reasonable task than specifying an entire information structure. Even so, a designer may concerned about misspecification of μ. Proposition 4 in Online Appendix B.1.2 shows that if we fix a mechanism $M = (X(k), m)$ and λ that solve (9) at μ and then change the prior to μ', then the associated lower bound on the guarantee for M can decrease by at most $\sum_\theta \lambda(\theta) [\mu(\theta) - \mu'(\theta)]$. In that sense, the model is robust to misspecification of the prior.

3.3 Discussion

The remainder of this section provides further results and commentary on the bounding programs. In particular, we discuss whether and when the bounding programs will be tight, a deeper explanation of why we focus on local equilibrium constraints, and the connection to the literature. This material is not necessary to understand our subsequent applications.

3.3.1 Tightness of the bounds and approximate duality

We say that *the bounds are tight* if

$$\lim_{k\to\infty} W(\text{UB-P-}k) - W(\text{LB-G-}k) = 0.$$

There is a *duality gap* if $W(\text{MIN-P}) > W(\text{MAX-G})$. Importantly, while Theorem 1 asserts that the bounding programs are in fact bounds for the max guarantee and the min potential, it does not assert that the bounds are tight. But if the bounds are tight, then there is no duality gap, and max guarantee is equal to min potential. In that case, by solving (LB-G-k) and (UB-P-k) for k sufficiently large, one can obtain arbitrarily good approximations of the max guarantee and min potential, and associated almost guarantee maximizing mechanisms and almost potential-minimizing information structures. Moreover, these approximate solutions have the property that there is a linear order on actions and signals, and the only relevant equilibrium constraints are those that are local in that order. Also, participation security is imposed only on the lowest action, and participation constraints are imposed only on the lowest type. We will give examples where the bounds are tight in Sections 4 and 6, and sufficient conditions for the bounds to be tight in Section 5. In these applications, the one-dimensionality of actions/signals is associated with very particular forms for guarantee-maximizing mechanisms and potential-minimizing information structures. We also give an example where there is a duality gap in Section 4.

To obtain more intuition for why we might expect the bounds to be tight, consider again the Lagrangian (6). This was an intermediate step in upper bounding the potential for an information structure of the form $(X(k), \sigma)$, and before we summed by parts and solved out the mechanism. But we could have stayed with the saddle point problem in which σ is chosen first and m is chosen second. This is a zero-sum game, where the actions σ and m are elements of compact and convex sets, and the objective is bilinear. By the minimax theorem, the optimal value does not depend on the order of moves. If we reverse the order and choose m first, then we can solve out σ—as we did in deriving (LB-G-k)—to conclude that the value of (UB-P-k) is

$$\max_{m\in\mathcal{M}_k} \sum_\theta \mu(\theta) \min_x \sum_\omega \left[w(\omega,\theta)m(\omega|x) + \sum_i u_i(\omega,\theta)\nabla_i^- m(\omega|x) \right]. \tag{11}$$

This program is almost the same as (LB-G-k), except that instead of imposing participation security as a constraint on the designer, we have priced individual rationality into the Lagrangian, as part of the definition of ∇_i^-. Also, the local equilibrium constraints point towards the types with the binding participation constraints, rather than pointing away from the participation secure actions. This leads to a modified strategic virtual objective as the minimand in (11).

Indeed, as in Remark 1, we can formulate (11) as the linear program

$$\max_{\substack{m:X(k)\times\Omega\to\mathbb{R}_+,\\ \lambda:\Theta\to\mathbb{R}}} \sum_\theta \mu(\theta)\lambda(\theta)$$
$$\text{s.t. } \lambda(\theta) \leqslant \sum_\omega \left[w(\omega,\theta)m(\omega|x) + \sum_i u_i(\omega,\theta)\nabla_i^- m(\omega|x)\right] \;\forall\theta,x \tag{12}$$
$$\sum_\omega m(\omega|x) = 1 \;\forall x$$

This is precisely the dual linear program to (10) (and hence (UB-P-k) as well), where $m(\omega|x)$ is the Lagrange multiplier on the constraint that the maximum informational virtual objective at x is at least that obtained at the outcome ω, and $\sigma(x,\omega)$ is the Lagrange multiplier on the constraint that the minimum (modified) strategic virtual objective at ω is at least that obtained at the action profile x. Indeed, this linear program has almost the same form as (9), except for the aforementioned differences regarding participation constraints and the direction of local equilibrium constraints. It is in this sense that (LB-G-k) and (UB-P-k) are "almost" a dual pair of linear programs.

Intuitively, the gap between (LB-G-k) and (UB-P-k) should disappear in the limit as k goes to infinity, as long as there are solutions (11) that converge to a differentiable function $m : \mathbb{R}_+^N \to \Delta(\Omega)$ and for which the actions 0 are participation secure. In this case, we can approximate that limit with feasible solutions to (LB-G-k) that have a similar value, and hence the bounds will be tight. This observation is formalized as Proposition 6 in Online Appendix B.1.4.

When the bounds are tight, one has the intuition that this approximate duality becomes exact in the limit as k goes to infinity. If we make the conceptual leap from approximate to exact duality, then the usual properties of saddle points of linear programming problems would apply, and a feasible pair (m,σ) are optimal if and only if they satisfy *complementary slackness*: $\sigma(x,\theta) > 0$ only if x minimizes the strategic virtual value for m at θ, and $m(\omega|x) > 0$ only if ω maximizes the informational virtual objective for σ at x. Moreover, the potential upper bound is exactly equal to the guarantee lower bound. To be clear: we do not formally establish that this exact complementary slackness is either necessary or sufficient for (σ,m) to be optimal for (LB-G-k) and (UB-P-k), nor do we think it is true for any finite k. Nonetheless, this form of complementary slackness is present in the limiting solutions that we have constructed thus far in cases where the bounds are tight, including in Brooks and Du (2021a,b, 2023). It has also proven to be a useful heuristic for *deriving* the analytical solution, as we will demonstrate with examples in Sections 4 and 6.

In light of this somewhat speculative discussion, we feel compelled to briefly mention a related phenomenon. When complementary slackness is exactly satisfied, the mechanism M that solves (LB-G-k) maximizes the informational virtual objective for the information structure I that solves (UB-P-k). As a result, if we were to view M as a direct mechanism on I, we know that no agent has an incentive to misreport as the next lower type. But in fact, for the solutions constructed for optimal auctions (Bergemann et al., 2016; Brooks and Du, 2021a,b) and for public goods and bilateral trade (Brooks and Du, 2023), the mechanism turns out to be *globally* incentive compatible, meaning that truthful reporting is an equilibrium of the game (M, I). Flipping the interpretation, we can also view I as a direct recommendation information structure on M, and obeying the recommendation is also an equilibrium. We have previously referred to this phenomenon as the *double revelation principle*. While we do not yet have a general explanation for why global constraints are implied by local at the saddle point, the fact that this rather mysterious structure has manifested itself in both of these applications suggests that it is more than a coincidence. While we do not discuss it further in this paper, the finding that truthful/obedient strategies are an equilibrium at the saddle point is addressed in our related work.

Finally, we note that the solutions to the bounding programs need not be unique. Moreover, even when the bounds are tight, there may be solutions to (MAX-G) and (MIN-P) that are not solutions to the bounding programs. For example, it is in principle possible that the designer could maximize the guarantee by asking the agents to report the information structure itself. Such solutions would be more in the spirit of the literature on full implementation that we referenced in the introduction. However, such solutions are implicitly ruled out by the focus on the bounding program (LB-G-k) and mechanisms that admit a tight lower bound on welfare derived from local obedience constraints.

3.3.2 Further explanation of the bounding programs

We now give a heuristic explanation for why one-dimensional equilibrium constraints should naturally appear in the bounding programs. At a key step in the derivation of the lower bounding program, we formulated the Lagrangian (3) by adding to the designer's objective the slack in local upward obedience constraints, formulated as (2). Similarly, in deriving the upper bounding program, we formulated the Lagrangian (6) by subtracting from the designer's objective the slack in local downward truthtelling constraints and individual rationality for the lowest type, formulated as (5). Implicit in these steps is a particular choice of Lagrange multipliers on obedience, truthtelling, and individual rationality constraints. The logic behind these multipliers can be understood by examining more general Lagrangian relaxations of the potential and the guarantee, where we allow an arbitrary choice of multipliers.

Consider first the lower bound. The full set of obedience constraints is

$$\sum_{x_{-i},\theta,\omega} \sigma(x,\theta) u_i(\omega,\theta) \left[m(\omega|x_i,x_{-i}) - m(\omega|x_i',x_{-i})\right] \geqslant 0 \tag{13}$$

for all i, x_i, and x_i'. For any choice of non-negative multipliers $\alpha_i^{obed}(x_i, x_i')$ on these constraints, we can *subtract* the product of multipliers and the non-negative left-hand side of the obedience constraints from the objective to obtain a saddle point problem that *lower bounds* the max guarantee:

$$\max_{\substack{m\in\mathcal{M}_k\\ \text{s.t. p.s.}}} \min_{\sigma\in\mathcal{I}_k} \sum_{x,\theta,\omega} \sigma(x,\theta) \left[w(\omega,\theta)m(\omega|x) - \sum_{i,x_i'} \alpha_i^{obed}(x_i,x_i') u_i(\omega,\theta) \left[m(\omega|x) - m(\omega|x_i',x_{-i})\right] \right]. \tag{14}$$

In the outer maximization, we have restricted to mechanisms in $\mathcal{M}_k$ that are participation secure (but not necessarily with 0 being the participation secure action).

Consider next the upper bound. Truthtelling constraints can be written in precisely the same manner as obedience in (13), and individual rationality requires that

$$\sum_{x_{-i},\theta,\omega} \sigma(x_i, x_{-i}, \theta) u_i(\omega, \theta) m(\omega | x_i, x_{-i}) \geqslant 0$$

for all i and x_i. For any choice of non-negative multipliers $\alpha_i^{truth}(x_i, x_i')$ and $\beta_i(x_i)$ on truthtelling and individual rationality constraints, we can *add* the product of multipliers and non-negative left-hand sides to the objective to obtain a saddle point problem that *upper bounds* the min potential:

$$\min_{\sigma \in \mathcal{I}_k} \max_{m \in \mathcal{M}_k} \sum_{x,\theta,\omega} \sigma(x,\theta) \Big[w(\omega,\theta) m(\omega|x) + \sum_{i,x_i'} \alpha_i^{truth}(x_i, x_i') u_i(\omega,\theta) \left[m(\omega|x) - m(\omega|x_i', x_{-i}) \right] + \sum_i \beta_i(x_i) u_i(\omega,\theta) m(\omega|x) \Big].$$

We emphasize that these bounds are valid for *any* choice of multipliers.

The generalized bounds differ from one another in three key respects: (i) the order of moves is reversed, (ii) the lower bound imposes participation security on the mechanism, whereas the upper bound prices individual rationality into the objective, and (iii) the programs have different signs on equilibrium constraints.

As mentioned above, the minimax theorem implies that (i) is not an issue: For any choice of multipliers, these are compact finite dimensional bilinear saddle point problems, and we can reverse the order of moves without changing the value.

The differences (ii) and (iii) are more substantive. In formulating (LB-G-k) and (UB-P-k), we engineered the multipliers and the choice of participation secure action to make the two bounds as "similar as possible." To finesse (ii), we fixed a particular action/signal for each agent (labeled as zero) to be the one which is participation secure/has a positive multiplier on individual rationality. To finesse (iii), we have reversed the sign on the equilibrium constraints by linearly ordering actions and signals, dropping non-local constraints, and flipping the direction of binding constraints between the two programs. Specifically, in the lower bound, the binding local obedience constraints point *away* from the participation secure action, and in the upper bound, the binding local truthtelling constraints point *towards* the type with a binding individual rationality constraint. Up to rescaling the multipliers, the resulting programs are simply

$$\max_{m \in \mathcal{M}_k^0} \min_{\sigma \in \mathcal{I}_k} \sum_{x,\theta,\omega} \sigma(x,\theta) \Big[w(\omega,\theta) m(\omega|x) + \sum_i \alpha_i^{obed}(x_i) u_i(\omega,\theta) \nabla_i^+ m(\omega|x) \Big]$$

$$\max_{m \in \mathcal{M}_k} \min_{\sigma \in \mathcal{I}_k} \sum_{x,\theta,\omega} \sigma(x,\theta) \Big[w(\omega,\theta) m(\omega|x) + \sum_i \alpha_i^{truth}(x_i) u_i(\omega,\theta) \nabla_i^- m(\omega|x) \Big],$$

where we write $\alpha_i^{obed}(x_i)$ for the rescaled multipliers on local upward obedience and $\alpha_i^{truth}(x_i)$ for the rescaled multipliers on local downward truthtelling and individual rationality for the lowest type.[15]

These steps get us almost to (LB-G-k) and (UB-P-k) (and more specifically, the latter's dual program (11)). To get the rest of the way, we set $\alpha_i^{obed}(x_i) = \alpha_i^{truth}(x_i) = 1$ for all i and x_i. This choice of multipliers may seem arbitrary, but there is a sense in which it is without loss when k is large. This can be seen most clearly in the continuous limit.[16] Suppose that the action/signal space is all of $\mathbb{R}_+$, and $m(\omega|x)$ is differentiable in x_i for all i. Then the continuous analogue of the strategic virtual objective is

$$\sum_{\theta}\left[w(\omega,\theta)m(\omega|x) + \sum_i u_i(\omega,\theta)\alpha_i(x_i)\frac{\partial}{\partial x_i}m(\omega|x)\right].$$

where $\alpha_i(x_i)$ is the Lagrange multiplier on local equilibrium constraints. Suppose further that α_i is bounded away from zero. Then we can change units so that an action $x_i \in \mathbb{R}_+$ is mapped to an action $g_i(x_i)$, where

$$g_i(x_i) = \int_{y=0}^{x_i} \frac{1}{\alpha_i(y_i)} dy_i.$$

With these new units, the mechanism becomes

$$\tilde{m}(\omega|y) = m(\omega|g_1^{-1}(y_1), \ldots, g_N^{-1}(y_N)),$$

and hence, by the inverse function theorem, we have that at $y = g(x)$,

$$\left.\frac{\partial}{\partial y_i}\tilde{m}(\omega|y)\right|_{y_j=g_j(x_j)\forall j} = \left.\frac{\partial}{\partial x_i}m(\omega|g^{-1}(y))\frac{1}{g_i'(g_i^{-1}(y_i))}\right|_{y_j=g_j(x_j)\ \forall j} = \alpha_i(x_i)\frac{\partial}{\partial x_i}m(\omega|x).$$

The strategic virtual objective in the new units is therefore

$$\sum_{\theta}\left[w(\omega,\theta)\tilde{m}(\omega|y) + \sum_i u_i(\omega,\theta)\frac{\partial}{\partial y_i}\tilde{m}(\omega|y)\right].$$

In effect, the change in units for actions rescales the "size" of a local deviation, so that a unit deviation in the new units is equivalent to a deviation of $1/\alpha_i(x_i)$ in the original units.

The point is that given any mechanism and optimal local multipliers, as long as that mechanism is sufficiently well behaved when k is large, we can adjust the units for actions so that the same mechanism (under the new units) would have exactly the same strategic virtual objective, except that the multipliers are normalized to one. In applications, we have found this to be a natural choice of units, but it is not a theoretical necessity.

[15] Note that these local multipliers are rescaled from the general multipliers, in order to align with the definitions of ∇_i^+ and ∇_i^-. For example, when $x_i < k$, $\alpha_i^{obed}(x_i) = (k-1)\alpha_i^{obed}(x_i, x_i + 1/k)$, and $\alpha_i^{truth}(0) = k\beta_i(0)$.

[16] A precise formulation of convergence to a continuous limit is given before Proposition 6 in Online Appendix B.1.4.

3.3.3 Context within the literature

The fact that local constraints appear so prominently in our theory is not entirely surprising. The pattern of binding local truthtelling constraints that point to a lone type with a binding participation constraint is familiar from the analysis of optimal auctions in Myerson (1981). And yet, strong assumptions on primitives are usually needed for these constraints to be the only ones that bind at the optimum, such as independence, private values, and concavity of the revenue curve (i.e., regularity). An important distinction is that in the classical analysis, the optimal multipliers and mechanism are derived from primitive assumptions about information, whereas in the present model, the multipliers and mechanism and information are all jointly determined.

Still, it is far from obvious that the local structure would emerge as optimal for the applications we describe, and with minimal assumptions on primitives. This structure is however suggested by prior work on guarantees and potentials in the optimal auctions problem (Bergemann, Brooks, and Morris, 2016, 2017, 2019, 2020; Du, 2018; Brooks and Du, 2021a,b). To begin with, Bergemann, Brooks, and Morris (2017) computed the guarantee of the first-price auction and showed that the optimal multipliers on obedience constraints, denoted α^{FPA}, have a particular form: All *upward* constraints bind, and the multiplier depends only on the deviation, and not the recommendation.[17] Recall the general lower bound program (14), and consider that lower bound applied to the guarantee for revenue in the auction setting. Further suppose we set $\alpha^{obed} = \alpha^{FPA}$ (in lieu of the local-upward multipliers used to obtain (LB-G-k)). We do not know what is the mechanism that maximizes (14) given these multipliers, but one feasible choice is the first-price auction itself, and for that mechanism, the optimal value of (14) with the multipliers α^{FPA} is precisely the guarantee of the first-price auction. Thus, the value of (14) with $\alpha^{obed} = \alpha^{FPA}$ and optimized over all participation secure mechanisms must be even higher.

This example shows that imposing a seemingly arbitrary order on actions and an associated pattern on multipliers α^{obed} can yield non-trivial lower bounds on the maximum guarantee. And yet, there were reasons to think that the particular multipliers α^{FPA} would not maximize the lower bound. In particular, in the limit where the number of bidders becomes large, the guarantee of the first-price auction is generally bounded away from total surplus. But Du (2018) showed that when values are common, there is a sequence of mechanisms whose guarantees converge to total surplus when the number of bidders goes to infinity. This is demonstrated using lower bounds on the mechanisms' guarantees which are derived from local obedience constraints that point away from an action which is participation secure.[18,19] Thus, for revenue maximization in common value auctions, this

[17] Here, "upward" is with respect to the natural order on bids, and the lowest action in the first-price auction, a bid of zero, is participation secure.

[18] Further evidence that α^{FPA} are not the optimal multipliers came from Bergemann, Brooks, and Morris (2020), who calculated the potential of information structure that minimizes expected revenue for the first-price auction. That potential turns out to be is strictly greater than the first-price auction's revenue guarantee. Thus, the pair of the first-price auction and its own worst-case information structure is not a saddle point, even though a potential-minimizing information structure necessarily minimizes the welfare of the guarantee-maximizing mechanism (that is, when the duality gap is zero).

[19] The working paper version of Du (2018) used discrete actions, whereas the published version worked in the continuum limit.

pattern of obedience constraints is approximately optimal with many bidders. Brooks and Du (2021b) pursued this logic even further in the context of common value auction, and showed that max guarantee equals min potential for any fixed number of bidders, and they prove it using the pattern of constraints underlying the bounding programs (LB-G-k) and (UB-P-k), specialized to that setting, and in the continuum limit. Our Theorem 1 distills and extends this logic to more general environments.

We conclude this section with two other comments on the literature. In contrast to the discrete model studied here, Brooks and Du (2021b) allow for mechanisms and information structures that have arbitrary measurable spaces of actions and signals. As alluded to previously, this appears to be necessary in order to exactly attain the max guarantee and min potential. Moreover, the critical action and signal spaces end up being the non-negative real line, and the optimal mechanism and information structures are almost everywhere differentiable. The obvious advantage of working directly in the continuum limit is that it allows one to use calculus in deriving and characterizing solutions. In our applications below, we will often work in the continuum limit, guided by the discrete bounding programs (LB-G-k) and (UB-P-k) as a heuristic. On the other hand, the discrete model allows us to rely on the elementary but powerful theory of finite dimensional linear programming. Also, the restriction to finite mechanisms and information structures dispels any concerns that the desirable properties of our solutions might be due to a controversial use of infinite action or signal spaces, as in the integer games commonly used in full implementation with mixed strategies.

Finally, when working directly with infinite mechanisms and information structures, Brooks and Du (2021b) finessed the issue of equilibrium existence by employing a novel solution concept called a *strong maxmin solution*, which is a triple (M, I, b), where $b \in \mathcal{E}(M, I)$, and $G(M) = P(I)$. In Online Appendix B.1.3, we define an analogous ϵ-strong maxmin solution, appropriate to the discrete setting where the optimum may only be attained in the large k limit. Proposition 5 shows the equivalence between "max guarantee equals min potential" and the existence of ϵ-strong maxmin solutions for arbitrary ϵ.

4 Social welfare and public expenditure

4.1 Setup

We now describe an application of our methodology to the public expenditure problem (Samuelson, 1954; Güth and Hellwig, 1986). This application is developed in full rigor and greater generality in Brooks and Du (2023). Here we present an informal overview of the solution, in a special case where the social value of the good is known. Afterwards, we reinterpret the case of $N = 2$ as a model of bilateral trade.

The model, previously introduced in Section 2, has the agents' contributions to the public expenditure represented by $e \in \mathbb{R}^N_+$ such that $E = \sum_i e_i \leqslant 1$. Each agent has a value θ_i from the public expenditure, and hence a utility $u_i = \theta_i E - e_i$. The designer wants to maximize the social welfare $w = \sum_i u_i = (\sum_i \theta_i - 1)E$. We further assume that the set of value profiles Θ consists of the vectors $\theta \in \mathbb{R}^N$ where for some i, $\theta_i = \overline{\theta}$ and $\theta_{-i} = 0$, and the prior μ is uniform on Θ. Thus, there is common knowledge that values are non-negative and the social value is $\sum_i \theta_i = \overline{\theta}$.[20] We further assume that $\overline{\theta} \geqslant 1$, so that the socially efficient outcome is full expenditure. While the social value of expenditure is known, what is unknown is the agents' idiosyncratic values. Moreover, each agent can opt out of the mechanism and pay nothing. This gives rise to a free-rider problem: Agents have the option of behaving as if their value is low, so as to avoid paying for the public good, while still enjoying its benefits.

4.2 Evidence from Simulations

In Figure 1, we plot features of numerical solutions to the bounding programs (LB-G-k) and (UB-P-k) when $N = 2$, $\overline{\theta} = 3$ and $k = 30$. Figure 1 reveals some striking structures, and clearly suggests the functional form of the saddle point. We use $\overline{E}$ and $\overline{e}_i$, and $\overline{\sigma}$ to denote the mechanism and information structure from the numerical solution (and subsequently the optimal mechanism and information structure at the saddle point).

To start, the top-left panel is a contour plot of the total expenditure function $\overline{E}(x) = 1 - \overline{m}(0 \mid x)$ from (LB-G-k). It is clear that total expenditure depends only on the aggregate action $\Sigma x \equiv \sum_i x_i$. The middle-left panel shows that $\overline{E}$ is increasing and concave in Σx, and hits 1 at a finite level, which we denote by $\overline{y}$. Finally, the bottom-left panel shows agent 1's expenditure share $\overline{e}_1(x)/\overline{E}(x)$ as a function of x_1, holding fixed Σx at various levels. The expenditure share is clearly linear in x_1, interpolating from 0 to 1, meaning that $\overline{e}_1(x)/\overline{E}(x)$ the proportional fraction $x_1/\Sigma x$. We refer to a mechanism of this form as a *proportional cost-sharing mechanism.*

Turning now to the information structure, the top-right panel is a contour plot of the log of the probability mass function of signals $\overline{\rho}(x) = \sum_{\theta \in \Theta} \overline{\sigma}(x, \theta)$ from (UB-P-k). Clearly, $\overline{\rho}$ depends only on the aggregate signal Σx. The middle-right panel shows that $\overline{\rho}$ discontinuously drops to near zero when the aggregate signal Σx exceeds a certain threshold, which is close to the point where $\overline{E}$ hits one.[21] The bottom-right panel shows agent 1's interim value $\overline{v}_1(x) = \sum_\theta \theta_1 \overline{\sigma}(x, \theta)/\overline{\rho}(x)$. Again, this function is linear in x_1, holding Σx fixed, indicating that it also has a proportional form: $v_1(x) = \overline{\theta} x_1/\Sigma x$.

We will presently use these functional forms and the complementary-slackness heuristic discussed in Section 3.3.1 to deduce the functional forms of $\overline{E}$ and $\overline{\rho}$ and the threshold $\overline{y}$. We will also sketch the argument for why the bounds are tight. The heuristic argument will appeal to a continuous approximation when k is large, and supposes that the total expenditure and signal density functions converge to limits that are differentiable functions on the action/signal space $\mathbb{R}^N_+$.

[20] Our results would remain the same if we relaxed the symmetry assumption and instead just assumed that values are positive and the social value is *at least* $\overline{\theta}$. This is the formulation adopted in Brooks and Du (2023). That paper also considers an extension where there are lower and upper bounds on the social value of the good, and a lower bound on the expectation.

[21] In the simulations, the density is never *exactly* zero. One reason is that the barrier algorithm used to compute the solution approaches the optimum from the interior of the feasible set. But more generally, one of the criteria used for convergence by numerical algorithms is that the duality gap is below a certain strictly positive threshold. For action/signal profiles above the boundary where the allocation hits one and the density drops discontinuously, the likelihood is so small that the associated violation of complementary slackness (and the corresponding contribution to the duality gap) is negligible.

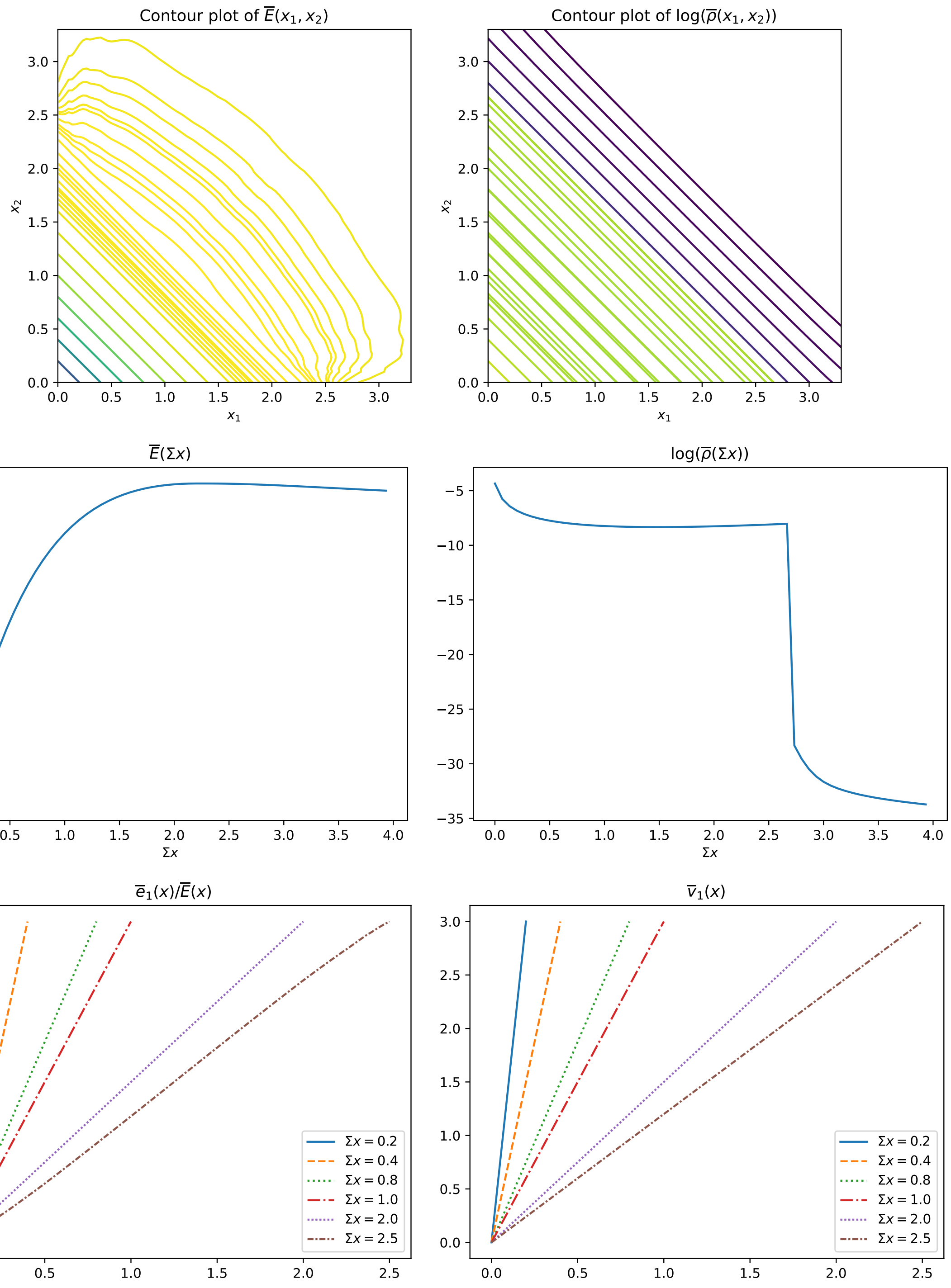

Figure 1: Numerical solutions for public goods when $N = 2$, $\overline{\theta} = 3$, and $k = 30$.

4.2.1 Guarantee-maximizing mechanism

The logic behind the proportional cost-sharing mechanism can be understood by examining its strategic virtual objective. First observe that social welfare is

$$\sum_{\omega} w(\omega, \theta) m(\omega|x) = (\overline{\theta} - 1) E(x),$$

where we allow total expenditure to be an arbitrary function of $x \in \mathbb{R}_+^N$. Moreover, the strategic adjustment is

$$\sum_{\omega} u_i(\omega, \theta)\nabla_i m(\omega|x) = \theta_i \nabla_i E(x) - \nabla_i e_i(x),$$

where $\nabla_i = \partial/\partial x_i$ is the partial derivative with respect to x_i. (To reiterate, in this heuristic derivation, we substitute the partial derivative ∇_i for the discrete upward derivative ∇_i^+.) Hence, the strategic virtual objective (4) is

$$(\overline{\theta} - 1)E(x) + \sum_i \left(\theta_i \nabla_i E(x) - \nabla_i e_i(x)\right).$$

Now, if $E(x) = \widehat{E}(\Sigma x)$ and $e_i(x) = (x_i/\Sigma x)\widehat{E}(\Sigma x)$, then (4) further reduces to

$$(\overline{\theta} - 1)(\widehat{E}(\Sigma x) + \widehat{E}'(\Sigma x)) - \frac{(N-1)\widehat{E}(\Sigma x)}{\Sigma x}.$$

In other words, the strategic virtual objective *depends only on the aggregate action and the social value.* As a result, the scope for information to depress welfare is limited to its effect on the aggregate action. Note that participation security of $x_i = 0$ is satisfied only if $\widehat{E}(0) = 0$, which is a feature of the solution we now derive.

In fact, the functional form for $\overline{E}$ is obtained by pursing this logic one step further, and making the strategic virtual objective independent of the aggregate action as well. This is also suggested by the complementary slackness heuristic of Section 3.3.1: the simulations indicate that $\sigma(x, \theta) > 0$ for all θ and all x for which $\Sigma x \leqslant \overline{y}$, and hence all such x must minimize the strategic virtual objective. Thus, for some constant $\lambda \geqslant 0$, $\widehat{E}$ must solve the following linear first-order ODE

$$(\overline{\theta} - 1)(\widehat{E}(y) + \widehat{E}'(y)) - \frac{N-1}{y}\widehat{E}(y) = \lambda.$$

The solution, subject to the initial condition that $\widehat{E} = 0$ when $y = 0$, is

$$\widehat{E}(y; \lambda) \equiv \frac{\lambda}{\overline{\theta} - 1} \int_{z=0}^{y} \exp(z - y) \left(\frac{z}{y}\right)^{-(N-1)/(\overline{\theta}-1)} dz. \tag{15}$$

The integral converges as long as $\overline{\theta}/N > 1$. We now maintain this as an assumption on parameters, which will be discussed further below.

In order for the mechanism to be feasible, we can only use this functional form until the total expenditure hits 1. This occurs at a boundary which we implicitly define as $\widehat{y}(\lambda)$. For λ sufficiently small, $\widehat{y}(\lambda)$ is infinite (the expenditure never hits one) but for λ sufficiently large, the boundary is finite. For $y > \widehat{y}(\lambda)$, we set $\widehat{E}(y; \lambda) = 1$, so that the strategic virtual objective for $y > \widehat{y}(\lambda)$ is

$$(\overline{\theta} - 1) - \frac{N-1}{y}.$$

Since this expression is increasing in y, the minimum strategic virtual objective for $y \geqslant \widehat{y}(\lambda)$ is attained at $y = \widehat{y}(\lambda)$, at a value of

$$\widehat{\lambda}(\lambda) \equiv (\overline{\theta} - 1) - \frac{N-1}{\widehat{y}(\lambda)}.$$

This is again consistent with complementary slackness: for the information structure suggested by the simulations, there is zero probability of $\Sigma x > \widehat{y}(\lambda)$, and hence these action profiles need not minimize the strategic virtual objective.

By construction, the strategic virtual objective below the boundary $\widehat{y}(\lambda)$ is constant and equal to λ. Therefore, the overall minimum strategic virtual objective is $\min\{\lambda, \widehat{\lambda}(\lambda)\}$. Note that $\widehat{E}(y;\lambda)$ is increasing pointwise in λ, so that both $\widehat{y}(\lambda)$ and $\widehat{\lambda}(\lambda)$ are decreasing in λ. Hence, the expected minimum strategic virtual objective is maximized by choosing λ as large as possible subject to $\lambda \leqslant \widehat{\lambda}(\lambda)$. This is achieved by the $\overline{\lambda}$ that solves $\lambda = \widehat{\lambda}(\lambda)$, the optimal boundary is $\overline{y} = \widehat{y}(\overline{\lambda})$, and the optimal aggregate expenditure function is $\overline{E}(y) = \widehat{E}(y;\overline{\lambda})$. Finally, we note that the boundary condition $\overline{E}(\overline{y}) = 1$ implies that the guarantee of this mechanism is

$$\overline{\lambda} = (\overline{\theta} - 1)\frac{\exp(\overline{y})\overline{y}^{-(N-1)/(\overline{\theta}-1)}}{\int_{z=0}^{\overline{y}} \exp(z) z^{-(N-1)/(\overline{\theta}-1)} dz}.$$

This formula will be useful for comparing this guarantee with the potential of the information structure that we construct in the next subsection.

4.2.2 Potential-minimizing information structure

Just as with the guarantee-maximizing mechanism, we can understand the information structure in the simulations by examining the informational virtual objective. Let $\rho(x)$ denote the density of the signal profile x and let $v_i(x)$ denote the conditional expectation of θ_i given x. Again, supposing that these are differentiable functions on $\mathbb{R}_+^N$, we have

$$\sum_{\theta} w(\omega,\theta)\sigma(x,\theta) = \mathbb{I}_{\omega\neq 0}(\overline{\theta} - 1)\rho(x)$$

and

$$\sum_{\theta} u_i(\omega,\theta)\nabla_i \sigma(x,\theta) = \nabla_i\left[(v_i(x)\mathbb{I}_{\omega\neq 0} - \mathbb{I}_{\omega=i})\rho(x)\right].$$

Hence, the informational virtual objective (8) at the outcome ω is

$$\mathbb{I}_{\omega\neq 0}(\overline{\theta} - 1)\rho(x) - \sum_i \nabla_i\left[(v_i(x)\mathbb{I}_{\omega\neq 0} - \mathbb{I}_{\omega=i})\rho(x)\right].$$

Now, with the functional forms $\rho(x) = \widehat{\rho}(\Sigma x)$ and $v_i(x) = \overline{\theta}x_i/\Sigma x$ suggested by the simulations, the informational virtual objective further reduces to

$$\mathbb{I}_{\omega\neq 0}\left((\overline{\theta} - 1)(\widehat{\rho}(\Sigma x) - \widehat{\rho}'(\Sigma x)) - \overline{\theta}\frac{N-1}{\Sigma x}\widehat{\rho}(\Sigma x)\right).$$

In other words, the informational virtual objective *only depends whether full expenditure is implemented, and not on how the cost is shared.* This limits the mechanism designer's ability to increase welfare by controlling the agents' individual shares.

The optimal $\widehat{\rho}$ can be deduced by pursuing this logic one step further, and making the informational virtual objective independent of ω. This is again suggested by the complementary slackness heuristic: the optimal expenditure is interior when $\Sigma x \leqslant \overline{y}$, meaning that at such signal profiles, all ω are implemented with positive probability. Hence, the informational virtual objective must be equal for all ω, and therefore it is exactly zero. This is equivalent to the first-order linear ODE

$$(\overline{\theta}-1)(\widehat{\rho}(y)-\widehat{\rho}'(y))-\overline{\theta}\frac{N-1}{y}\widehat{\rho}(y)=0,$$

whose solution is

$$\overline{\rho}(y)=\begin{cases}\frac{\exp(y)y^{-(N-1)\overline{\theta}/(\overline{\theta}-1)}}{\int_{z=0}^{\overline{y}}\exp(z)z^{-(N-1)\overline{\theta}/(\overline{\theta}-1)}\frac{z^{N-1}}{(N-1)!}dz} & \text{if } y\leqslant\overline{y};\\ 0 & \text{if } y>\overline{y}.\end{cases}\tag{16}$$

Note that the constant of integration is determined so that $\overline{\rho}$ integrates to 1 on the simplex $\{x\in\mathbb{R}_+^M|x\leqslant\overline{y}\}$. Again, the integrals in this expression converge as long as $\overline{\theta}/N>1$.

Equation (16) ensures that the informational virtual objective is zero for every outcome when $\Sigma x<\overline{y}$. Moreover, when $\Sigma x>\overline{y}$, the density and the informational virtual objective are both zero. But this leads to a minor paradox: If the informational virtual objective is zero everywhere, then it seems that the potential is zero as well. However, we must keep in mind that when $\Sigma x=\overline{y}$, the density $\overline{\rho}$ drops discontinuously to zero, and so the corresponding informational virtual objective is infinite. Thus, in order to calculate the contribution of the boundary to the potential, we have have to reintroduce a discrete upward deviation of size $1/k$, for which the associated informational virtual objective at $\Sigma x=\overline{y}$ is

$$\mathbb{I}_{\omega\neq 0}(\overline{\theta}-1)\overline{\rho}(\Sigma x)-\sum_i\tilde{\nabla}_i^+\left[\left(\overline{\theta}\frac{x_i}{\Sigma x}\mathbb{I}_{\omega\neq 0}-\mathbb{I}_{\omega=i}\right)\overline{\rho}(\Sigma x)\right]\approx\mathbb{I}_{\omega\neq 0}(\overline{\theta}-1)\overline{\rho}(\overline{y})k.$$

since the left-hand side is dominated by $\tilde{\nabla}_i^+\overline{\rho}(\overline{y})=(0-\overline{\rho}(\overline{y}))k$.[22] In particular, the informational virtual objective is positive when $\omega\neq 0$ and blows up as k goes to infinity. As a result, the optimal outcome at the boundary $\Sigma x=\overline{y}$ is full expenditure, just as we constructed in the previous subsection and also just as we observed in the simulations.

At the same time, as k goes to infinity, the mass on the boundary goes to zero, and is approximately $\overline{y}^{N-1}/(k(N-1)!)$. The boundary's overall contribution to the informational

[22]The discrete derivative here can be interpreted as a deviation in the continuous mechanism, wherein an agent increases their reported signal by $1/k$. Such deviations are needed to obtain a tight upper bound on the potential, given the discontinuity in ρ at the upper bound of the support. See Brooks and Du (2023) for details. The public expenditure problem thus demonstrates that discrete local equilibrium constraints are not the same as first-order conditions, and the former may be needed to pin down the value when utilities are not smooth.

virtual objective is therefore approximately

$$(\overline{\theta}-1)\overline{\rho}(\overline{y})\frac{\overline{y}^{N-1}}{(N-1)!}=(\overline{\theta}-1)\frac{\exp(\overline{y})\overline{y}^{-(N-1)/(\overline{\theta}-1)}}{\int_{z=0}^{\overline{y}}\exp(z)z^{-(N-1)/(\overline{\theta}-1)}dz}=\overline{\lambda}. \tag{17}$$

Thus, the boundary's contribution to the potential does not vanish in the limit as k goes to infinity. Moreover, this expression for the potential exactly coincides with the guarantee constructed in the previous subsection. A fortiori, $\overline{\lambda}$ is both the max guarantee and the min potential, and moreover, the bounds in Theorem 1 coincide. See Brooks and Du (2023) for a rigorous proof.

4.3 Discussion

To our knowledge, the proportional cost-sharing mechanisms are new to the literature. In mitigating free riding, it is natural to consider agents' marginal incentives to move their actions in a direction that reduces their contribution. The strategic virtual objective is precisely the sum of welfare and the agents' marginal utilities with respect to their own actions. By making the strategic virtual objective invariant to who pays for the good, these mechanisms are resistant to adverse outcomes in which funding disproportionately depends on agents who have an outsized marginal incentive to free ride, due to the particulars of the information structure or equilibrium. Moreover, the form of the expenditure function exactly balances welfare against the aggregate marginal incentive. As a result, if expected total welfare were too low, then the expected aggregate marginal utility would be too high, and some agent would benefit by increasing their action. This is the logic that sustains the welfare guarantee.

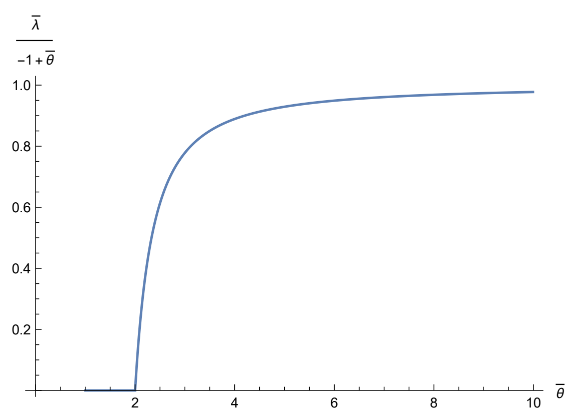

Figure 2: Optimal guarantee as a fraction of the efficient surplus, $N = 2$.

For the case of $N = 2$, Figure 2 depicts the max guarantee as a fraction of the efficient welfare, $\overline{\lambda}/(\overline{\theta}-1)$, as we vary $\overline{\theta}$. For $\overline{\theta} > 2$, the guarantee starts near zero, increases, and eventually converges to the efficient surplus. This is the range satisfying our parametric assumption that $\overline{\theta}/N > 1$, which we adopted after equation (15). When $\overline{\theta} \leqslant 2$, the max guarantee is zero. This is not surprising when $\overline{\theta} \leqslant 1$, since in this case production is inefficient. But when $\overline{\theta} \in [1, 2]$, the max guarantee is zero, even though there is common knowledge that full expenditure is efficient.

Indeed, if $\overline{\theta}/N \leqslant 1$, then the guarantee of any participation secure mechanism is zero: Suppose the agents have no information, so that each agent's interim expected value is $\overline{\theta}/N$. Under such information, it is an equilibrium for all agents to play the participation secure action with probability one. To see why, suppose agent i deviates and induces total expenditure E. The other agents' actions are participation secure, so $e_j = 0$ for all $j \neq i$. Hence, $e_i = E$, and agent i's payoff is $(\overline{\theta}/N - 1)E \leqslant 0$. In effect, any agent who deviates has to provide all of the funds, which may not be worthwhile when the social-value per capita is less than 1.

A naïve reaction might be that the issue is equilibrium selection: Under no information, there are obviously alternative mechanisms and equilibria under which there is full expenditure.[23] However, it turns out that when $\overline{\theta}/N \leqslant 1$, there are information structures for which the potential is arbitrarily close to zero, meaning that all mechanisms and all equilibria generate negligible surplus.[24]

These perverse information structures sometimes involve large differences in the agents' interim values. In the most extreme cases, the value is zero for all but one of the agents. One way to forestall the collapse to zero potential would be to limit the heterogeneity in values. For example, suppose $\Theta = \{(0, \ldots, 0), (\overline{\theta}/N, \ldots, \overline{\theta}/N)\}$ and $\mu(\{(\overline{\theta}/N, \ldots, \overline{\theta}/N)\}) = 1$, where $\overline{\theta} \in (1, N)$. Then the potential is the efficient surplus $\overline{\theta} - 1$, and it can be achieved by the binary-action mechanism described in Footnote 23. However, since Θ contains a state with $\theta_i = 0$ for every i, any participation secure mechanism will still have an equilibrium with zero expenditure, so that the max guarantee is zero, and the duality gap is positive.[25]

That the potential may be zero even though there is common knowledge that full expenditure is efficient indicates how demanding participation security is, and it may be deemed too demanding, depending on the circumstances. Even so, as long as the social value is relatively large, proportional cost-sharing mechanisms attain non-trivial guarantees, even when there can be extreme heterogeneity across agents and even with such a strong assurance that agents will be willing to participate.

As mentioned above, Brooks and Du (2023) rigorously develops the public expenditure application. The analysis is unchanged if we regard $\overline{\theta}$ as only a lower bound on the social

[23] For example, consider the mechanism in which agents can either opt out or opt in; there is full expenditure only if all agents opt in, in which case they share the cost equally, and otherwise the total expenditure is zero. Under this mechanism and no information, it is an equilibrium for all agents to opt in (under the hypothesis that full expenditure is socially efficient).

[24] Such is the case for information structures of the same form as in (16) but with a positive lower bound for the signals. The resulting potential would be the same as (17), but the integral in the denominator would range from $z = \underline{y} > 0$ to $\overline{y}$. As $\underline{y} \to 0$, the denominator blows up, and the potential converges to zero. See Brooks and Du (2023) for details.

[25] In this example, Θ is a strict superset of the support of the prior. But it is straightforward to enrich the example so that Θ is equal to the support of the prior. Suppose that μ puts probability $1 - \epsilon$ on all agents having a value of $\overline{\theta}/N \in (1/N, 1)$ and probability ϵ that all agents have a value zero. For any information structure consistent with this prior, consider the direct mechanism that implements full expenditure if and only if the expected social value given the reported signal profile is greater than 1. Clearly, the probability that full expenditure is interim efficient converges to one as $\epsilon \to 0$, so the potential converges to the efficient surplus, even though the max guarantee is zero.

value. They also show that proportional-cost sharing maximizes the guarantee when $\overline{\theta}$ has a known expectation and known bounds on the support.

Finally, Brooks and Du (2023) show that the model with $N = 2$ can be reinterpreted as a model of bilateral trade, in which the seller's value is $\theta_1 \in \{0, \overline{\theta}\}$, both equally likely, and the buyer's value is $\theta_1 + \overline{\theta} - 1$. Whether or not the good is produced is reinterpreted as whether or not trade occurs, and the sharing of the cost corresponds to the terms of trade. The proportional cost-sharing mechanism is reinterpreted as a *proportional-price trading mechanism*, where trade occurs with probability $\overline{E}(x_1 + x_2)$ at the price

$$p(x) = \overline{\theta} - \frac{x_1}{x_1 + x_2}.$$

This mechanism is participation secure, because if $x_1 = 0$, trade only occurs at a price equal to the highest value of the seller, and if $x_2 = 0$, then trade only occurs at a price equal the lowest value of the buyer. The results on the public expenditure problem imply that the min potential for gains from trade is zero when $1 < \overline{\theta} < 2$, even though it is common knowledge that trade is efficient. This illustrates that welfare maximization in bilateral trade may be even more challenging than suggested by either the lemons market of Akerlof (1970) or the low-welfare information structures constructed by Carroll (2016) in the context of posted price mechanisms.

5 Optimal multi-good auctions

We next consider the optimal auctions problem introduced in Section 2 and prove non-constructively that the bounds are tight. In this section we maintain the full support assumption of $\mu(\theta) > 0$ for every $\theta \in \Theta$.

5.1 Solving out transfers

As a preliminary step, we solve out the transfers from the bounding programs, and replace them with a simpler object, the *aggregate excess growth*.[26] In a slight abuse of notation, we define $\mathcal{M}_k^0$ to be the set of allocation and transfer rules (q, t) defined on the action profile space $X(k)$, and for which $t_i(0, a_{-i}) = 0$ for all i and a_{-i}. Using the functional forms for the agents' and the designer's preferences, the bounding programs are

$$\max_{(q(\cdot),t(\cdot))\in\mathcal{M}_k^0} \sum_{\theta} \mu(\theta) \min_{x} \left[\sum_i t_i(x) + \sum_i \left(\sum_l \theta_{i,l} \nabla_i^+ q_{i,l}(x) - \nabla_i^+ t_i(x) \right) \right]; \tag{18}$$

$$\min_{\sigma\in\mathcal{I}_k} \sum_x \max_{(\tilde{q},\tilde{t})\in(\Delta\{0,\ldots,N\})^L\times\mathbb{R}^N} \sum_{\theta} \left[\sum_i \tilde{t}_i \sigma(x,\theta) - \sum_i \left(\sum_l \theta_{i,l} \tilde{q}_{i,l} - \tilde{t}_i \right) \tilde{\nabla}_i^+ \sigma(x,\theta) \right]. \tag{19}$$

[26] In Online Appendix B.3, we prove several further results about transfers. In particular, we construct "canonical" transfers associated with balanced aggregate excess growth functions, and we characterize all transfer rules with a given aggregate excess growth.

In writing the programs in this manner, we have simply integrated out ω and replaced the terms corresponding to m with q and t, which are the allocation probabilities and expected transfers, respectively.

Because the transfers are allowed to be unbounded, the bounding programs (18) and (19) are not a special case of (LB-G-k) and (UB-P-k). However, it is straightforward to extend the proof of Theorem 1 to cover the case of unbounded transfers. For completeness, we have included a statement and proof of the analogue of Theorem 1 for the optimal auctions problem in Appendix B.2. A difference with the analysis in Section 3 is that it is no longer immediate the upper bound on min potential is bounded. But as we prove shortly, there is a choice of σ that causes transfers to drop out and be replaced by a simpler object, the *aggregate excess growth* (following terminology established in Brooks and Du (2021b)). This reduction is the main advantage of working with unbounded transfers. Moreover, it will be self-evident that the reduced program has a finite value.

To that end, observe that the coefficient on $\tilde{t}_i$ in (19) is

$$\xi_i(x) \equiv \sum_{\theta} \left[\sigma(x,\theta) + \tilde{\nabla}_i^+ \sigma(x,\theta) \right].$$

If σ is such that $\xi_i(x)$ is non-zero for some x, then the designer can take $\tilde{t}_i$ to be a large number with the same sign as $\xi_i(x)$ and make the inner maximum arbitrarily large. Thus, in order for the value of the inner program to be finite, it must be that $\xi_i(x) = 0$. This yields a difference equation for σ, for which the solution is

$$\sum_{\theta} \sigma(x,\theta) = \rho_i(x_i) \sum_{\theta} \sigma(0, x_{-i}, \theta),$$

where

$$\rho_i(x_i) \equiv \left(1 - \frac{1}{k}\right)^{kx_i} \frac{1}{k^{\mathbb{I}_{x_i<k}}}.$$

Iterating over i, and using the fact that $\sum_{x_i} \rho_i(x_i) = 1$, we conclude that

$$\sum_{\theta} \sigma(x,\theta) = \prod_i \rho_i(x_i) \equiv \rho.$$

We have proven the following:

Proposition 1. *The value of the inner program of* (19) *is finite only if the marginal of* σ *on* x *is* ρ.

Remark 3. This result relies on the fact that the designer places positive weight on transfers. It would remain true if, say, the designer's objective were a weighted sum of revenue and social welfare, with the weight on revenue normalized to be one. At a higher level, it is not surprising that the potential-minimizing signals should be independent when the objective is revenue maximization, because of the well-known result that correlation between signals can be exploited to make participation constraints bind (Myerson, 1981; Crémer and McLean, 1988; Luz, 2013). Note that independence of signals is not a general property of potential-minimizing information: in the public expenditure problem studied in Section 4, where the objective is social welfare maximization and transfers have to satisfy ex post budget balance, the potential-minimizing signal distribution is actually correlated.

Remark 4. There is a curious connection between Proposition 1 and the characterization of revenue maximizing mechanisms in the independent private value model due to Myerson (1981). As mentioned above in Section 3.3.3, in that model, all of the local downward incentive constraints and the lowest participation constraint bind. Moreover, in the regular case, the optimal multiplier on the local downward constraint is the inverse hazard rate of the marginal distribution of the signal.[27] (See also discussions in Vohra, 2011; Cai, Devanur, and Weinberg, 2019). Proposition 1 provides a partial converse: If these are the binding constraints at the optimum, then signals must be independent, with the multiplier being the inverse hazard rate of the marginal.

Continuing with our analysis, in light of Proposition 1, any optimal solution σ of (19) must be such that the marginal on x is ρ, and $\tilde{t}_i$ drops out of the inner program. Moreover, the optimal $\tilde{q}$ will simply allocate good l to whichever bidder maximizes the informational virtual objective, $-\sum_\theta \theta_{i,l}\tilde{\nabla}_i^+\sigma(x,\theta)$, as long as the maximum is positive, and otherwise the good will not be sold. We can therefore rewrite (19) as the linear program

$$\begin{aligned} &\min_{\substack{\sigma:X(k)\times\Theta\to\mathbb{R}_+,\\ \gamma:X(k)\to\mathbb{R}_+^L}} \sum_{x,l}\gamma_l(x) \\ \text{s.t. } &\gamma_l(x) \geqslant -\sum_\theta \theta_{i,l}\tilde{\nabla}_i^+\sigma(x,\theta)\ \forall x,i,l \\ &\sum_\theta \sigma(x,\theta) = \rho(x)\ \forall x \\ &\sum_x \sigma(x,\theta) = \mu(\theta)\ \forall \theta, \end{aligned} \tag{20}$$

It is evident that (20) has a finite value, and therefore so does (19).

Remark 5. There is a tight connection between the informational virtual objective and the classical virtual value. Let

$$v_{i,l}(x) \equiv \frac{1}{\rho(x)}\sum_\theta \theta_{i,l}\sigma(x,\theta)$$

denote agent i's interim value for good l. Hence, for $x_i < k - 1/k$,

$$\begin{aligned} -\sum_\theta \theta_{i,l}\tilde{\nabla}_i^+\sigma(x,\theta) &= k\left[v_{i,l}(x)\rho(x) - v_{i,l}(x_i+1/k,x_{-i})\rho(x_i+1/k,x_{-i})\right] \\ &= \left[v_{i,l}(x) - \frac{k-1}{k}\tilde{\nabla}_i^+ v_{i,l}(x)\right]\rho(x). \end{aligned}$$

[27]For this result, it is necessary to formulate the truthtelling constraint in ex ante probability units. Otherwise, the optimal multiplier is the upward cumulative distribution.

The term in brackets is no more than agent i's virtual value for good l. This is a discrete analogue of the virtual value derived in Bulow and Klemperer (1996) in a continuous and differentiable independent-signal interdependent-values model:

$$v_{i,l}(x) - \frac{1 - F_i(x_i)}{f_i(x_i)} \nabla_i v_{i,l}(x),$$

where F_i is the cumulative distribution of i's signal and f_i is the density. This formula reduces to that of Myerson (1981) in the special case of private values and a single good, and under the normalization that $v_i(x) = x_i$. Note that for the distribution ρ_i, the discrete inverse hazard rate is precisely $(k-1)/k$. Thus, in the optimal auctions problem, the upper bounding program reduces to choosing an independent-signal information structure to minimize the classical Myersonian upper bound on revenue, i.e., the expected highest virtual value for each good.

We now complete the task of solving out transfers. The dual of (20) is:

$$\begin{aligned} \max_{\substack{q:X(k)\to\mathbb{R}_+^{NL},\\ \Xi:X(k)\to\mathbb{R},\ \lambda:\Theta\to\mathbb{R}}} \quad & \sum_\theta \mu(\theta)\lambda(\theta) + \sum_x \rho(x)\Xi(x) \\ \text{s.t. } \lambda(\theta) + \Xi(x) &\leqslant \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}(x) \ \forall \theta, x \\ \sum_i q_{i,l}(x) &\leqslant 1 \ \forall x, l. \end{aligned} \tag{21}$$

The last step is to manipulate (18) into a form that is comparable to (21). Let

$$\Xi(x) \equiv \sum_i \left[\nabla_i^+ t_i(x) - t_i(x)\right]. \tag{22}$$

Clearly, Ξ is the only feature of the transfer that matters for the value of (18). So, we could substitute Ξ for the transfers, but we have to restrict ourselves to Ξ that satisfy (22) for *some* participation secure transfer rule. The following lemma reformulates that constraint without the existential quantifier, where $\nabla^+ \cdot t(x) = \sum_i \nabla_i^+ t_i(x)$ (cf. equation (1)):[28]

Lemma 1. *Given $\Xi : X(k) \to \mathbb{R}$, there exists a $t : X(k) \to \mathbb{R}^N$ that solves*

$$\Xi(x) = \nabla^+ \cdot t(x) - \Sigma t(x) \ \forall x; \tag{23}$$

$$t_i(0, x_{-i}) = 0 \ \forall i, x_{-i} \tag{24}$$

if and only if

$$\sum_x \rho(x)\Xi(x) = 0. \tag{25}$$

[28]We defined ∇_i^+ to be scaled by $(k-1)$ rather than k so that the distribution that appears in the balance condition (25) is precisely ρ, which also appears independently in Proposition 1. By defining ∇_i^+ so that these two distributions are the same, the only remaining difference between the upper and lower bounds is the direction of discrete derivatives on the allocation rule.

Proof. By Fredholm's alternative, there exists a t that solves (23) and (24) if and only if there does not exist a $\tilde{\rho}$ such that

$$\sum_x \tilde{\rho}(x)\Xi(x) \neq 0 \tag{26}$$

$$\tilde{\rho}(x) = \begin{cases} \frac{k-1}{k}\tilde{\rho}(x_i - 1/k, x_{-i}) & \text{if } 0 < x_i < k; \\ (k-1)\tilde{\rho}(k - 1/k, x_{-i}) & \text{if } x_i = k. \end{cases}$$

Thus, the choice of $\tilde{\rho}(0)$ pins down the rest of $\tilde{\rho}$, and in fact

$$\tilde{\rho}(x) = \rho(x)\frac{\tilde{\rho}(0)}{\rho(0)}.$$

As a result, (26) holds if and only if $\sum_x \rho(x)\Xi(x) \neq 0$, and therefore (23)–(24) has a solution if and only if Ξ satisfies (25). □

We refer to aggregate excess growth functions that satisfy (25) as *balanced*. This condition appeared in earlier work on the optimal auctions problem (Brooks and Du, 2021b). We comment more on the connection in Online Appendix B.3.

By Lemma 1, program (18) is equivalent to the following:

$$\begin{aligned} \max_{\substack{q:X(k)\to\mathbb{R}_+^{NL}, \\ \Xi:X(k)\to\mathbb{R},\ \lambda:\Theta\to\mathbb{R}}} \quad & \sum_\theta \mu(\theta)\lambda(\theta) \\ \text{s.t. } \lambda(\theta) + \Xi(x) &\leqslant \sum_{i,l} \theta_{i,l}\nabla_i^+ q_{i,l}(x)\ \forall \theta, x \\ \sum_x \rho(x)\Xi(x) &= 0 \\ \sum_i q_{i,l}(x) &\leqslant 1\ \forall x, l. \end{aligned} \tag{27}$$

Here we have substituted Ξ according to (22), and added the balance constraint (25). Alternatively, we can just add the expectation of Ξ to the objective, to obtain a program that is still equivalent to (18) but is closer in form to (21):

$$\begin{aligned} \max_{\substack{q:X(k)\to\mathbb{R}_+^{NL}, \\ \Xi:X(k)\to\mathbb{R},\lambda:\Theta\to\mathbb{R}}} \quad & \sum_\theta \mu(\theta)\lambda(\theta) + \sum_x \rho(x)\Xi(x) \\ \text{s.t. } \lambda(\theta) + \Xi(x) &\leqslant \sum_{i,l} \theta_{i,l}\nabla_i^+ q_{i,l}(x)\ \forall \theta, x \\ \sum_i q_{i,l}(x) &\leqslant 1\ \forall x, l. \end{aligned} \tag{28}$$

To see why adding the expectation of Ξ under ρ to the objective is equivalent to imposing the balance condition (25), note that if (λ, Ξ) is feasible for (27), then it is also feasible for

(28) with the same objective value. On the other hand, if we take (λ, Ξ) that are feasible for (28), then we can define

$$\widehat{\lambda}(\theta) \equiv \lambda(\theta) + \sum_{x'} \rho(x')\Xi(x'), \quad \widehat{\Xi}(x) \equiv \Xi(x) - \sum_{x'} \rho(x')\Xi(x').$$

Then $\widehat{\Xi}$ is balanced, so that $(\widehat{\lambda}, \widehat{\Xi})$ is feasible for (27), and

$$\sum_{\theta} \mu(\theta)\widehat{\lambda}(\theta) = \sum_{\theta} \mu(\theta)\lambda(\theta) + \sum_{x} \rho(x)\Xi(x).$$

We have proven the following:

Proposition 2. *The program* (18) *has the same value as* (28)*. The program* (19) *has the same value as* (21)*.*

Thus, the task of showing the bounds are tight is reduced to showing that (21) and (28) have approximately the same value, in the limit as k becomes large.

5.2 Tightness of the bounds

We now arrive at the main result for this section:

Theorem 2. *For the optimal auctions problem,*

$$\lim_{k\to\infty} W(19) = W(\text{MIN-P}) = W(\text{MAX-G}) = \lim_{k\to\infty} W(18).$$

The formal proof appears in Appendix A.2. Here we sketch the argument. In light of Proposition 2, it suffices to show that (21) has the same value as (28), in the limit as k goes to infinity. The only difference between these two programs is the direction of local derivatives. We show that there is an optimal solution (λ, Ξ, q) to (21) that can be manipulated into a feasible solution (λ', Ξ', q') of (28), such that the difference in value is small when k is large. In fact, this is easy to do when q is non-decreasing, in which case we can set $\lambda' = \lambda\frac{k}{k-1}$ and $\Xi' = \Xi\frac{k}{k-1}$ and define, for $0 < x_i < k$,

$$q'_{i,l}(x) = q_{i,l}(x_i - 1/k, x_{-i}). \tag{29}$$

As a result,

$$\nabla_i^- q_{i,l} = \frac{k}{k-1}\nabla_i^+ q'_{i,l},$$

and (setting aside delicate boundary cases) the values of the two solutions in their respective programs will be close as k becomes large. But if q decreases at x, then it could be that the "shifted" allocation q' defined by (29) is infeasible, because $\sum_i q'_{i,l}(x) > 1$. However, as long as the absolute decrease in q is small when k is large, we can simply rescale q' so that it is feasible, without significantly changing the discrete upward derivative.

In fact, we establish an even stronger property: Lemma 4 in Appendix A.2 shows that for every $\epsilon > 0$, we can find a k large enough so that there is an allocation $\tilde{q}$ that is ϵ-optimal for (21) and for which

$$|\tilde{q}_{i,l}(x_i + 1/k, x_{-i}) - \tilde{q}_{i,l}(x)| \leqslant \epsilon$$

for every i, l, and x. Thus, the allocation $\tilde{q}$ is approximately optimal and *almost continuous.*

At a high level, the argument is as follows. Let (λ^*, Ξ^*, q^*) be optimal for (21). Suppose we change the allocation to q, hold fixed (λ^*, q), and partially optimize the value of (21) over Ξ. This gives us a value, denoted $W(q)$, which is clearly concave in q.[29] We can use this fact to "smooth out" q^* to produce the desired $\tilde{q}$.[30]

In particular, given $y \in X(k)$, define

$$q^y(x) = \begin{cases} q^*(x - y) & \text{if } x_i \geqslant y_i \ \forall i; \\ 0 & \text{otherwise.} \end{cases}$$

This is the allocation in which all actions are translated down by the vector y (and the allocation is zero if any of the translated actions are negative). Lemma 4 shows that each q^y is almost optimal, as long as y_i is small relative to k. In particular, if y_i is less than $\sqrt{k}$, then the approximate optimality result holds, as we now explain. Except at the boundaries, the term in $W(q^*)$ that involves $\nabla_i^- q_{i,l}^*(x)$ also appears in the calculation of $W(q^y)$, except that the probability weighting changes from $\rho(x)$ to $\rho(x+y)$. The likelihood ratio is on the order of $(1 - 1/k)^{\sqrt{k}} \approx \exp(-1/\sqrt{k})$, which converges to 1 as k goes to infinity, so that for k large, the contribution of these terms is essentially the same. There are also terms that appear in $W(q^y)$ but have no counterpart in $W(q^*)$, which are when $x_i < y_i$ for some i, so q^y is zero. But these terms have vanishingly small probability weight according to ρ, on the order of $1 - \exp(-1/\sqrt{k})$. Moreover, we show that the optimal λ^* must be bounded uniformly in k. This result crucially relies on the full support hypothesis (which we assumed at the beginning of this section) that $\mu(\theta) > 0$ for all $\theta \in \Theta$. Boundedness of λ^* implies that the optimal $\Xi(x)$ is bounded for regions in which q^y is zero, and hence the contribution of these terms is negligible as well. Finally, there are terms in $W(q^*)$, for $x_i > k - y_i$, which have no counterpart in the translated allocation. These terms may grow on the order of k, but the weight assigned to these terms under ρ is on the order of $\exp(-\sqrt{k})$, so that the overall contribution to $W(q^*)$ is again small when k is large.

Hence, we can define a new solution $\tilde{q}$ to be the arithmetic average of the q^y for $y \in X(k)$ and for which $y_i \leqslant \sqrt{k}$ for all i. By concavity, $W(\tilde{q})$ is at least the minimum $W(q^y)$ across y, which is close to $W(q^*)$. Finally, as long as the number of terms in that are averaged in $\tilde{q}(x)$ grows without bound as k goes to infinity, very few terms in the average will change when we increment x, so that $\tilde{q}$ is almost continuous.

[29] To see that W is concave, suppose (λ^*, Ξ, q) and (λ^*, Ξ', q') are both feasible, with values $W(q)$ and $W(q')$. Then for any $\alpha \in [0,1]$, the mixture $(\lambda^*, \alpha\Xi + (1-\alpha)\Xi', \alpha q + (1-\alpha)q)$ is also feasible and has a value $\alpha W(q) + (1-\alpha)W(q')$, which is a lower bound on the value $W(\alpha q + (1-\alpha)q')$.

[30] A previous version of this paper contained an erroneous proof of the smoothness of optimal solutions to (21). We are grateful to Gabriel Carroll for suggesting this correct proof strategy.

Incidentally, this argument also establishes an upper bound on the rate of convergence of the values of the bounding programs, which is shown to be on the order of $1/\sqrt{k}$. This explained in detail in Proposition 9 in Online Appendix B.4.

5.3 Exchangeable values

An interesting special case is when the prior is *exchangeable across goods*, meaning that if θ' is obtained from θ by permuting agent i's values for the different goods, then both value profiles are equally likely. In this case, it is without loss to restrict attention to mechanisms in which the seller only offers the goods as a grand bundle (meaning that probabilities of being allocated each good is the same), and to information structures in which agents only receive information about the value of the grand bundle. The reason is as follows. Clearly, if the mechanism only offers the grand bundle, then in computing the guarantee, it is without loss to consider information structures that are only informative about the value of the grand bundle. In the other direction, if the prior is exchangeable across goods and the information structure is only informative about the value of the grand bundle, then agents will have the same interim expected value for each good, so it is without loss to restrict attention to mechanisms for which the allocation is the same for all goods, i.e., the mechanism only offers the grand bundle.

In independent and concurrent work, Deb and Roesler (2023) studied informationally robust optimal auctions with a single agent.[31] They also conclude that when the prior is exchangeable across goods, there is a guarantee-maximizing mechanism in which the seller only offers the grand bundle. Online Appendix B.5 contains numerical examples in which the exchangeability condition is violated, and guarantee-maximizing mechanisms offer more than just the grand bundle. We also report simulations for the cases where there is a single good and values are either perfectly correlated or independently distributed.

6 Optimal Auctions with a Known Empirical Distribution

We now apply our theory to the optimal auctions problem when there is a single good $L = 1$, and the empirical distribution of the agents' ex post values for the good is known.[32] This assumption is especially natural in a large market, although for ease of analysis, we will primarily focus on $N = 2$. We set $\Theta = \{(1, 0), (0, 1)\}$, with the two value profiles being equally likely. By Theorem 2, the bounding programs are tight. We will analyze their solutions when k becomes large.

[31] Che and Zhong (2021) study a related model, but rather than fixing the distribution of ex post values, they consider value distributions with fixed mean values for each good, goods are divided into "product groups," and some convex moment is known for the sum of the values within each product group. They similarly find that the maxmin mechanism involves bundling all of the goods within a product group.

[32] The case of known empirical distribution may be contrasted with the common value model studied by Brooks and Du (2021b), where there is uncertainty about the empirical distribution, but no uncertainty about heterogeneity across agents.

6.1 Potential-minimizing information structures

If the agents have no information about their values, no mechanism can guarantee more revenue than 1/2, which is each agent's ex ante expected value for the good. The min potential is therefore less than 1/2.

Moreover, the potential of any information structure $I = (S, \sigma)$ is at least 1/2. To see why, let $v_i(s)$ denote the conditional expectation of θ_i given s, i.e., $v_i(s) = \sum_\theta \theta_i \sigma(s,\theta)/\sum_\theta \sigma(s,\theta)$. Note that $v_1(s) + v_2(s) = 1$ for all s. Now, consider the direct mechanism that allocates the good to whichever agent has $v_i(s) \geqslant 1/2$ (breaking ties arbitrarily when $v_1(s) = v_2(s) = 1/2$), and charge a price of 1/2 to whoever is allocated the good. This mechanism is clearly incentive compatible and individually rational and it generates revenue of 1/2. We therefore conclude that the min potential is exactly 1/2, and no information is a potential minimizer.

6.2 Guarantee-maximizing mechanism

Theorem 2 implies that there are mechanisms with profit guarantees arbitrarily close to 1/2. Constructing such mechanisms turns out to be a subtle matter, as we now explain. A natural guess is to simply post a price of $p = 1/2 - \epsilon$ for ϵ small (so as to break ties in favor of buying). Such mechanisms would be approximately optimal under no information. However, the guarantee of such a posted price is actually bounded away from 1/2. To see why, consider the following information structure: $S_i = \{0, 1\}$, signals are conditionally independent, and $s_i = \theta_i$ with probability 3/4 conditional on θ. When ϵ is small, there is an equilibrium of the posted price mechanism with this information structure in which agents purchase if and only if $s_i = 1$.[33] Thus, under this information structure and equilibrium, a sale occurs only if at least one agent has a signal $s_i = 1$, which occurs with probability $13/16 < 1$.

We conclude that posted prices do not maximize the guarantee. Intuitively, what is needed is a mechanism that will aggregate the agents' private information in order to determine who has the higher value and should therefore purchase the good. But rather than proceeding from first principles, we will simply construct feasible solutions to (18) with value close to 1/2.

To that end, let $\lambda(v) \equiv 1/2$ for all v. Motivated by simulations of the sort described in Section 4, we guess and verify a solution of the following form. Fix a positive integer m, and define

$$q_i(x) \equiv \begin{cases} 1 & \text{if } x_i > x_j + m; \\ 0 & \text{if } x_i < x_j - m; \\ \frac{x_i - x_j + m}{2m} & \text{otherwise} \end{cases}$$

[33] Conditional on a signal $s_i = 0$ and asking to purchase the good, the posterior probability that $v_i = 1$ is only 1/4, so that the expected utility from buying the good is

$$\frac{1}{4}\left(\frac{3}{4} + \frac{1}{4}\frac{1}{2}\right)(1-p) + \frac{3}{4}\left(\frac{1}{4} + \frac{3}{4}\frac{1}{2}\right)(0-p) = -\frac{1}{8} + \frac{11}{16}\epsilon,$$

which is negative when $\epsilon < 2/11$.

and

$$\Xi(x) \equiv \min_v v \cdot \nabla^+ q(x) - \lambda(v) = \min_{i=1,2} \nabla_i^+ q(x) - 1/2.$$

By construction, (λ, Ξ, q) is feasible for (28). Note that

$$\nabla_i^+ q(x) = \begin{cases} \frac{k-1}{k}\frac{1}{2m} & \text{if } m > x_i - x_j \geqslant -m \text{ and } x_i < k; \\ 0 & \text{otherwise.} \end{cases}$$

Hence,

$$\Xi(x) = -1/2 + \begin{cases} \frac{k-1}{k}\frac{1}{2m} & \text{if } |x_1 - x_2| < m \text{ and } \max(x_1, x_2) < k; \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

The above construction ensures that the strategic virtual objective is equalized across bidders all x.

Now, when k is large, x converges in distribution to independent exponential, so $x_1 - x_2$ converges to Laplace, and $|x_1 - x_2|$ converges to exponential. Thus, when k is large we have

$$\begin{aligned} \sum_{x \in X(k)} \rho(x)\Xi(x) &\approx -1/2 + \frac{1}{2m}\int_{y=0}^{m} \exp(-y)dy \\ &= -1/2 + \frac{1-\exp(-m)}{2m}. \end{aligned}$$

The limit profit guarantee associated with this mechanism is therefore arbitrarily close to $(1 - \exp(-m))/2m$. Using L'Hôpital's rule, we find that

$$\lim_{m \to 0} \frac{1-\exp(-m)}{2m} = \lim_{m \to 0} \frac{\exp(-m)}{2} = \frac{1}{2}.$$

Hence, by first taking k large and then m small, the seller can guarantee profit arbitrarily close to 1/2. In this limit, the good is essentially allocated to whichever agent has the highest action.

Note that for finite k, the function Ξ given by (30) is not balanced. But it is straightforward to modify the solution by setting C equal to the expectation of Ξ, replacing λ and Ξ with $\lambda + C$ and $\Xi - C$, respectively. Then $\Xi - C$ is balanced, and hence by Lemma 1, there exist participation secure transfers t with aggregate excess growth $\Xi - C$, so that $(\lambda + C, q, t)$ is feasible for (18) and has value close to 1/2.

Finally, we relate the optimal Ξ to the guarantee-maximizing transfers. Propositions 7 and 8 in Online Appendix B.3 give a general construction of transfers that induce a given balanced aggregate excess growth. In Online Appendix B.3.4, we show that applying this construction to the present model leads to fairly complicated transfers, given in equation (51). However, by leveraging additional results in the Online Appendix, and Proposition 8 in particular, we constructed an alternative transfer rule that induces the same aggregate excess growth and is considerably simpler. In the limit where we first take $k \to \infty$ and then $m \to 0$, these transfers converge to

$$t_i(x) = \begin{cases} 1/2 & \text{if } x_i > x_j > 0; \\ 0 & \text{if } x_i < x_j \text{ or } x_i = 0; \\ 1/4 & \text{if } x_i = x_j > 0 \text{ or } x_i > x_j = 0. \end{cases}$$

Thus, setting aside ties and boundary cases, as $m \to 0$ the good is allocated to the high bidder for a posted price of 1/2.

The bottom line is that the seller can guarantee revenue of 1/2 with mechanisms that are, in a sense, discrete approximations of the following enriched posted price mechanism: the agents bid non-negative real numbers, the high bidder wins, and the winner pays a posted price of 1/2. The extra actions allow the agents to express intensity of preference in a manner that aggregates private information and determines which bidder has the higher expected value. Importantly though, in the discrete approximations, it is necessary to smooth out the allocation, so that a change in the bid has negligible effect on the allocation when k is large.

6.3 Extensions

The construction can be generalized to priors μ supported on value profiles for which $\theta_1 + \theta_2 = \overline{\theta}$ for some constant $\overline{\theta}$, and the two agents have the same ex ante expected value $\overline{\theta}/2$. We can proceed with the same allocation q as before, $\lambda(\theta) = \overline{\theta}/2$ for all θ,

$$\Xi(x) = \overline{\theta} \min_{i=1,2} \nabla_i^+ q(x) - \overline{\theta}/2 \leqslant \min_{\theta} \theta \cdot \nabla^+ q(x) - \lambda(\theta)$$

for all x, and multiplying the transfers by a factor of $\overline{\theta}$. Our constructed solution (λ, Ξ, q) is still feasible for (28), and the mechanism remains optimal.

The generalization to $N > 2$ is both more interesting and less straightforward. The critical step is to construct an allocation that satisfies

$$\sum_{x \in X(k)} \rho(x) \min_{i=1,\ldots,N} \nabla_i^+ q(x) \approx 1/N.$$

Simulations indicate that such an allocation exists for $N = 3$. If existence of such an allocation can be established theoretically, then it is straightforward to extend the rest of our construction to prove that (19) has value $\overline{\theta}/N$.

7 Conclusion

This paper has developed new tools for the characterization of guarantee-maximizing mechanisms and potential-minimizing information structures. The bounding programs we derived have a natural economic interpretation in terms of the strategic and informational virtual objectives, which adjust the designer's welfare to account for agents' incentives to deviate to nearby actions or to mimic nearby types. We used the bounding programs to construct solutions for public expenditure, bilateral trade, and optimal auctions, and we showed non-constructively that the bounds are tight for optimal auctions with multiple goods and interdependent values. In all of these cases, we conclude that max guarantee equals min potential.

There remain many promising applications which we have not yet explored, in different environments and under different conditions on primitives. We also suspect that our non-constructive tightness result can be extended beyond the optimal auctions problem. In addition, we have gone back and forth between the discrete model for general theory and the continuous model in applications. It would be useful to formulate the bounding problems and their duality directly in the continuum limit. Finally, an ambitious and challenging goal is to incorporate more flexible restrictions on the agents' information into the framework.

We conclude by discussing the interpretation of our results. The guarantee-maximization program literally represents the preferences of a designer who evaluates each mechanism by its minimum welfare across all information structures and equilibria. We do not believe that real-world designers generally exhibit such extreme pessimism and paranoia. At the same time, we suspect that designers in a practical setting may be unable or unwilling to commit to a single information structure and equilibrium as the correct description of behavior, as required by the classical Bayesian mechanism design paradigm. The truth is likely somewhere in between: Designers may know some features of agents' information without being able to give a complete description. Of course, uncertainty about agents' information may be accompanied by distinct concerns about the complexity of the mechanism or the empirical validity of the equilibrium hypothesis. It is beyond our present abilities to incorporate all such concerns into the theory of optimal mechanism design. We can, however, ask which mechanisms are robust to uncertainty about agents' information and strategies in an extreme sense, provided we are still willing to accept the common prior and Bayes Nash equilibrium as an "as-if" description of behavior. Our results show that it is not necessary for the agents to explicitly articulate or communicate all of their private information in order for a mechanism to attain the optimal guarantee, and in that sense, the approach does not unduly strain the credulity of our assumptions.

In our view, the greatest promise of this approach is that it may lead to the discovery of novel mechanisms, such as proportional auctions, proportional cost-sharing mechanisms, and proportional-price trading mechanisms, that are compelling both for their optimal worst-case performance as well as for their simplicity.[34] The guarantee of a mechanism is, in a sense, a measure of how "safe" it is. To be sure, it is just one of many criteria that might be considered in applied mechanism design. For example, one may also wish to consider how the mechanism performs in benchmark environments, such as affiliated values in the auction context. Importantly, these criteria need not conflict: when values are common and the number of agents is large, the maximum guarantee for profit is approximately the entire surplus, so that profit-guarantee-maximizing mechanisms are nearly optimal in all information structures (Du, 2018; Brooks and Du, 2021b); and likewise for the welfare-guarantee-maximizing mechanism in the public expenditure problem when the social value

[34]The worst-case analysis naturally leads to a great deal of structure on information and mechanisms, which we view as being relatively "simple." This judgment is of course subjective.

is large. This will not always be the case, however, and an important task for future work is to evaluate guarantee-maximizing mechanisms on a variety of information structures and under different solution concepts. Such analyses will lead to a more balanced view of the merits and demerits of these new mechanisms, and of the tradeoff between informational robustness and Bayesian optimality.

References

Akerlof, G. A. (1970): "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 488–500.

Bachrach, N., Y.-C. Chen, I. Talgam-Cohen, X. Yang, and W. Zhang (2022): "Distributionally Robust Auction Design," Working paper.

Bergemann, D., B. Brooks, and S. Morris (2016): "Informationally Robust Optimal Auction Design," Tech. rep., Princeton University and the University of Chicago and Yale University, working paper.

——— (2017): "First-Price Auctions with General Information Structures: Implications for Bidding and Revenue," *Econometrica*, 85, 107–143.

——— (2019): "Revenue guarantee equivalence," *The American Economic Review*, 109, 1911–29.

——— (2020): "Countering the winner's curse: Optimal auction design in a common value model," *Theoretical Economics*, 15, 1399–1434.

Bergemann, D. and S. Morris (2005): "Robust mechanism design," *Econometrica*, 1771–1813.

——— (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

Brooks, B. and S. Du (2021a): "Maxmin Auction Design with Known Expected Values," Tech. rep., The University of Chicago and University of California-San Diego, working paper.

——— (2021b): "Optimal auction design with common values: An informationally robust approach," *Econometrica*, 89, 1313–1360.

——— (2023): "Robust Mechanisms for the Financing of Public Goods," Tech. rep., The University of Chicago and University of California-San Diego, working paper.

Bulow, J. and P. Klemperer (1996): "Auctions Versus Negotiations," *The American Economic Review*, 180–194.

Cai, Y., N. R. Devanur, and S. M. Weinberg (2019): "A duality-based unified approach to bayesian mechanism design," *SIAM Journal on Computing*, STOC16–160.

Carroll, G. (2016): "Informationally robust trade and limits to contagion," *Journal of Economic Theory*, 166, 334–361.

Che, E. (2020): "Distributionally Robust Optimal Auction Design under Mean Constraints," *arXiv:1911.07103*.

Che, Y.-K. and W. Zhong (2021): "Robustly-Optimal Mechanism for Selling Multiple Goods," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 314–315.

Chen, Y.-C. and J. Li (2018): "Revisiting the Foundations of Dominant-Strategy Mechanisms," *Journal of Economic Theory*, 178, 294–317.

Chung, K.-S. and J. C. Ely (2007): "Foundations of Dominant-Strategy Mechanisms," *The Review of Economic Studies*, 74, 447–476.

Crémer, J. and R. P. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 1247–1257.

Dasgupta, P., P. Hammond, and E. Maskin (1979): "The implementation of social choice rules: Some general results on incentive compatibility," *The Review of Economic Studies*, 46, 185–216.

Deb, R. and A.-K. Roesler (2023): "Multi-Dimensional Screening: Buyer-Optimal Learning and Informational Robustness," *arXiv preprint arXiv:2105.12304*.

Du, S. (2018): "Robust mechanisms under common valuation," *Econometrica*, 86, 1569–1588.

Güth, W. and M. Hellwig (1986): "The private supply of a public good," *Journal of Economics*, 46, 121–159.

He, W. and J. Li (2022): "Correlation-robust auction design," *Journal of Economic Theory*, 200, 105403.

Jehiel, P., M. Meyer-ter Vehn, B. Moldovanu, and W. R. Zame (2006): "The limits of ex post implementation," *Econometrica*, 74, 585–610.

Luz, V. F. (2013): "Surplus extraction with rich type spaces," *Journal of Economic Theory*, 148, 2749–2762.

Maskin, E. (1999): "Nash Equilibrium and Welfare Optimality," *The Review of Economic Studies*, 66, 23–38.

McAfee, R. P. and P. J. Reny (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395–421.

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

Samuelson, P. A. (1954): "The pure theory of public expenditure," *The Review of Economics and Statistics*, 387–389.

Serrano, R. and R. Vohra (2010): "Multiplicity of mixed equilibria in mechanisms: A unified approach to exact and approximate implementation," *Journal of Mathematical Economics*, 46, 775–785.

Vohra, R. V. (2011): *Mechanism Design: A Linear Programming Approach*, vol. 47, Cambridge University Press.

Yamashita, T. (2016): "Revenue Guarantee in Auction with a (Correlated) Common Prior and Additional Information," Working paper.

Yamashita, T. and S. Zhu (2018): "On the foundations of ex post incentive compatible mechanisms," *American Economic Journal: Microeconomics*.

A Omitted proofs

A.1 Proof of Theorem 1

A.1.1 Summation by parts formula

Lemma 2. *For functions $f : X(k) \to \mathbb{R}$ and $g : X(k) \to \mathbb{R}$,*

$$\sum_x (\nabla_i^- f(x))g(x) = -\sum_x f(x)(\tilde{\nabla}_i^+ g(x)).$$

Proof. Using the definitions, we have

$$\begin{aligned}
\sum_x (\nabla_i^- f(x))g(x) &= \sum_{x_{-i}} \Bigg[kf(0, x_{-i})g(0, x_{-i}) + \sum_{0<x_i<k} k(f(x) - f(x_i - 1/k, x_{-i}))g(x) \\
&\qquad + (f(k, x_{-i}) - f(k - 1/k, x_{-i}))g(k, x_{-i}) \Bigg] \\
&= -\sum_{x_{-i}} \Bigg[-kf(0, x_{-i})g(0, x_{-i}) - \sum_{0<x_i<k} kf(x)g(x) + \sum_{0\leqslant x_i<k-1/k} kf(x)g(x_i + 1/k, x_{-i}) \\
&\qquad - f(k, x_{-i})g(k, x_{-i}) + f(k - 1/k, x_{-i})g(k, x_{-i}) \Bigg] \\
&= -\sum_{x_{-i}} \Bigg[\sum_{0\leqslant x_i<k-1/k} f(x)k(g(x_i + 1/k, x_{-i}) - g(x)) + f(k - 1/k, x_{-i})g(k, x_{-i}) \\
&\qquad - kf(k - 1/k, x_{-i})g(k - 1/k, x_{-i}) - f(k, x_{-i})g(k, x_{-i}) \Bigg]
\end{aligned}$$

$$= -\sum_x f(x)\tilde{\nabla}_i^+ g(x).$$

□

A.1.2 W(MIN-P) is greater than W(MAX-G)

For any $M \in \mathcal{M}^*$, $I \in \mathcal{I}$, and $b \in \mathcal{E}(M, I)$,

$$P(I) = \sup_{M' \in \mathcal{M}^*, b' \in \mathcal{E}(M', I)} W(M', I, b') \geqslant W(M, I, b) \geqslant \inf_{I' \in \mathcal{I}, b' \in \mathcal{E}(M, I')} W(M, I', b') = G(M).$$

and hence $W(\text{MIN-P}) = \inf_{I \in \mathcal{I}} P(I) \geqslant \sup_{M \in \mathcal{M}^*} G(M) = W(\text{MAX-G})$ as desired.

A.1.3 (UB-P-k) is an upper bound on (MIN-P)

For each k, an upper bound on W(MIN-P) is the infimum potential across all information structures of the form $(X(k), \sigma)$ for $\sigma \in \mathcal{I}_k$.

Now, fix $I = (X(k), \sigma)$, $\sigma \in \mathcal{I}_k$. For any $M \in \mathcal{M}^*$ and $b \in \mathcal{E}(M, I)$, it must be that for all i and x_i, agent i's interim payoff given a signal x_i is non-negative. If not, then agent i could obtain a higher payoff by playing any participation secure action with probability one. Thus, participation security implies interim individual rationality. An upper bound on $P(I)$ may therefore be computed by applying the revelation principle and maximizing the designer's payoff over all incentive compatible and individually rational direct mechanisms, i.e.,

$$\max_{m: X(k) \times \Omega \to \mathbb{R}_+} \sum_{x, \theta, \omega} w(\omega, \theta) m(\omega|x) \sigma(x, \theta)$$

$$\text{s.t.} \sum_{x_{-i}, \theta, \omega} u_i(\omega, \theta) \left[m(\omega|x_i, x_{-i}) - m(\omega|x_i', x_{-i})\right] \sigma(x_i, x_{-i}, \theta) \geqslant 0 \ \forall i, x_i, x_i' \tag{31a}$$

$$\sum_{x_{-i}, \theta, \omega} u_i(\omega, \theta) m(\omega|x_i, x_{-i}) \sigma(x_i, x_{-i}, \theta) \geqslant 0 \ \forall i, x_i \tag{31b}$$

$$\sum_{\omega} m(\omega|x) = 1 \ \forall x \tag{31c}$$

This program has a bounded feasible set, and by hypothesis it is non-empty because a participation-secure mechanism exists. By strong duality, this program and its dual have the same optimal value.

Let $\alpha_i(x_i, x_i') \geqslant 0$ be the multiplier on the truthtelling constraint (31a), let $\beta_i(x_i) \geqslant 0$ be the multiplier on individual rationality (31b), and let $\gamma(x)$ be the multiplier on feasibility

(31c). The dual to (31) is

$$
\begin{aligned}
&\min_{\substack{\alpha:X_1(k)\times X_1(k)\to\mathbb{R}^N_+,\ \beta:X_1(k)\to\mathbb{R}^N_+,\\ \gamma:X(k)\to\mathbb{R}}} \sum_x \gamma(x) \\
\text{s.t. } \gamma(x) &\geqslant \sum_\theta w(\omega,\theta)\sigma(x,\theta) \\
&\quad - \sum_{\theta,i,x_i'} u_i(\omega,\theta)\left[\alpha_i(x_i',x_i)\sigma(x_i',x_{-i},\theta) - \alpha_i(x_i,x_i')\sigma(x_i,x_{-i},\theta)\right] \\
&\quad + \sum_{\theta,i} u_i(\omega,\theta)\beta_i(x_i)\sigma(x,\theta)\ \forall x,\omega
\end{aligned}
\tag{32}
$$

We obtain an upper bound on the optimal value of (32) by fixing

$$
\alpha_i(x_i,x_i') = \begin{cases} 1 & \text{if } x_i = x_i' + 1/k = k; \\ k & \text{if } x_i = x_i' + 1/k < k; \\ 0 & \text{otherwise,} \end{cases} \qquad \beta_i(x_i) = \begin{cases} k & \text{if } x_i = 0; \\ 0 & \text{otherwise.} \end{cases}
$$

and optimizing over γ, i.e.,

$$
\min_{\gamma:X(k)\to\mathbb{R}} \sum_x \gamma(x) \text{ s.t. } \gamma(x) \geqslant \sum_\theta \left[w(\omega,\theta)\sigma(x,\theta) - \sum_i u_i(\omega,\theta)\tilde{\nabla}_i^+\sigma(x,\theta)\right]\ \forall x,\omega, \tag{33}
$$

where $\tilde{\nabla}_i^+$ is defined in (7). For any $\sigma \in \mathcal{I}_k$, the value of (33) is an upper bound on (32), which is in turn an upper bound on $P(X(k),\sigma)$, which is in turn an upper bound on W(MIN-P). Minimizing (33) over all $\sigma \in \mathcal{I}_k$ is precisely (UB-P-k).

A.1.4 W(LB-G-k) is a lower bound on W(MAX-G)

For each k, a lower bound on W(MAX-G) is the supremum guarantee over all mechanisms of the form $(X(k), m)$ for $m \in \mathcal{M}_k^0$.

For a fixed $M = (X(k), m)$, $m \in \mathcal{M}_k^0$, we compute $G(M)$ by applying the revelation principle for information design and minimizing welfare over BCE, i.e.,

$$
\min_{\sigma:X(k)\times\Theta\to\mathbb{R}_+} \sum_{x,\theta,\omega} w(\omega,\theta)m(\omega|x)\sigma(x,\theta)
$$

$$
\text{s.t. } \sum_{x_{-i},\theta,\omega} u_i(\omega,\theta)\left[m(\omega|x_i,x_{-i}) - m(\omega|x_i',x_{-i})\right]\sigma(x_i,x_{-i},\theta) \geqslant 0\ \forall i, x_i, x_i' \tag{34a}
$$

$$
\sum_x \sigma(x,\theta) = \mu(\theta)\ \forall\theta \tag{34b}
$$

This program has a feasible set that is bounded and, by Nash's theorem, is non-empty. As a result, by strong duality, it has an optimal value which is equal to the optimal value of its dual. Let $\alpha_i(x_i,x_i') \geqslant 0$ be the multiplier on (34a), and let $\lambda(\theta)$ be the multiplier on

(34b). Then the dual of (34) is

$$\max_{\substack{\alpha:X_1(k)\times X_1(k)\to\mathbb{R}^N_+,\\ \lambda:\Theta\to\mathbb{R}}} \sum_\theta \mu(\theta)\lambda(\theta)$$
$$\text{s.t. } \lambda(\theta) \leqslant \sum_\omega \left[w(\omega,\theta)m(\omega|x) + \sum_{i,x_i'} u_i(\omega,\theta)\alpha_i(x_i,x_i')\left(m(\omega|x_i',x_{-i}) - m(\omega|x_i,x_{-i})\right)\right] \;\forall\theta, x \tag{35}$$

We obtain a lower bound on (35) by fixing

$$\alpha_i(x_i,x_i') = \begin{cases} k-1 & \text{if } x_i = x_i' - 1/k; \\ 0 & \text{otherwise,}\end{cases}$$

and optimizing over λ, i.e.,

$$\max_{\lambda:\Theta\to\mathbb{R}} \sum_\theta \mu(\theta)\lambda(\theta)$$
$$\text{s.t. } \lambda(\theta) \leqslant \sum_\omega \left[w(\omega,\theta)m(\omega|x) + \sum_i u_i(\omega,\theta)\nabla_i^+ m(\omega|x)\right] \;\forall\theta, x, \tag{36}$$

where ∇_i^+ is defined in (1). For any $m \in \mathcal{M}_k^0$, the value of (36) is a lower bound on (35), which is equal to $G(X(k), m)$, which is in turn a lower bound on (MAX-G). Maximizing (36) over all $m \in \mathcal{M}_k^0$ is precisely (LB-G-k). This concludes the proof of Theorem 1.

A.2 Proof of Theorem 2

A.2.1 Boundedness of optimal λ

Let $\overline{\theta} \equiv \max_{i,l,\theta} \theta_{i,l}$.

Lemma 3. *For all k and θ, if (λ^*, Ξ^*, q^*) is an optimal solution to (21) such that Ξ^* satisfies (25), then*

$$|\lambda^*(\theta)| \leqslant \frac{L\overline{\theta}}{\min_{\theta'}\mu(\theta')} \equiv C_\lambda$$

for all θ. Hence, the optimal value of (21) is at most C_λ.

Proof. We first show that $\lambda^*(\theta) \leqslant L\overline{\theta}$. To obtain a contradiction, suppose there exists a θ' such that $\lambda^*(\theta') > L\overline{\theta}$. Consider the program (21) but where we hold fixed $\lambda = \lambda^*$, which has the same optimal value as (20), equal to $\sum_\theta \mu(\theta)\lambda^*(\theta)$. This program has the dual:

$$\min_{\substack{\sigma:X(k)\times\Theta\to\mathbb{R}_+,\\ \gamma:X(k)\to\mathbb{R}^L_+}} \sum_{x,l}\gamma_l(x) - \sum_\theta \lambda^*(\theta)\left(\sum_x \sigma(x,\theta) - \mu(\theta)\right)$$
$$\text{s.t. } \gamma_l(x) \geqslant -\sum_\theta \theta_{i,l}\tilde{\nabla}_i^+\sigma(x,\theta) \;\forall i,x,l \tag{37}$$
$$\sum_\theta \sigma(x,\theta) = \rho(x) \;\forall x.$$

Note that both (21) with fixed λ^* and (37) are feasible.[35] Hence, by the strong duality theorem, these two programs must have the same optimal value, equal to $\sum_\theta \mu(\theta)\lambda^*(\theta)$.

Let $\overline{\sigma}(x,\theta) = \rho(x)\mathbb{I}_{\theta=\theta'}$ and $\overline{\gamma}_l(x) = \rho(x)\max_i \theta'_{i,l}$ (in other words, put probability one on θ', x is distributed according to ρ, and assign good l to whichever agent has the highest value for good l). It is easy to check that $(\overline{\sigma}, \overline{\gamma})$ is feasible for (37) (because $\rho(x) = -\tilde{\nabla}_i^+ \rho(x)$) and the resulting objective is $\sum_l \max_i \theta'_{i,l} - \lambda^*(\theta') + \sum_\theta \mu(\theta)\lambda^*(\theta)$, which is strictly less than (21) because $\lambda^*(\theta') > L\overline{\theta} \geqslant \sum_l \max_i \theta'_{i,l}$, which contradicts weak duality.

Now we show that $\lambda^*(\theta) \geqslant -L\overline{\theta}/\mu(\theta)$ for all θ. Note that $q = \Xi = \lambda = 0$ is always feasible for (21)), so the optimal value $\sum_\theta \mu(\theta)\lambda^*(\theta)$ must be non-negative. Using that and the fact that $\lambda^*(\theta') \leqslant L\overline{\theta}$ for all θ, we have

$$\begin{aligned}\lambda^*(\theta) &= \left(\sum_{\theta'} \mu(\theta')\lambda^*(\theta') - \sum_{\theta' \neq \theta} \mu(\theta')\lambda^*(\theta')\right)/\mu(\theta) \\ &\geqslant -L\overline{\theta}/\mu(\theta),\end{aligned}$$

as desired. □

A.2.2 Continuity

Lemma 4. *Fix a function $h : \mathbb{N} \to \mathbb{N}$ such that $\lim_{k\to\infty} h(k) = \infty$ and $\lim_{k\to\infty} h(k)/k = 0$. Then there exists a function $\epsilon(k)$ such that $\epsilon(k) \to 0$ as $k \to \infty$, and with the following additional property: For any k, there exists a feasible solution (λ, Ξ, q) to (21) with value at least $W(21){-}\epsilon(k)$, that satisfies $|\lambda(\theta)| \leqslant C_\lambda$ for all θ, and*

$$|q_{i,l}(x) - \mathbb{I}_{x_i>0} q_{i,l}(x_i - 1/k, x_{-i})| \leqslant \frac{2}{h(k)+1} \;\forall i, x, l. \tag{38}$$

Proof. Fix an optimal solution (λ^*, Ξ^*, q^*) to (21). Without loss, we may assume that Ξ^* satisfies (25). (If not, we can add and subtract a constant from Ξ^* and λ^* so that (25) is satisfied, and without changing the value of the solution.)

Let $W(q)$ be the value of (21) with fixed (λ^*, q) and under the partially optimal Ξ, i.e.,

$$W(q) \equiv \sum_\theta \mu(\theta)\lambda^*(\theta) + \sum_x \rho(x) \min_\theta \left\{ \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}(x) - \lambda^*(\theta) \right\}.$$

Note that $W(q)$ is concave (as a minimum of linear functions of $\nabla_i^- q_{i,l}$, and as ∇_i^- is a linear operator). Moreover, $W(q^*) = W(21)$.

[35] For (21) with fixed λ^*, we can define $q = 0$ and set $\Xi(x)$ to be the mimimum across θ of

$$\sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}(x) - \lambda^*(\theta).$$

For (37), given an arbitrary choice of σ, we can simply define $\gamma_l(x)$ to be the maximum of the right-hand side across i for each (x, l).

Let $Y = \{0, 1/k, \dots h(k)/k\}^N$. For a $y \in Y$, define q^y by

$$q^y(x) = \begin{cases} q^*(x-y) & \text{if } x_i \geqslant y_i \ \forall i; \\ 0 & \text{otherwise.} \end{cases}$$

Hence, if $x_i < y_i$ for some i, we have $\nabla_i^- q_{i,l}^y(x) = 0$, and if $k > x_i \geqslant y_i$ for all i, we have $\nabla_i^- q_{i,l}^y(x) = \nabla_i^- q_{i,l}^*(x-y)$.[36] We therefore have

$$\begin{aligned} W(q^y) = &\sum_\theta \mu(\theta)\lambda^*(\theta) + \sum_{x:\exists i\, x_i<y_i} \rho(x) \min_\theta \{-\lambda^*(\theta)\} \\ &+ \sum_{x:\forall i\, y_i \leqslant x_i < k} \rho(x) \min_\theta \left\{ \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}^*(x-y) - \lambda^*(\theta) \right\} \\ &+ \sum_{\substack{x:x\geqslant y,\\ \exists i\, x_i = k}} \rho(x) \min_\theta \left\{ \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}^*(x-y) \frac{1}{k^{\mathbb{I}_{x_i=k, y_i>0}}} - \lambda^*(\theta) \right\}. \end{aligned} \tag{39}$$

(The third line of this equation is adjusting for the fact that a boundary case in $\nabla_i^- q_{i,l}^y(x)$ may not be a boundary case in $\nabla_i^- q_{i,l}^*(x-y)$.) Now,

$$\sum_{x:\exists i\, x_i<y_i} \rho(x) \min_\theta \{-\lambda^*(\theta)\} \geqslant -C_\lambda \sum_{x:\exists i\, x_i<y_i} \rho(x) \geqslant -C_\lambda N \sum_{x_i:x_i<y_i} \rho_i(x_i) \geqslant -C_\lambda N \left[1 - \left(1 - \frac{1}{k}\right)^{h(k)} \right].$$

Since $h(k)/k \to 0$, the term in square brackets goes to zero as $k \to \infty$, so that the lower bound goes to zero.[37] Next,

$$\begin{aligned} &\sum_{\substack{x:\forall i\\ y_i \leqslant x_i < k}} \rho(x) \min_\theta \left\{ \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}^*(x-y) - \lambda^*(\theta) \right\} \\ &= \left(1 - \frac{1}{k}\right)^{k\Sigma y} \sum_{x:\forall i\, y_i \leqslant x_i < k} \rho(x-y) \min_\theta \left\{ \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}^*(x-y) - \lambda^*(\theta) \right\} \\ &= \left(1 - \frac{1}{k}\right)^{k\Sigma y} \left(W(q^*) - \sum_\theta \mu(\theta)\lambda^*(\theta) - \sum_{x:\exists i\, x_i \geqslant k - y_i} \rho(x) \min_\theta \left\{ \sum_{i,l} \theta_{i,l} \nabla_i^- q_{i,l}^*(x) - \lambda^*(\theta) \right\} \right) \end{aligned}$$

---

[36] This second case subtly depends on the definition of ∇_i^-. In particular, if $x_i = y_i$, then we have

$$\nabla_i^- q_{i,l}^*(x-y) = k q_{i,l}^*(x-y) = k(q_{i,l}^y(x) - 0) = k\left(q_{i,l}^y(x) - q_{i,l}^y(x_i - 1/k, x_{-i})\right) = \nabla_i^- q_{i,l}^y(x).$$

[37] One way to see this limit is that

$$1 \geqslant \left(1 - \frac{1}{k}\right)^{h(k)} = \left(1 - \frac{h(k)/k}{h(k)}\right)^{h(k)} \geqslant \left(1 - \frac{x}{h(k)}\right)^{h(k)} \to \exp(-x),$$

for any $x > 0$, since $h(k)/k \to 0$. Since $\exp(-x) \to 1$ as $x \to 0$, $\lim_{k\to\infty}(1 - 1/k)^{h(k)} = 1$.

$$\geqslant \left(1-\frac{1}{k}\right)^{k\Sigma y}\left(W(q^*)-\sum_\theta \mu(\theta)\lambda^*(\theta) - N\left(1-\frac{1}{k}\right)^{k^2-h(k)}(kNL\overline{\theta}+C_\lambda)\right)$$

$$=W(q^*)-\sum_\theta \mu(\theta)\lambda^*(\theta)-\left(1-\left(1-\frac{1}{k}\right)^{k\Sigma y}\right)\left(W(q^*)-\sum_\theta \mu(\theta)\lambda^*(\theta)\right)$$

$$-\left(1-\frac{1}{k}\right)^{k\Sigma y} N\left(1-\frac{1}{k}\right)^{k^2-h(k)}(kNL\overline{\theta}+C_\lambda)$$

$$\geqslant W(q^*)-\sum_\theta \mu(\theta)\lambda^*(\theta)-\left(1-\left(1-\frac{1}{k}\right)^{Nh(k)}\right)2C_\lambda - N\left(1-\frac{1}{k}\right)^{k^2-h(k)}(kNL\overline{\theta}+C_\lambda).$$

The last line uses the result of Lemma 3 that $|W(q^*)| \leqslant C_\lambda$ and $|\lambda^*(\theta)| \leqslant C_\lambda$. The third term above goes to zero as $k \to \infty$ (see footnote 37), and the fourth term goes to zero as well, since $(1-1/k)^{k^2}k$ converges to zero.[38]

Finally, since $\rho_i(k) = \left(1-\frac{1}{k}\right)^{k^2}$, the third line of (39) is at least $-N\left(1-\frac{1}{k}\right)^{k^2}(kNL\overline{\theta}+C_\lambda)$, which again converges to zero.

We have proven that for each $y \in Y$, $W(q^y) \geqslant W(q^*) - \epsilon(k)$, where

$$\begin{aligned}\epsilon(k) \equiv\ & C_\lambda N\left[1-\left(1-\frac{1}{k}\right)^{h(k)}\right] + N\left(1-\frac{1}{k}\right)^{k^2}(kNL\overline{\theta}+C_\lambda) \\ &+\left(1-\left(1-\frac{1}{k}\right)^{Nh(k)}\right)2C_\lambda + N\left(1-\frac{1}{k}\right)^{k^2-h(k)}(kNL\overline{\theta}+C_\lambda),\end{aligned} \tag{40}$$

and that $\epsilon(k) \to 0$ as $k \to \infty$.

Now, let $\tilde{q}(x) \equiv \frac{1}{|Y|}\sum_{y\in Y} q^y(x)$. By concavity of W, we have that

$$W(\tilde{q}) \geqslant \frac{1}{|Y|}\sum_{y\in Y} W(q^y) \geqslant W(q^*) - \epsilon(k).$$

Moreover, for each x_{-i}, we have

$$\tilde{q}_{i,l}(x) - \mathbb{I}_{x_i>0}\tilde{q}_{i,l}(x_i - 1/k, x_{-i})$$

[38] To see this, first note that for $z \in (-\infty, 0)$, we have $\log(1/(1-z)) \geqslant z$. This follows because the two expressions are equal when $z = 0$, the derivative of the left-hand side with respect to z is $1/(1-z) < 1$, which is the derivative of the right-hand side. As a result,

$$\begin{aligned}\frac{d}{dl}\left(1+\frac{x}{l}\right)^l &= \left(1+\frac{x}{l}\right)^l \frac{d}{dl}\left[\log\left(1+\frac{x}{l}\right)l\right] \\ &= \left(1+\frac{x}{l}\right)^l\left[\log\left(\frac{1}{1-\frac{x}{l+x}}\right) - \frac{x}{l+x}\right] \geqslant 0,\end{aligned}$$

as long as $l \geqslant -x$. Hence, $(1+x/l)^l$ is less than its limit as $l \to \infty$, which is $\exp(x)$, so that

$$\left(1-\frac{1}{k}\right)^{k^2} k = \left(1-\frac{k}{k^2}\right)^{k^2} k \leqslant \exp(-k)k,$$

which converges to zero as $k \to \infty$.

$$= \frac{1}{|Y|} \sum_{y \in Y} \left(q_i^y(x) - \mathbb{I}_{x_i > 0} q_{i,l}^y(x_i - 1/k, x_{-i}) \right)$$
$$= \frac{1}{|Y|} \sum_{y_{-i} \in \{0, 1/k, \ldots, h(k)/k\}^{N-1}} \left(q_i^{(0, y_{-i})}(x) - \mathbb{I}_{x_i > 0} q_{i,l}^{(h(k)/l, y_{-i})}(x_i - 1/k, x_{-i}) \right).$$

We therefore have $|\tilde{q}_{i,l}(x) - \mathbb{I}_{x_i > 0} \tilde{q}_{i,l}(x_i - 1/k, x_{-i})| \leqslant 2/(h(k)+1)$, as desired. □

A.2.3 Shifting

We now complete the proof of Theorem 2:

Proof of Theorem 2. Let $h(k)$ satisfy the hypotheses of Lemma 4 (for example, $h(k)$ can be the smallest integer larger than $\sqrt{k}$). Hence, for every k, there exists a feasible solution (λ, Ξ, q) to (21) that is within $\epsilon(k)$ of being optimal, satisfies (38) and $|\lambda(\theta)| \leqslant C_\lambda$. To simplify expressions, we define $\tilde{\epsilon}(k) = 2/(h(k)+1)$.

We modify (λ, Ξ, q) to obtain a feasible solution for (28). Define

$$q'_{i,l}(x) = \begin{cases} \frac{q_{i,l}(x_i - 1/k, x_{-i})}{1 + N\tilde{\epsilon}(k)} & \text{if } 0 < x_i < k; \\ 0 & \text{if } x_i = 0 \text{ or } x_i = k; \end{cases}$$
$$\lambda'(\theta) = \frac{k-1}{k(1+N\tilde{\epsilon}(k))} \lambda(\theta) \quad \forall \theta;$$
$$\Xi'(x) = \begin{cases} \frac{k-1}{k(1+N\tilde{\epsilon}(k))} \Xi(x) & \text{if } x \notin \partial X(k); \\ -(k-1)NL\overline{\theta} - \max_\theta \lambda'(\theta) & \text{if } x \in \partial X(k), \end{cases}$$

where $\partial X(k) = \{x \in X(k) \mid x_i \geqslant k - 1/k \text{ for some } i\}$.

We claim that (λ', Ξ', q') is feasible for (28): First, the constraint on $\lambda'(\theta) + \Xi'(x)$ holds for all θ and $x \in \partial X(k)$ because

$$\Xi'(x) = -(k-1)NL\overline{\theta} - \max_\theta \lambda'(\theta) \leqslant \sum_{i,l} \theta_{i,l} \nabla_i^+ q'_l(x) - \lambda'(\theta) \ \forall \theta;$$

it also holds for $x \notin \partial X(k)$ because $\nabla_i^+ q'_l(x) = \frac{k-1}{k(1+N\tilde{\epsilon}(k))} \nabla_i^- q_l(x)$, $\Xi'(x) = \frac{k-1}{k(1+N\tilde{\epsilon}(k))} \Xi(x)$, $\lambda'(\theta) = \frac{k-1}{k(1+N\tilde{\epsilon}(k))} \lambda(\theta)$, and $\Xi(x) + \lambda(\theta) \leqslant \sum_{i,l} \theta_{i,l} \nabla_i^- q_l(x)$. Also, q' is feasible, as

$$\sum_i q'_{i,l}(x) \leqslant \sum_i \frac{q_{i,l}(x_i - 1/k, x_{-i})}{1 + N\tilde{\epsilon}(k)} \mathbb{I}_{0 < x_i < k} \leqslant \sum_i \frac{q_{i,l}(x) + \tilde{\epsilon}(k)}{1 + N\tilde{\epsilon}(k)} \leqslant 1.$$

Finally, using Lemma 4 and defining $\epsilon(k)$ as in (40), we know that the difference between the optimal value of (21) and the value of (28) under (λ', Ξ', q') is at most

$$\epsilon(k) + \sum_x \rho(x)(\Xi(x) - \Xi'(x)) + \sum_\theta \mu(\theta)(\lambda(\theta) - \lambda'(\theta))$$
$$= \epsilon(k) + \left(1 - \frac{k-1}{k(1+N\tilde{\epsilon}(k))}\right) \left[\sum_x \rho(x)\Xi(x) + \sum_\theta \mu(\theta)\lambda(\theta) \right] + \sum_{x \in \partial X(k)} \rho(x) \left(\frac{k-1}{k(1+N\tilde{\epsilon}(k))} \Xi(x) - \Xi'(x) \right)$$

$$
\begin{aligned}
&\leqslant \epsilon(k) + \left(1 - \frac{k-1}{k(1+N\tilde{\epsilon}(k))}\right) C_\lambda \\
&\quad + \sum_{x \in \partial X(k)} \rho(x) \left(\frac{k-1}{k(1+N\tilde{\epsilon}(k))} \Xi(x) + (k-1)NL\overline{\theta} + \max_\theta \frac{k-1}{k(1+N\tilde{\epsilon}(k))} \lambda(\theta) \right) \\
&\leqslant \epsilon(k) + \left(1 - \frac{k-1}{k(1+N\tilde{\epsilon}(k))}\right) C_\lambda + N(1-1/k)^{k^2-1} \left(\frac{k-1}{k(1+N\tilde{\epsilon}(k))} kNL\overline{\theta} + (k-1)NL\overline{\theta} \right).
\end{aligned}
$$

In the first inequality, we have used the result of Lemma 3 that the value of (21) is at most C_λ. In the last inequality, we use the fact that $\rho(\partial X(k)) \leqslant N(1-1/k)^{k^2-1}$ and $\Xi(x) + \lambda(\theta) \leqslant \sum_{i,l} \theta_{i,l} \nabla_i^- q_l(x) \leqslant kNL\overline{\theta}$. The last line vanishes as $k \to \infty$ because $\epsilon(k) \to 0$, $\tilde{\epsilon}(k) \to 0$ and $(1-1/k)^{k^2-1} k \to 0$.

Thus, we conclude that the optimal value of (28) is at least that of (21), minus a term that converges to zero as k goes to infinity. By Proposition 2, (28) has the same value as (18) Moreover, by strong duality, (21) has the same value as (20), which by Proposition 2 has the same value as (19). Hence, the value of (18) is at least the value of (19), minus a term that goes to zero as k goes to infinity. Moreover, by Theorem 3 in Online Appendix B.2 (the analogue of Theorem 1 for the optimal auctions problem), for all k, the value of (19) is at least the minimum potential, which is greater than the max guarantee, which is greater than the value of (18). We conclude that (18) and (19) have the same value in the limit as $k \to \infty$, as desired. □