University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 56

A Geometric Approach to Mechanism Design

Jacob K. Goeree and Alexey Kushnir

December 2011

A Geometric Approach to Mechanism Design

Jacob K. Goeree and Alexey Kushnir*

December 20, 2011

Abstract

An important result in convex analysis is the duality between a closed convex set and its support function. We exploit this duality to develop a novel geometric approach to mechanism design. For a general class of social choice problems we characterize the feasible set, which is closed and convex, and its support function. We next provide a geometric interpretation of incentive compatibility and refine the support function to include incentive constraints using arguments from majorization theory. The optimal mechanism can subsequently be derived from the support function using Hotelling's lemma.

We first assume that values are linear in types and types are independent, private, and one-dimensional. For this environment we provide a simple geometric proof that Bayesian and dominant strategy implementation are equivalent by showing that the feasible sets that remain after imposing either type of incentive constraints coincide. Furthermore, we derive the optimal mechanism for *any* social choice problem and *any* linear objective, including revenue and surplus maximization. As an illustration, we determine the optimal multi-unit auction for a class of value functions that exhibit decreasing marginal valuations. Other types of constraints, such as capacity constraints and budget balancedness, can be interpreted geometrically as well, which facilitates a unified approach to a range of social choice problems, including auctions, bargaining, and public goods provision.

We discuss how our geometric approach extends to environments with value interdependencies, non-linear valuations, and correlated or multi-dimensional types. Specifically, we illustrate that with interdependent valuations the equivalence between Bayesian and dominant strategy implementation breaks down, and our approach naturally produces the second-best outcomes for both types of incentive constraints.

Keywords: *convex sets, support functions, majorization, Hotelling's lemma, mechanism design, revenue equivalence, BIC-DIC equivalence, multi-unit auctions, bargaining, public goods provision, capacity constraints, budget balance, interdependent values, second best efficiency*

*Chair for Organizational Design, Department of Economics, University of Zürich, Blümlisalpstrasse 10, CH-8006, Zürich, Switzerland. We gratefully acknowledge financial support from the European Research Council (ERC Advanced Investigator Grant, ESEI-249433). We thank Yuanchuan Lien, Konrad Mierendorff, and Phil Reny as well as seminar participants at SHUFE, HKUST, and Humboldt University for useful suggestions.

1. Introduction

Mechanism design is the science of creating optimal social systems by maximizing a well-defined social welfare function taking into account resource constraints and participants' incentives and hidden information. It provides a framework to address social engineering questions like "what auction format assigns goods most efficiently or yields the highest seller revenue," "when should a public project such as building a highway be undertaken," and "which trading rules maximize the gains from trade?" The difficulty in answering these questions stems from the fact that the designer, or public authority, typically does not possess detailed information about the relevant parameters, e.g. bidders' valuations for the goods for sale or voters' preferences for the public project. A well-designed mechanism should therefore both truthfully elicit participants' private information and implement the corresponding social optimum.

Hurwicz (1960) was among the first to recognize the prevalence and importance of economically relevant information that is dispersed in the population.[1] He introduced a formal model of communication where agents send messages to a central planner who selects an outcome based on a pre-specified rule. Hurwicz (1972) also introduced the key notion of *incentive compatibility*, which emphasizes the need for collecting agents' private information in a manner that is coherent with their incentives. The study of incentive compatible mechanisms was significantly simplified through the observation of the *revelation principle* by Gibbard (1973) and subsequent extensions to incomplete information environments by Dasgupta, Hammond, and Maskin (1979) and Myerson (1979). This principle implies that general mechanisms or institutions can be analyzed through equivalent direct revelation mechanisms, where participants' only form of communication or action is the revelation of their private information.

Notwithstanding this simplification, the constraints imposed by incentive compatibility are generally treated separately from other more basic constraints, such as resource constraints. As a result, mechanism design theory appears to have developed quite differently from classical approaches to consumer and producer choice theory despite some obvious parallels. For example, in producer choice theory, the firm also maximizes a well-defined objective, its profit, over a feasible production set that reflects its resource constraints. A well-known result is that a firm's optimal production plan follows by evaluating the gradient of the profit function at output and factor prices – Hotelling's lemma. One contribution of this paper is to draw a parallel between classical choice theory and mechanism design by showing how the revenue-maximizing or surplus-maximizing mechanism can be derived using standard micro-economics tools.

[1] Early contributors include Hayek (1945) who contemplated the feasibility of a centralized socialist economy.

Our approach is geometric in nature and utilizes convexity of the set of feasible, incentive compatible outcomes. In particular, starting from the basic feasible set that is the product of probability simplices, we determine the set that remains after imposing incentive constraints. We do so by providing a geometric interpretation of the incentive compatibility constraints, which puts them on an equal footing with the basic resource constraints. Exploiting convexity of the resulting set, we subsequently derive the optimal mechanism using Hotelling's lemma.

The challenge lies in keeping track of the (high-dimensional) set that remains after feasibility and incentive compatibility constraints are imposed. To this end we employ techniques from convex analysis, a subfield of mathematics that studies properties of convex sets and functions. A key result in convex analysis is the duality between a closed convex set and its support function, which is convex and homogeneous of degree one (e.g. a firm's profit function). Conversely, any convex function that is homogeneous of degree one defines a convex set. We exploit this duality to derive the support function of the basic feasible set for a general class of social choice problems. Furthermore, support functions possess convenient algebraic properties that facilitate the description of the union, sum, and intersection of convex sets. These algebraic properties allow us to refine the support function to include incentive constraints. Using arguments from majorization theory, we show that this approach naturally generates the "ironing" procedure first described by Mussa and Rosen (1978) and Myerson (1981).

A major question in mechanism design is whether dominant strategy incentive compatibility is more stringent than Bayesian incentive compatibility. For example, does requiring dominant strategy incentive compatibility limit a seller's revenue or overall welfare? A recent contribution by Manelli and Vincent (2010) shows this is not true for the special case of single-unit, private-value auctions: Bayesian incentive compatibility (BIC) and dominant strategy incentive compatibility (DIC) are equivalent in this setting. Goeree and Kushnir (2011) extend this BIC-DIC equivalence result to a broad class of social choice problems by generalizing a theorem due to Gutmann et al. (1991), which was introduced to the economics literature by Gershkov, Moldovanu, and Shi (2011).

In this paper we take a new perspective on the issue, one that fits with our geometric approach. What matters to agents at the time they make their decisions is how BIC and DIC constraints compare at the interim stage, i.e. when agents know only their own types and the distributions of others' types. We first show that each ex post DIC constraint can be represented by a vector in some high-dimensional space and then study how this vector transforms under the linear transformation (of taking expectations over others' types) that represents going from the ex post to the interim stage. We demonstrate that, at the interim stage, the projected DIC constraints coincide with the BIC constraints. We make these arguments precise by proving that the support functions for both types of incentive constraints are identical.

Importantly, support functions provide a useful tool when maximizing some linear social objective over the set of feasible, incentive compatible outcomes. Both revenue maximization and surplus maximization fit the linear framework and, as a result, the revenue-maximizing and surplus-maximizing mechanisms follow from the support function by applying Hotelling's lemma. More generally, we determine the optimal mechanism for *any* social choice problem and *any* linear objective and show that the resulting mechanism is dominant strategy incentive compatible and ex post individually rational. We illustrate the power of our approach by deriving the optimal multi-unit auction for a class of value functions that exhibit decreasing marginal valuations, a result that is new to the auction literature.

The geometric interpretation of the incentive compatibility constraints extends to other types of constraints, which allows us to revisit and unify a number of important applications of mechanism design. For instance, in auctions with many items for sale, bidders typically face some budget or capacity constraints. We show that such constraints can easily be incorporated into the support function from which the second-best mechanism follows using Hotelling's lemma. Moreover, we demonstrate that budget balancedness, which is a natural requirement in bargaining and public goods provision, can be dealt with in a similar manner.

Finally, we show that our methods apply beyond the main model of the paper, which assumes that values are linear in types and that types are independent, private, and one-dimensional. For example, we illustrate that with interdependent values the equivalence between Bayesian and dominant strategy incentive compatibility no longer holds. Nevertheless, our support function approach naturally extends to this setting and produces the second-best mechanisms for both types of incentive constraints.

This paper is organized as follows. In section 2 we describe the basic duality result from convex analysis and list several other useful facts.[2] Section 3 first describes the support functions for the ex post and interim feasible sets and then incorporates incentive constraints. We prove BIC-DIC equivalence and revenue equivalence and derive the optimal dominant strategy mechanism for a broad class of social choice problems. In section 4 we apply our geometric procedure to other types of constraints, which naturally occur in multi-unit auctions, bargaining, and public goods provision. Section 5 discusses extensions that allow for value interdependencies, non-linear valuations, and correlated or multi-dimensional types. Section 6 concludes and the Appendix contains the proofs.

[2]Proofs of these facts can be found in *Convex Analysis* by Rockafellar (1997).

2. Preliminaries from Convex Analysis

An important concept in convex analysis is the duality between a closed convex set $C \subset \mathbb{R}^n$ and its *support function* $\mathcal{S}^C : \mathbb{R}^n \to \mathbb{R}$, which is defined as

$$\mathcal{S}^C(\mathbf{w}) \;=\; \sup\{\mathbf{v}\cdot\mathbf{w} \,|\, \mathbf{v} \in C\}$$

with $\mathbf{v}\cdot\mathbf{w} = \sum_{i=1}^n v_i w_i$ the usual inner product. The support function is homogeneous of degree 1, i.e. $\mathcal{S}^C(\lambda\mathbf{w}) = \lambda\mathcal{S}^C(\mathbf{w})$ for any $\lambda \geq 0$, and convex, i.e. $\mathcal{S}^C(\alpha\mathbf{w}_1 + (1-\alpha)\mathbf{w}_2) \leq \alpha\mathcal{S}^C(\mathbf{w}_1) + (1-\alpha)\mathcal{S}^C(\mathbf{w}_2)$ for any $\alpha \in [0,1]$. Conversely, any lower semi-continuous function defined over $\mathbb{R}^n$ that is convex and homogeneous of degree 1 is the support function of a closed convex set, which can be defined as the intersection of half spaces

$$C \;=\; \big\{\mathbf{v} \in \mathbb{R}^n \,|\, \mathbf{v}\cdot\mathbf{w} \;\leq\; \mathcal{S}^C(\mathbf{w}) \;\forall\,\mathbf{w} \in \mathbb{R}^n\big\}$$

To illustrate, consider the two-dimensional simplex $C = \{(v_1, v_2) \,|\, v_1 \geq 0, v_2 \geq 0, v_1 + v_2 \leq 1\}$ shown in the left panel of Figure 1. In this panel, the blue arrows represent arbitrary vectors $\mathbf{w} \in \mathbb{R}^2$ and the label next to the arrow shows the outcome of the maximization problem $\sup\{\mathbf{v}\cdot\mathbf{w} \,|\, \mathbf{v} \in C\}$. It is readily verified that the support function for the two-dimensional simplex can be summarized as

$$\mathcal{S}^C(\mathbf{w}) \;=\; \max(0, w_1, w_2)$$

In turn, the two-dimensional simplex can be recovered from the support function by considering, for each $\mathbf{w} \in \mathbb{R}^2$, the inequality

$$\mathbf{v}\cdot\mathbf{w} \;\leq\; \mathcal{S}^C(\mathbf{w})$$

This inequality defines a half space of possible $\mathbf{v} \in \mathbb{R}^2$ for each $\mathbf{w} \in \mathbb{R}^2$. In the right panel of Figure 1 these half spaces are bounded by the green lines and their intersection reproduces the two-dimensional simplex. It is straightforward to generalize the example to n dimensions.

Fact 1. *For the n-dimensional simplex*

$$C = \{\mathbf{v} \in \mathbb{R}^n_+ \,|\, \sum_{i=1}^n v_i \;\leq\; 1\}$$

the support function is given by $\mathcal{S}^C(\mathbf{w}) = \max_i(0, w_i)$ for $\mathbf{w} \in \mathbb{R}^n$.

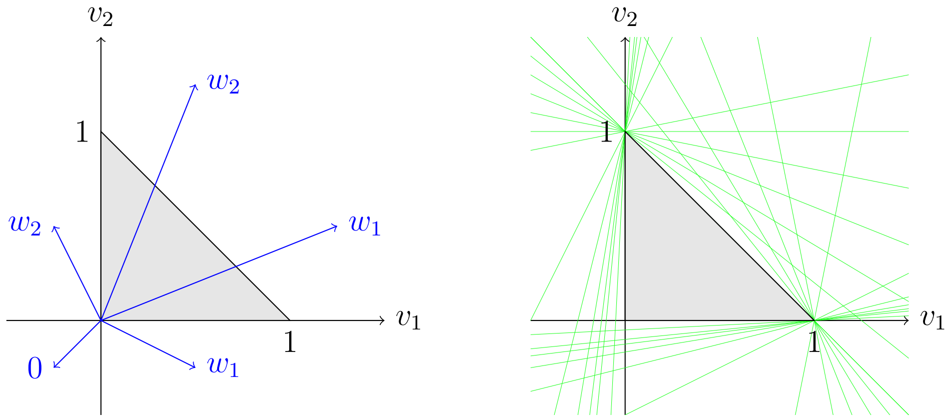

Figure 1. Illustration of duality. The left panel shows how the support function $\mathcal{S}^C(w_1, w_2)$ follows by maximizing $\mathbf{v} \cdot \mathbf{w}$ over the simplex $C = \{(v_1, v_2) \,|\, v_1 \geq 0, v_2 \geq 0, v_1 + v_2 \leq 1\}$. The right panel shows how the simplex can be recovered from the inequalities $\mathbf{v} \cdot \mathbf{w} \leq \mathcal{S}^C(w_1, w_2)$ for all $\mathbf{w} \in \mathbb{R}^2$, where the support function is given by $\mathcal{S}^C(w_1, w_2) = \max(0, w_1, w_2)$.

The support functions for the convex sets C_1 and C_2 can be used to construct the support function for associated sets. Two relevant cases are the sum

$$C_1 + C_2 \;=\; \{\mathbf{v}_1 + \mathbf{v}_2 \,|\, \mathbf{v}_1 \in C_1,\, \mathbf{v}_2 \in C_2\}$$

and the intersection $C_1 \cap C_2$, both of which are convex. An example of the sum is shown in the left panel of Figure 2, where C_1 and C_2 are one-dimensional simplices embedded in $\mathbb{R}^2$ (indicated by the thick lines on the axes) and their sum is the unit square. The example in the right panel shows the intersection of the two-dimensional simplex $C_1 = \{\mathbf{v} \in \mathbb{R}^2_+ | v_1 + v_2 \leq 1\}$ with the half space $C_2 = \{\mathbf{v} \in \mathbb{R}^2 \,|\, v_2 \geq v_1\}$.

Fact 2. *If C_1 and C_2 are non-empty closed convex sets with non-empty intersection then*

$$S^{C_1+C_2}(\mathbf{w}) = S^{C_1}(\mathbf{w}) + S^{C_2}(\mathbf{w})$$

and

$$S^{C_1 \cap C_2}(\mathbf{w}) \;=\; \inf\{S^{C_1}(\mathbf{w}_1) + S^{C_2}(\mathbf{w}_2) \,|\, \mathbf{w}_1 + \mathbf{w}_2 \;=\; \mathbf{w}\}$$

The left panel of Figure 2 illustrates the support function of the sum of two one-dimensional simplices. Recall from Fact 1 that $\mathcal{S}^{C_1}(\mathbf{w}) = \max(0, w_1)$ and $\mathcal{S}^{C_2}(\mathbf{w}) = \max(0, w_2)$. It is readily verified that the support function for the unit square is simply the sum, i.e. $\mathcal{S}^{C_1+C_2}(\mathbf{w}) =$

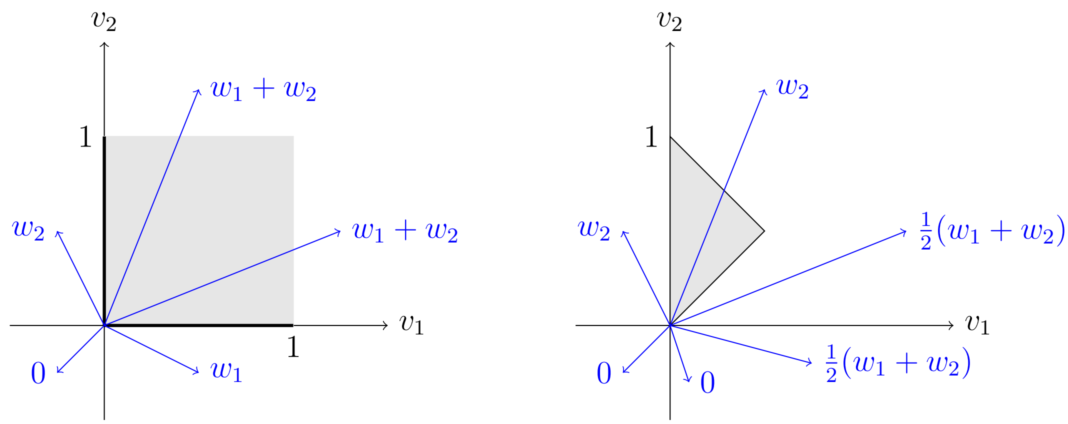

Figure 2. The left panel shows the support function for the sum of two one-dimensional simplices. The right panel shows the support function for the intersection of the two-dimensional simplex and the half-space above the 45 degree line.

$\max(0, w_1) + \max(0, w_2)$, as indicated by the labels next to the arrows in the left panel of Figure 2.

To compute the support function for the intersection in the right panel, we first need to determine the support function for the unbounded half space C_2. Define $\boldsymbol{\beta} = (-1, 1)$ so that the constraint $v_2 \geq v_1$ can be written as $\boldsymbol{\beta} \cdot \mathbf{v} \geq 0$. We have

$$\mathcal{S}^{C_2}(\mathbf{w}) = \begin{cases} \infty & \text{if} \quad \mathbf{w} \neq -\lambda\boldsymbol{\beta} \\ 0 & \text{if} \quad \mathbf{w} = -\lambda\boldsymbol{\beta} \end{cases}$$

for $\lambda \geq 0$. The support function of the intersection can thus be written as

$$S^{C_1 \cap C_2}(\mathbf{w}) = \inf_{\lambda \geq 0} \max(0, w_1 - \lambda, w_2 + \lambda)$$

The infimum is attained when $\lambda = \max(0, \frac{1}{2}(w_1 - w_2))$ and the resulting support function is

$$S^{C_1 \cap C_2}(\mathbf{w}) = \begin{cases} \max(0, w_1, w_2) & \text{if} \quad w_1 \leq w_2 \\ \max(0, \frac{1}{2}(w_1 + w_2)) & \text{if} \quad w_1 \geq w_2 \end{cases}$$

as shown by the labels next to the arrows in the right panel of Figure 2.

In the applications below, we will frequently need the support function for a closed convex set on which multiple constraints $\boldsymbol{\beta}_m \cdot \mathbf{v} \geq K_m$ for $m = 1, \ldots, M$ are imposed. The support function can be derived by repeatedly applying Fact 2, and we list the result here for convenience.

Fact 3. *The support function of a closed convex set C intersected with the half spaces $\boldsymbol{\beta}_m \cdot \mathbf{v} \geq K_m$ for $m = 1, \ldots, M$ is given by*

$$\inf_{\lambda_1,\ldots,\lambda_M \geq 0} S^C(\mathbf{w} + \sum_{m=1}^{M} \lambda_m \boldsymbol{\beta}_m) - \sum_{m=1}^{M} \lambda_m K_m$$

We next determine how the support function transforms under a linear mapping $A : \mathbb{R}^n \to \mathbb{R}^m$. For any $\mathbf{v} \in \mathbb{R}^n$ and $\mathbf{w} \in \mathbb{R}^m$ we have $A\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot A^T\mathbf{w}$ where A^T denotes the transpose of A, i.e. $(A^T)_{ij} = A_{ji}$. This well-known property of the inner product can be used to derive the support function of the convex set AC obtained by applying the linear transformation A to elements of the convex set C.

Fact 4. *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. For any closed convex set C in $\mathbb{R}^n$*

$$S^{AC}(\mathbf{w}) = S^C(A^T\mathbf{w})$$

for any $\mathbf{w} \in \mathbb{R}^m$.

The left panel of Figure 3 demonstrates this fact when the linear transformation

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

is applied to the two-dimensional simplex shown in the left panel of Figure 1. The gray shaded area depicts the resulting convex set and the associated support function is given by $S^{AC}(\mathbf{w}) = \max(0, w_1, w_1 + w_2)$ as indicated by the labels next to the arrows.

An alternative way to represent a convex set is in terms of its extreme points or "vertices." In the right panel of Figure 3 these vertices are indicated by the red dots. It is well known that a bounded closed convex set is simply the convex hull of its vertices, which can be obtained by computing the gradient of the support function at points where it is differentiable. To illustrate, consider the support function of the two-dimensional simplex $\mathcal{S}^C(\mathbf{w}) = \max(0, w_1, w_2)$, which is differentiable when (i) $\max(w_1, w_2) < 0$, (ii) $\max(0, w_1) < w_2$, and (iii) $\max(0, w_2) < w_1$. The corresponding gradients yield the three vertices shown in the right panel of Figure 3, i.e. $\mathcal{V}_1 = (0, 0)$, $\mathcal{V}_2 = (0, 1)$, and $\mathcal{V}_3 = (1, 0)$ respectively.

What about the edges or "faces" of the two-dimensional simplex, which are labeled $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{F}_3$ in the right panel of Figure 3. The faces correspond to points of non-differentiability of

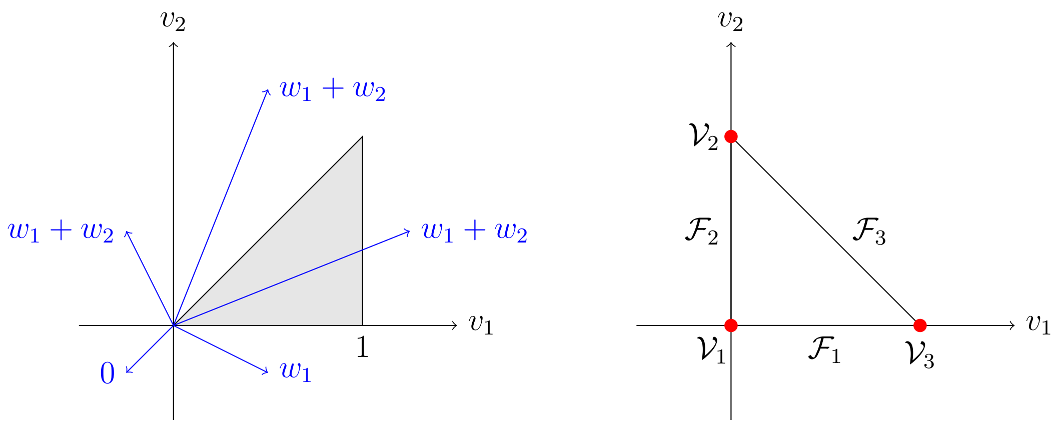

Figure 3. The left panel illustrates the support function for a transformed two-dimensional simplex. In the right panel the two-dimensional simplex is seen as the convex hull of its vertices, $\mathcal{V}_i$, which can be calculated by taking the gradient of the support function at points of differentiability. Similarly, the edges or faces, $\mathcal{F}_i$, of the two-dimensional simplex correspond to the subgradient of the support function at points of non-differentiability.

the support function in which case the gradient should be replaced by the subgradient. Recall that a vector $\mathbf{g} \in \mathbb{R}^n$ is a subgradient of the support function at $\mathbf{w} \in \mathbb{R}^n$ if

$$\mathcal{S}^C(\mathbf{z}) \geq \mathcal{S}^C(\mathbf{w}) + \mathbf{g} \cdot (\mathbf{z} - \mathbf{w})$$

for all $\mathbf{z} \in \mathbb{R}^n$. Consider, for example, the support function $\mathcal{S}^C(\mathbf{w}) = \max(0, w_1, w_2)$ of the two-dimensional simplex and the following points of non-differentiability, $w_1 = 0$ and $w_2 < 0$. For any such $\mathbf{w}$ the subgradient is any vector $\mathbf{g} = (g_1, 0)$ with $0 \leq g_1 \leq 1$.[3] The set of all subgradients at $\mathbf{w}$ is called the subdifferential of the support function at $\mathbf{w}$. To keep the notation simple we will also denote the subdifferential by $\nabla\mathcal{S}^C(\mathbf{w})$. For $\mathbf{w} = (0, w_2)$ with $w_2 < 0$, the subdifferential consists of the face $\mathcal{F}_1$ that connects $\mathcal{V}_1$ and $\mathcal{V}_3$ in the right panel of Figure 3. The other faces can be recovered similarly by considering other points of non-differentiability, i.e. $\mathcal{F}_2$ follows from $\mathbf{w} = (w_1, 0)$ with $w_1 < 0$ and $\mathcal{F}_3$ follows from $\mathbf{w} = (w, w)$ with $w > 0$.

To summarize, any point on the boundary of the convex set C can be written as $\nabla\mathcal{S}^C(\mathbf{w})$ for some $\mathbf{w} \in \mathbb{R}^n$. This allows for the following characterization of the maximization of a linear function over the convex set C.

[3]Let $\mathbf{g} = (g_1, g_2)$. For $z_1 < 0$ and $z_2 = w_2 < 0$ the subgradient inequality yields $g_1 \geq 0$ and for $z_1 > 0$ and $z_2 = w_2$ it yields $g_1 \leq 1$. Likewise, for $z_1 = 0$ and $z_2 = 0$ the subgradient inequality yields $g_2 \leq 0$ and for $z_1 = 0$ and $z_2 < w_2$ it yields $g_2 \geq 0$.

Fact 5. *Consider some closed convex set C in $\mathbb{R}^n$ and some vector $\boldsymbol{\omega} \in \mathbb{R}^n$. Then*

$$\begin{aligned}\sup\{\mathbf{v}\cdot\boldsymbol{\omega}\,|\,\mathbf{v}\in C\} &= \mathcal{S}^C(\boldsymbol{\omega})\\ \arg\sup\{\mathbf{v}\cdot\boldsymbol{\omega}\,|\,\mathbf{v}\in C\} &= \nabla\mathcal{S}^C(\boldsymbol{\omega})\end{aligned}$$

The second part of Fact 5 is a generalization of the envelope theorem, or Hotelling's lemma, that allows for points of non-differentiability. This form of Hotelling's lemma will play an important role in the applications below, where it is used to derive the optimal allocation rule directly from the support function.

We end this section with a result from majorization theory. Let $p_1, \ldots, p_n$ denote arbitrary non-negative numbers and consider two sequences $\boldsymbol{\sigma}$ and $\boldsymbol{\varsigma}$ with elements σ_i, ς_i for $i = 1, \ldots, n$. We will write $\boldsymbol{\sigma} \succeq_p \boldsymbol{\varsigma}$ if

$$\begin{aligned}\sum_{i=1}^{j} p_i\sigma_i &\geq \sum_{i=1}^{j} p_i\varsigma_i \qquad \text{for } \ j \;=\; 1,\ldots,n-1\\ \sum_{i=1}^{n} p_i\sigma_i &= \sum_{i=1}^{n} p_i\varsigma_i\end{aligned}$$

The following result, due to Fuchs (1947), can be found in Marshall, Olkin, and Arnold (2011).

Fact 6. *If $\boldsymbol{\sigma}$, $\boldsymbol{\varsigma}$ are non-decreasing sequences and $\boldsymbol{\sigma} \succeq_p \boldsymbol{\varsigma}$ then we say that $\boldsymbol{\sigma}$ p-majorizes $\boldsymbol{\varsigma}$ and we have*

$$\sum_{i=1}^{n} p_i g(\sigma_i) \leq \sum_{i=1}^{n} p_i g(\varsigma_i)$$

for any continuous convex function $g : \mathbb{R} \to \mathbb{R}$.

Consider any sequence $\boldsymbol{\sigma}$, not necessarily non-decreasing, and let $\boldsymbol{\sigma}^+$ denote the *non-decreasing* sequence such that (i) $\boldsymbol{\sigma} \succeq_p \boldsymbol{\sigma}^+$ and (ii) any other *non-decreasing* sequence $\boldsymbol{\varsigma}$ that satisfies $\boldsymbol{\sigma} \succeq_p \boldsymbol{\varsigma}$ is p-majorized by $\boldsymbol{\sigma}^+$.[4] The second property motivates calling $\boldsymbol{\sigma}^+$ the *largest* non-decreasing sequence that satisfies $\boldsymbol{\sigma} \succeq_p \boldsymbol{\sigma}^+$. Its usefulness stems from the following fact.

Lemma 1. *For any sequence $\boldsymbol{\sigma}$ and any convex function $g : \mathbb{R} \to \mathbb{R}$, $\boldsymbol{\varsigma} = \boldsymbol{\sigma}^+$ solves*

$$\min_{\boldsymbol{\sigma} \succeq_p \boldsymbol{\varsigma}} \sum_{i=1}^{n} p_i g(\varsigma_i)$$

[4] See Bapat (1991) for arguments that ensure existence of such a sequence.

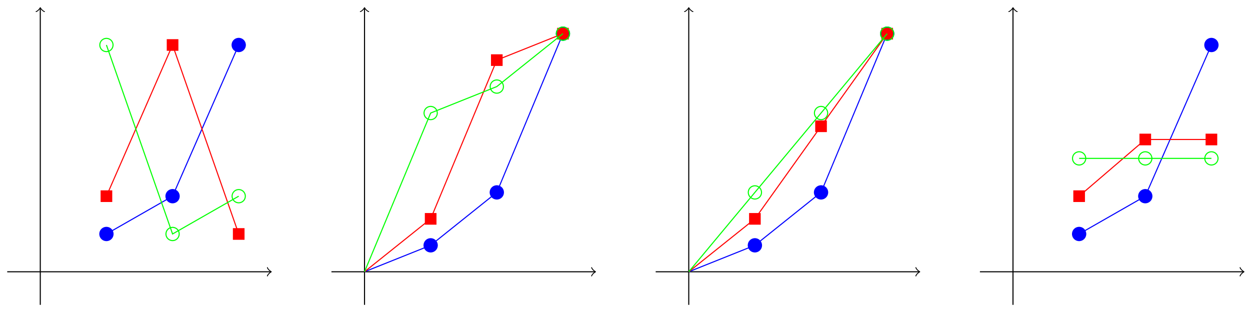

Figure 4. Illustration of majorization. The three sequences in the leftmost panel are $\boldsymbol{\sigma}_1 = (1, 2, 6)$ (solid blue circles), $\boldsymbol{\sigma}_2 = (2, 6, 1)$ (red squares), and $\boldsymbol{\sigma}_3 = (6, 1, 2)$ (open green circles). The rightmost panel shows the corresponding $\boldsymbol{\sigma}^+$ sequences: $\boldsymbol{\sigma}_1^+ = (1, 2, 6)$, $\boldsymbol{\sigma}_2^+ = (2, \frac{7}{2}, \frac{7}{2})$, and $\boldsymbol{\sigma}_3^+ = (3, 3, 3)$. The two middle panels (with rescaled y-axis) show the cumulative sequences for $\boldsymbol{\sigma}$ (middle-left) and $\boldsymbol{\sigma}^+$ (middle-right). The cumulative of $\boldsymbol{\sigma}^+$ is the largest convex function below the cumulative of $\boldsymbol{\sigma}$.

Figure 4 illustrates the construction when $n = 3$ and $p_i = 1$ for $i = 1, \ldots, n$. The leftmost panel shows the sequences $\boldsymbol{\sigma}_1 = (1, 2, 6)$, $\boldsymbol{\sigma}_2 = (2, 6, 1)$, and $\boldsymbol{\sigma}_3 = (6, 1, 2)$. The rightmost panel shows the corresponding $\boldsymbol{\sigma}_1^+ = (1, 2, 6)$, $\boldsymbol{\sigma}_2^+ = (2, \frac{7}{2}, \frac{7}{2})$, and $\boldsymbol{\sigma}_3^+ = (3, 3, 3)$. Note that $\boldsymbol{\sigma} = \boldsymbol{\sigma}^+$ if and only if $\boldsymbol{\sigma}$ is non-decreasing. The middle panels show the cumulative sequences for $\boldsymbol{\sigma}$ (left) and $\boldsymbol{\sigma}^+$ (right) and demonstrates that the cumulative of $\boldsymbol{\sigma}^+$ is the largest convex function that is below the cumulative of $\boldsymbol{\sigma}$. Our discrete majorization procedure thus parallels Myerson's (1981) "ironing" technique for continuous type spaces. Lemma 1 will be important when minimizing the support function with respect to the constraint parameters as in Fact 3.

3. Social Choice Implementation

We consider an environment with a finite set $\mathcal{I} = \{1, 2, \ldots, I\}$ of agents and a finite set $\mathcal{K} = \{1, 2, \ldots, K\}$ of social alternatives. When alternative k is selected, agent i's value is $a_i^k x_i$ where a_i^k is some non-negative constant and agent i's type, x_i, is distributed according to probability distribution $f_i(x_i)$ with discrete support $X_i = \{x_i^1, \ldots, x_i^{N_i}\}$, where the x_i^j are non-negative with $x_i^{j-1} < x_i^j$ for $j = 2, \ldots, N_i$. This formulation is rich enough to include many important applications, e.g. single or multi-unit auctions, public goods provision, bargaining, etc. For example, single-unit auctions are captured by setting $a_i^k = \delta_i^k$ for $i = 1, \ldots, I$ and $k = 1, \ldots, I + 1$, where alternative $i = 1, \ldots, I$ corresponds to the case where bidder i wins the object and alternative $I + 1$ corresponds to the case where the seller keeps the object. As another example, public goods provision can be summarized by two alternatives, i.e. $k = 1$ when the public good is implemented and $k = 2$ when it is not, and $a_i^k = \delta_1^k$ for $i = 1, \ldots, I$.

We denote the profile of all agents' types by $\mathbf{x} = (x_1, ..., x_I) \in X = \prod_{i\in\mathcal{I}} X_i$. Without loss of generality we restrict attention to direct mechanisms characterized by $K + I$ functions, $\{q^k(\mathbf{x})\}_{k\in\mathcal{K}}$ and $\{t_i(\mathbf{x})\}_{i\in\mathcal{I}}$, where $t_i(\mathbf{x}) \in \mathbb{R}$ is agent i's payment and $q^k(\mathbf{x})$ is the probability that alternative k is implemented. We define $v_i(\mathbf{x}) = \sum_{k\in\mathcal{K}} a_i^k q^k(\mathbf{x})$ so that agent i's utility from truthful reporting, assuming others report truthfully as well, is $u_i(\mathbf{x}) = x_i v_i(\mathbf{x}) - t_i(\mathbf{x})$.

3.1. Feasibility

The probabilities with which the alternatives occur satisfy the usual feasibility conditions $q^k(\mathbf{x}) \geq 0$ for $k \in \mathcal{K}$, $\mathbf{x} \in X$ and $\sum_{k\in\mathcal{K}} q^k(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in X$. In other words, the feasible $q^k(\mathbf{x})$ define a k-dimensional simplex for each type profile. We can invoke Fact 2 and simply write the support function $\mathcal{S} : \mathbb{R}^{K|X|} \to \mathbb{R}$ as a sum, over all type profiles, of support functions for k-dimensional simplices (cf. left panel of Figure 2)

$$\mathcal{S}(\mathbf{w}) \;=\; \sum_{x\,\in\, X} \;\max_{k\,\in\,\mathcal{K}} \,(0, w_k(\mathbf{x})) \tag{1}$$

From agent i's perspective it is the linear combination $v_i(\mathbf{x}) = \sum_{k\in\mathcal{K}} a_i^k q^k(\mathbf{x})$ that determines her possible ex post values. Let us define the ex post value possibility set ("vps") as

$$\text{vps} \;=\; \Big\{\mathbf{v} \in \mathbb{R}_+^{I|X|} \,|\, \exists \,\text{feasible } \mathbf{q}(\mathbf{x}) \text{ s.t. } v_i(\mathbf{x}) \;=\; \sum\nolimits_{k\in\mathcal{K}} a_i^k q^k(\mathbf{x}) \;\forall\, i \in \mathcal{I},\, \mathbf{x} \in X\Big\}$$

which is a convex, closed, and bounded set. Its support function $\mathcal{S} : \mathbb{R}^{I|X|} \to \mathbb{R}$ follows by applying Fact 4 to the support function in (1):

$$\mathcal{S}_{ex\ post}(\mathbf{w}) \;=\; \sum_{\mathbf{x}\,\in\, X} \;\max_{k\,\in\,\mathcal{K}} \,(0, \sum_{i\,\in\,\mathcal{I}} a_i^k w_i(\mathbf{x})) \tag{2}$$

The latter can be used to derive the support function of the feasible set of interim expected values:

$$\text{VPS} \;=\; \Big\{\mathbf{V} \in \mathbb{R}_+^{\sum_i |X_i|} \,|\, \exists \,\text{feasible } \mathbf{q}(\mathbf{x}) \text{ s.t. } V_i(x_i) \;=\; \sum\nolimits_{k\in\mathcal{K}} E_{\mathbf{x}_{-i}}(a_i^k q^k(\mathbf{x})) \;\forall\, i \in \mathcal{I},\, x_i \in X_i\Big\}$$

Throughout we distinguish interim variables using capital letters, e.g. the interim expected values are denoted $V_i(x_i)$ for $i \in \mathcal{I}$, $x_i \in X_i$. Since $V_i(x_i) = E_{x_{-i}}(v_i(\mathbf{x}))$, going from the ex post to the interim stage entails taking a sum over others' types, $\mathbf{x}_{-i}$, weighted with the product probability $\prod_{j\neq i} f_j(x_j)$. This is a linear transformation so once more we can invoke Fact 4.

To arrive at expressions that are symmetric in the probabilities we multiply the weight $W_i(x_i)$ associated with $V_i(x_i)$ by $f_i(x_i)$ so that all terms are weighted with $f(\mathbf{x}) = \prod_i f_i(x_i)$. In other words, when we define the interim support function $\mathcal{S}_{interim} : \mathbb{R}^{\sum_i |X_i|} \to \mathbb{R}$ as

$$\mathcal{S}_{interim}(\mathbf{W}) \;=\; \sup\{\mathbf{V} \circ \mathbf{W} \,|\, \mathbf{V} \in \text{VPS}\}$$

the inner product on the right side is probability weighted, i.e.

$$\mathbf{V} \circ \mathbf{W} \;=\; \sum_{i \in \mathcal{I}} E_{x_i}(V_i(x_i)W_i(x_i))$$

where $E_{x_i}(V_i(x_i)W_i(x_i)) = \sum_{x_i} f_i(x_i)V_i(x_i)W_i(x_i)$.

Proposition 1. *The support function for the set of feasible interim expected values is*

$$\mathcal{S}_{interim}(\mathbf{W}) \;=\; E_{\mathbf{x}}\left(\max_{k \in \mathcal{K}}(0, \sum_{i \in \mathcal{I}} a_i^k W_i(x_i))\right) \tag{3}$$

The intuition for the support-function inequalities $\mathbf{V} \circ \mathbf{W} \leq \mathcal{S}_{interim}(\mathbf{W})$, is as follows. For given weights, $\mathbf{W}$, the expected value implied by the interim values $\mathbf{V}$ is equal to $\mathbf{V} \circ \mathbf{W}$, which can be no higher than the maximum possible expected value $\mathcal{S}_{interim}(\mathbf{W})$ at these weights.

3.2. Incentive Compatibility

A mechanism is *dominant strategy incentive compatible* (DIC) if truthful reporting is a dominant strategy equilibrium. Necessary and sufficient conditions for a mechanism $(\mathbf{q}, \mathbf{t})$ to be DIC is that $v_i(x_i, \mathbf{x}_{-i})$ is non-decreasing in x_i for all $i \in \mathcal{I}$, $\mathbf{x} \in X$, and that the payments satisfy[5]

$$(v_i(x_i^n, \mathbf{x}_{-i}) - v_i(x_i^{n-1}, \mathbf{x}_{-i}))x_i^{n-1} \;\leq\; t_i(x_i^n, \mathbf{x}_{-i}) - t_i(x_i^{n-1}, \mathbf{x}_{-i}) \;\leq\; (v_i(x_i^n, \mathbf{x}_{-i}) - v_i(x_i^{n-1}, \mathbf{x}_{-i}))x_i^n \tag{4}$$

for $n = 2, \ldots, N_i$. Moreover, ex post individual rationality (EXIR) requires that $u_i(\mathbf{x}) \geq 0$ for $\mathbf{x} \in X$, $i \in \mathcal{I}$, which is most binding for the lowest-type agent and determines the following range of payments for this agent: $0 \leq t_i(x_i^1, \mathbf{x}_{-i}) \leq v_i(x_i^1, \mathbf{x}_{-i})x_i^1$. Note that the ex post individual rationality condition can be included as one of the incentive compatibility constraints in (4), namely for $n = 1$, if we set $x_i^0 = 0$ and $v_i(x_i^0, \mathbf{x}_{-i}) = t_i(x_i^0, \mathbf{x}_{-i}) = 0$.

[5]Notice that we consider only "adjacent" incentive constraints, which are necessary and sufficient when a bidder's value is linear in her private, one-dimensional type (see, e.g., Goeree and Kushnir, 2011).

Similarly, a mechanism $(\mathbf{q}, \mathbf{t})$ is *Bayesian incentive compatible* (BIC) if truthful reporting is a Bayes-Nash equilibrium. BIC holds if and only if $V_i(x_i)$ is non-decreasing in x_i for all $i \in \mathcal{I}$, $x_i \in X_i$, and the payments satisfy

$$(V_i(x_i^n) - V_i(x_i^{n-1}))x_i^{n-1} \;\leq\; T_i(x_i^n) - T_i(x_i^{n-1}) \;\leq\; (V_i(x_i^n) - V_i(x_i^{n-1}))x_i^n \tag{5}$$

for $n = 2, \ldots, N_i$. Furthermore, interim individual rationality (INIR) requires that $U_i(x_i) \geq 0$ for all $x_i \in X_i$, $i \in \mathcal{I}$, which holds if $0 \leq T_i(x_i^1) \leq V_i(x_i^1)x_i^1$. Also this individual rationality constraint can be obtained from (5) for $n = 1$ if we set $V_i(x_i^0) = T_i(x_i^0) = 0$.

3.2.1. BIC–DIC Equivalence

We postpone solving for the payments that satisfy the incentive compatibility conditions to Section 3.2.2. First, we demonstrate that the set of feasible interim expected values that are dominant strategy implementable coincide with those that are Bayesian implementable. In other words, from an interim perspective, dominant strategy incentive compatibility is no more stringent than Bayesian incentive compatibility.

To glean some intuition for this result, we start with a simple example. Consider the case of two agents and two equally likely types $\underline{x} < \overline{x}$. For each of the four type profiles, $(\underline{x}, \underline{x})$, $(\overline{x}, \underline{x})$, $(\underline{x}, \overline{x})$, and $(\overline{x}, \overline{x})$, the feasible set is a two-dimensional simplex, as shown in the left panel of Figure 1. The vps, i.e. the ex post feasible set, is thus the product of four two-dimensional simplices and an element $\mathbf{v}(\mathbf{x}) \in$ vps is an eight-dimensional vector

$$\mathbf{v}(\mathbf{x}) \;=\; (v_1(\underline{x}, \underline{x}), v_1(\overline{x}, \underline{x}), v_1(\underline{x}, \overline{x}), v_1(\overline{x}, \overline{x}), v_2(\underline{x}, \underline{x}), v_2(\underline{x}, \overline{x}), v_2(\overline{x}, \underline{x}), v_2(\overline{x}, \overline{x}))$$

The dominant strategy incentive compatibility constraints that $v_i(x_i, \mathbf{x}_{-i})$ be non-decreasing in x_i for all $\mathbf{x}_{-i}$, $i \in \mathcal{I}$, can be written as $\boldsymbol{\beta}_m \cdot \mathbf{v} \geq 0$ for $m = 1, \ldots, 4$ where

$$\begin{aligned}
\boldsymbol{\beta}_1 &= (-1, 1, 0, 0, 0, 0, 0, 0) \\
\boldsymbol{\beta}_2 &= (0, 0, -1, 1, 0, 0, 0, 0) \\
\boldsymbol{\beta}_3 &= (0, 0, 0, 0, -1, 1, 0, 0) \\
\boldsymbol{\beta}_4 &= (0, 0, 0, 0, 0, 0, -1, 1)
\end{aligned}$$

The VPS, i.e. the set of interim expected values, is a four-dimensional set consisting of vectors

$$\mathbf{V} \;=\; (V_1(\underline{x}), V_1(\overline{x}), V_2(\underline{x}), V_2(\overline{x}))$$

that follow from applying the linear transformation

$$P = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

to elements $\mathbf{v}(\mathbf{x})$ of the vps. The Bayesian incentive compatibility constraints that $V_i(x_i)$ be non-decreasing in x_i for all $i \in \mathcal{I}$, can be written as $\boldsymbol{\mathcal{B}}_m \circ \mathbf{V} \geq 0$ for $m = 1, 2$ where

$$\begin{aligned} \boldsymbol{\mathcal{B}}_1 &= (-1, 1, 0, 0) \\ \boldsymbol{\mathcal{B}}_2 &= (0, 0, -1, 1) \end{aligned}$$

Note that the constraint generated by $\mathcal{B}_1$ is the same as that generated by $P\boldsymbol{\beta}_1$ or $P\boldsymbol{\beta}_2$. Similarly, $\mathcal{B}_2$ generates the same constraint as $P\boldsymbol{\beta}_3$ or $P\boldsymbol{\beta}_4$. In other words, the DIC constraints, $\boldsymbol{\beta}$, are mapped exactly onto the BIC constraints, $\mathcal{B}$. At the interim stage, requiring dominant strategy incentive compatibility is the same as requiring Bayesian incentive compatibility.

To make this intuition more precise we use the support function to keep track of the set of feasible, incentive compatible outcomes. Generalizing the above example, a geometric characterization of the DIC constraints is given by

$$\big(\boldsymbol{e}(x_i^j, \mathbf{x}_{-i}) - \boldsymbol{e}(x_i^{j-1}, \mathbf{x}_{-i})\big) \cdot \mathbf{v} \geq 0$$

for $j = 2, \ldots, N_i$, $\mathbf{x}_{-i} \in \prod_{j \neq i} X_j$, $i \in \mathcal{I}$, where each $\boldsymbol{e}(\mathbf{x})$ is a unit vector in $\mathbb{R}^{I|X|}$. Likewise, the BIC constraints can be written as

$$\Big(\frac{1}{f_i(x_i^j)}\,\boldsymbol{E}(x_i^j) - \frac{1}{f_i(x_i^{j-1})}\,\boldsymbol{E}(x_i^{j-1})\Big) \circ \mathbf{V} \geq 0$$

for $j = 2, \ldots, N_i$, $i \in \mathcal{I}$, where each $\boldsymbol{E}(x_i)$ is a unit vector in $\mathbb{R}^{\sum_i |X_i|}$. The inverse probabilities appear because the inner product used at the interim stage is probability weighted.

To compare BIC and DIC constraints from an agent's viewpoint we determine the interim support functions for both cases. To this end, we define for $j = 1, \ldots, N_i$,

$$\begin{aligned} \Delta\lambda_i(x_i^j, \mathbf{x}_{-i}) &\equiv \lambda_i(x_i^j, \mathbf{x}_{-i}) - \lambda_i(x_i^{j-1}, \mathbf{x}_{-i}) \\ \Delta\Lambda_i(x_i^j) &\equiv \Lambda_i(x_i^j) - \Lambda_i(x_i^{j-1}) \end{aligned}$$

with $\lambda_i(x_i^0, \mathbf{x}_{-i}) = \lambda_i(x_i^{N_i}, \mathbf{x}_{-i}) = 0$ and $\Lambda_i(x_i^0) = \Lambda_i(x_i^{N_i}) = 0$.

Proposition 2. *The support function for the set of feasible interim expected values that satisfy dominant strategy incentive compatibility is given by*

$$\mathcal{S}^{DIC}_{interim}(\mathbf{W}) \;=\; \inf_{0\,\leq\,\lambda_i(\mathbf{x})}\;\; E_{\mathbf{x}}\Big(\max_{k\,\in\,\mathcal{K}}\big(0, \sum_{i\,\in\,\mathcal{I}} a_i^k(W_i(x_i) - \frac{\Delta\lambda_i(\mathbf{x})}{f_i(x_i)})\big)\Big) \tag{6}$$

Likewise, the support function for the set of feasible interim expected values that satisfy Bayesian incentive compatibility is

$$\mathcal{S}^{BIC}_{interim}(\mathbf{W}) \;=\; \inf_{0\,\leq\,\Lambda_i(x_i)}\;\; E_{\mathbf{x}}\Big(\max_{k\,\in\,\mathcal{K}}\big(0, \sum_{i\,\in\,\mathcal{I}} a_i^k(W_i(x_i) - \frac{\Delta\Lambda_i(x_i)}{f_i(x_i)})\big)\Big) \tag{7}$$

The minimization problem that defines the DIC support function involves more parameters and, hence, could result in a lower support function (reflecting a smaller set). This is not the case, however, if an agent's ex post DIC constraints for different profiles of others' types are all mapped to the same BIC constraint when we take an interim viewpoint, as in the example above. In terms of the minimization problems in Proposition 2 this would imply that the optimal parameters satisfy $\lambda_i(x_i, \mathbf{x}_{-i}) = \Lambda_i(x_i)$ for all $\mathbf{x}_{-i}$.

Consider again the above example with two agents and two equally-likely types and suppose we set $a_i^k = \delta_i^k$ and impose symmetry so that we can drop agent-specific subscripts. The two support functions are then given by[6]

$$\mathcal{S}^{DIC}_{interim}(\mathbf{W}) \;=\; \inf_{0\,\leq\,\underline{\lambda},\overline{\lambda}}\; \frac{1}{4}\max(0, \underline{W} - \underline{\lambda}) + \frac{1}{2}\max(0, \underline{W} - \overline{\lambda}, \overline{W} + \underline{\lambda}) + \frac{1}{4}\max(0, \overline{W} + \overline{\lambda})$$

where $\underline{W}, \overline{W}$ are the weights associated with $\underline{x}$ and $\overline{x}$ respectively, and

$$\mathcal{S}^{BIC}_{interim}(\mathbf{W}) \;=\; \inf_{0\,\leq\,\Lambda}\; \frac{1}{4}\max(0, \underline{W} - \Lambda) + \frac{1}{2}\max(0, \underline{W} - \Lambda, \overline{W} + \Lambda) + \frac{1}{4}\max(0, \overline{W} + \Lambda)$$

The solution to the latter problem is readily calculated as $\Lambda = \max(0, \frac{1}{2}(\underline{W} - \overline{W}))$. Furthermore, if we set $\underline{\lambda} = \Lambda$ in the DIC minimization problem then the solution for $\overline{\lambda}$ is Λ. Conversely, if we set $\overline{\lambda} = \Lambda$ then the solution for $\underline{\lambda}$ is Λ. In other words, $(\underline{\lambda}, \overline{\lambda}) = (\Lambda, \Lambda)$ generates a local minimum, and by convexity of the support function, it is the global solution. For this example, the resulting BIC and DIC support functions are thus the same.

[6]Without loss of generality we can scale the λ and Λ parameters by $\frac{1}{2}$.

We next establish this equivalence more generally and determine the support function that results from minimizing over the Λ (or λ) parameters. Define the shifted weights

$$\widetilde{W}_i(x_i) \;=\; W_i(x_i) - \frac{\Delta\Lambda_i(x_i)}{f_i(x_i)}$$

It is straightforward to verify[7] that for all $i \in \mathcal{I}$

$$\begin{aligned} \sum_{j=1}^{k} f_i(x_i^j)W_i(x_i^j) \;&\geq\; \sum_{j=1}^{k} f_i(x_i^j)\widetilde{W}_i(x_i^j) \qquad \text{for} \;\; k \;=\; 1,\ldots,N_i-1 \\ \sum_{j=1}^{N_i} f_i(x_i^j)W_i(x_i^j) \;&=\; \sum_{j=1}^{N_i} f_i(x_i^j)\widetilde{W}_i(x_i^j) \end{aligned}$$

i.e. $\mathbf{W}_i \succeq_{f_i} \widetilde{\mathbf{W}}_i$. Let $\mathbf{W}_i^+$ denote the largest non-decreasing sequence that satisfies $\mathbf{W}_i \succeq_{f_i} \mathbf{W}_i^+$ for $i \in \mathcal{I}$, and let $\mathbf{W}^+$ denote their concatenation (cf. Lemma 1).

Proposition 3 (BIC–DIC Equivalence). *The support function for the set of feasible interim expected values that satisfy BIC or DIC incentive compatibility is given by*

$$\mathcal{S}^{BIC}_{interim}(\mathbf{W}) \;=\; \mathcal{S}^{DIC}_{interim}(\mathbf{W}) \;=\; \mathcal{S}_{interim}(\mathbf{W}^+)$$

for any $\mathbf{W} \in \mathbb{R}^{\sum_i |X_i|}$ *where* $\mathcal{S}_{interim}$ *is given by equation (3) in Proposition 1.*

We next construct equivalent payments for the BIC and DIC mechanisms so that they deliver the same interim expected utilities to all agents. Importantly, we show that common objectives such as revenue or surplus maximization can be interpreted as optimizing a linear function over the set of feasible, incentive compatible outcomes. This allows us to derive the optimal mechanism by applying Hotelling's lemma to the support function in Proposition 3.

3.2.2. Payments and Revenues

The incentive constraints in (4) and (5) bound the difference in payments in terms of the difference in values times a number that lies between x_i^{n-1} and x_i^n. So let us define the convex combination

$$x_i^n(\alpha_i^n) \;\equiv\; (1-\alpha_i^n)x_i^{n-1} + \alpha_i^n x_i^n$$

where $0 \leq \alpha_i^n \leq 1$ for $n \;=\; 1,\ldots,N_i$, $i \in \mathcal{I}$.

[7] Since $\sum_{j=1}^{k} \Delta\Lambda_i(x_i^j) = \Lambda_i(x_i^k) - \Lambda_i(x_i^0) \geq 0$ for $k = 1,\ldots,N_i$ with equality for $k = N_i$.

Lemma 2. *In any dominant strategy incentive compatible, ex post individually rational mechanism the payments are given by*

$$t_i(x_i^n, \mathbf{x}_{-i}) \;=\; v_i(x_i^n, \mathbf{x}_{-i}) x_i^n(\alpha_i^n) - \sum_{j=1}^{n-1} \big(x_i^{j+1}(\alpha_i^{j+1}) - x_i^j(\alpha_i^j)\big) v_i(x_i^j, \mathbf{x}_{-i})$$

for $n = 1, \ldots, N_i$ *and* $i \in \mathcal{I}$*. Likewise, in any Bayesian incentive compatible, interim individually rational mechanism the payments are given by*

$$T_i(x_i^n) \;=\; V_i(x_i^n) x_i^n(\alpha_i^n) - \sum_{j=1}^{n-1} \big(x_i^{j+1}(\alpha_i^{j+1}) - x_i^j(\alpha_i^j)\big) V_i(x_i^j)$$

for $n = 1, \ldots, N_i$ *and* $i \in \mathcal{I}$*. The lowest and highest BIC and DIC payments follow by setting* $\alpha_i^n = 0$ *and* $\alpha_i^n = 1$ *respectively.*

The BIC-DIC equivalence result of Proposition 3 implies that for any increasing $V_i(x_i)$ one can construct $v_i(x_i, \mathbf{x}_{-i})$ such that $E_{\mathbf{x}_{-i}}(v_i(x_i, \mathbf{x}_{-i})) = V_i(x_i)$. The interim expected values of the DIC payments are therefore equal to the BIC payments, i.e. $E_{\mathbf{x}_{-i}}(t_i(x_i, \mathbf{x}_{-i})) = T_i(x_i)$. An important consequence is that the BIC and DIC mechanism yield the same expected utilities for all agents, i.e. they are equivalent (see also Goeree and Kushnir, 2011). In particular, the expected revenue from any BIC mechanism can be obtained from an equivalent DIC mechanism.

To characterize the range of possible expected revenues, let us define the cumulative probabilities $F_i(x_i^n) = \sum_{j=1}^n f_i(x_i^j)$ and the marginal revenues

$$MR_i(x_i^n) \;=\; x_i^n(\alpha_i^n) - \big(x_i^{n+1}(\alpha_i^{n+1}) - x_i^n(\alpha_i^n)\big) \frac{1 - F_i(x_i^n)}{f_i(x_i^n)}$$

for $n = 1, \ldots, N_i$, $i \in \mathcal{I}$ with $x_i^{N_i+1} = x_i^{N_i}$. Let $\mathbf{MR}(\boldsymbol{\alpha})$ denote the vector with elements $MR_i(x_i^n)$, where we make explicit the dependence on the α_i^n parameters. Any vector $\boldsymbol{\alpha}$ with elements between 0 and 1 ensures incentive compatibility. Of special interest are the lowest and highest marginal revenues $\underline{\mathbf{MR}} = \mathbf{MR}(\mathbf{0})$ and $\overline{\mathbf{MR}} = \mathbf{MR}(\mathbf{1})$.

Lemma 3. *The expected revenue,* R*, and the expected social surplus,* S*, can be written as*

$$R \;=\; \mathbf{V} \circ \mathbf{MR}(\boldsymbol{\alpha})$$

$$S \;=\; \mathbf{V} \circ \mathbf{x}$$

where $\mathbf{V} \circ \mathbf{W} = \sum_i E_{x_i}(V_i(x_i) W_i(x_i))$ *for any* $\mathbf{W} \in \mathbb{R}^{\sum_i |X_i|}$.

An important corollary to this lemma is a revenue equivalence result for general social choice environments. For mechanisms that employ the same allocation rule and, hence, result in the same interim values, $\mathbf{V}$, the revenue is completely determined by the $\boldsymbol{\alpha}$ constants.[8]

Corollary 1 (Revenue Equivalence). *Any interim individually rational, incentive compatible mechanism that results in interim expected values* $\mathbf{V}$ *yields an expected revenue in the range*

$$\mathbf{V} \circ \underline{\mathbf{MR}} \;\le\; R \;\le\; \mathbf{V} \circ \overline{\mathbf{MR}}$$

Another consequence of Lemma 3 is that the revenue and surplus maximizing mechanisms can be obtained by applying Hotelling's lemma to the support function in Proposition 3. The next proposition establishes this more generally, i.e. for arbitrary social choice problems and arbitrary linear objectives. For notational simplicity we set $\boldsymbol{\alpha} = \mathbf{1}$, corresponding to the highest incentive compatible payments.

Proposition 4 (Optimal Mechanism). *For any social choice problem and for any linear objective* $\mathbf{V} \circ \boldsymbol{\omega}$*, the optimal dominant strategy incentive compatible, ex post individual rational mechanism is*

$$q^k(\mathbf{x}) \;=\; \begin{cases} 1/|\mathcal{M}| & \text{if } k \,\in\, \mathcal{M} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where $\mathcal{M} \equiv \arg\max_{k \,\in\, \mathcal{K}}(0, \sum_{i \in \mathcal{I}} a_i^k \omega_i^+(x_i))$ *and*

$$t_i(\mathbf{x}) \;=\; \sum_{k \,\in\, \mathcal{K}} a_i^k \big(x_i q^k(\mathbf{x}) - \sum_{x_i^j \,<\, x_i} (x_i^{j+1} - x_i^j) q^k(x_i^j, \mathbf{x}_{-i}) \big) \tag{9}$$

In particular, the highest possible revenue, $\mathcal{S}_{interim}(\overline{\mathbf{MR}}^+)$*, follows by choosing* $\boldsymbol{\omega} = \overline{\mathbf{MR}}$ *and the highest possible surplus,* $\mathcal{S}_{interim}(\mathbf{x})$*, follows by choosing* $\boldsymbol{\omega} = \mathbf{x}$*.*

The optimal mechanism is in dominant strategies since $q^k(\mathbf{x})$ and, hence, $v_i(\mathbf{x}) = \sum_k a_i^k q^k(\mathbf{x})$ is non-decreasing in agent i's type, x_i, for any profile of others' types. Ex post individual rationality follows since $t_i(\mathbf{x}) \le \sum_k a_i^k x_i q^k(\mathbf{x}) = x_i v_i(\mathbf{x})$ so $u_i(\mathbf{x}) = x_i v_i(\mathbf{x}) - t_i(\mathbf{x}) \ge 0$.

To illustrate, consider a multi-unit auction with $I \ge 1$ ex ante symmetric bidders and K perfectly divisible units. Bidders' types are distributed according to a common probability

[8] Note that with discrete types there is a larger range of possible revenues than with continuous types. In the continuous case, the expected revenue is pinned down by $\mathbf{V}$ and the payments of the lowest types – stated differently, given $\mathbf{V}$, the only degrees of freedom are $0 \le \alpha_i^1 \le 1$ for $i \in \mathcal{I}$. In the discrete case, the degrees of freedom are $0 \le \alpha_i^n \le 1$ for $n = 1, \ldots, N_i$, $i \in \mathcal{I}$.

distribution $f(\cdot)$ with support $X = \{x^1, \ldots, x^N\}$ for some $N \geq 1$. Suppose bidders have diminishing marginal valuations, i.e. they value winning q units at $v(q)$ where $v(\cdot)$ is some concave function. While the diminishing marginal value assumption is natural for many multi-unit auctions, e.g. treasury auctions, it has proven intractable to derive closed-form solutions for this case (see Ausubel and Cramton, 1998). In contrast, the next proposition establishes the dominant strategy auction when bidders' value functions obey a power law.

Proposition 5. *When $v(q_i) = q_i^{1-\gamma}$ for $i = 1, \ldots, I$, with $0 < \gamma < 1$, the ex post support function is*

$$\mathcal{S}_{ex\ post}(\mathbf{w}) = v(K) \sum_{\mathbf{x} \in X} \Big(\sum_{i=1}^{I} \max(0, w_i(\mathbf{x}))^{1/\gamma} \Big)^{\gamma}$$

The efficient allocation rule is

$$q_i(\mathbf{x}) = K \frac{x_i^{1/\gamma}}{\sum_{j=1}^{I} x_j^{1/\gamma}}$$

Similarly, the optimal allocation rule is[9]

$$q_i(\mathbf{x}) = K \frac{\max(0, \overline{MR}^{+}(x_i))^{1/\gamma}}{\sum_{j=1}^{I} \max(0, \overline{MR}^{+}(x_j))^{1/\gamma}}$$

In both cases, the ex post payment rule is given by

$$t_i(\mathbf{x}) = x_i v(q_i(\mathbf{x})) - \sum_{x_i^j < x_i} (x_i^{j+1} - x_i^j) v(q_i(x_i^j, \mathbf{x}_{-i}))$$

In the limit when γ tends to one, the efficient allocation rule assigns units proportionally to bidders' types while the optimal allocation rule assigns units proportionally to bidders' marginal revenues. For intermediate values, $0 < \gamma < 1$, the efficient and optimal allocation rules resemble "Tullock-type" success functions. Finally, Myerson's (1981) familiar result for the optimal auction is obtained in the limit when γ tends to zero, which corresponds to the linear valuation case $v(q) = q$. Now the efficient allocation rule is to assign all units to the highest-type bidder while the revenue-maximizing allocation rule assigns all units to the bidder with the highest positive marginal revenue (and assigns no units if all marginal revenues are negative).

[9] Where we interpret 0/0 as 0.

4. Incorporating Other Types of Constraints

In this section we demonstrate how our geometric approach facilitates the inclusion of other types of ex ante or ex post constraints, such as budget balancedness in bargaining and public goods provision. First, we consider multi-unit auctions where the bidders have capacity constraints, i.e. their demand is "flat" up to a certain number of units.

4.1. Capacity Constraints in Multi-Unit Auctions

There are $K \geq 1$ perfectly divisible units for sale and $I \geq 1$ bidders with linear valuations. In the absence of any capacity constraints the ex post support function is given by

$$\mathcal{S}_{ex\ post}(\mathbf{w}) \;=\; K \sum_{\mathbf{x}\in X} \max_i(0, w_i(\mathbf{x}))$$

Now suppose that bidder i has capacity $K_i < K$. This implies the following constraint on the ex post allocation rule: $q_i(\mathbf{x}) \leq K_i$, or, equivalently, $\mathbf{q}\cdot\mathbf{e}_i \leq K_i$, where $\mathbf{e}_i$ is the i-th unit vector in $\mathbb{R}^{I|X|}$. Using Fact 3 the support function of the constrained set is

$$\mathcal{S}_{constrained}(\mathbf{w}) \;=\; \inf_{0\leq\lambda_i(\mathbf{x})} \sum_{\mathbf{x}\in X} K \max_i(0, w_i(\mathbf{x}) - \lambda_i(\mathbf{x})) + \sum_{\mathbf{x}\in X}\sum_i K_i\lambda_i(\mathbf{x})$$

For $r = 1, \ldots, I$, let $w_{i(r)}(\mathbf{x})$ denote the weight with rank r for each $\mathbf{x}$. The capacity of the bidder with rank 1 is less than K, i.e. $K_{i(1)} < K$, so raising $\lambda_{i(1)}(\mathbf{x})$ from 0 lowers the objective. Suppose we raise $\lambda_{i(1)}(\mathbf{x})$ to $w_{i(1)}(\mathbf{x}) - w_{i(2)}(\mathbf{x})$ then, as long as $K_{i(1)} + K_{i(2)} < K$, subsequently raising both $\lambda_{i(1)}(\mathbf{x})$ and $\lambda_{i(2)}(\mathbf{x})$ at the same speed lowers the objective, etc. Let $1 \leq r^* \leq I$ denote the largest rank such that

$$\begin{cases} \sum_{r=1}^{r^*} K_{i(r)} \;\leq\; K \\ w_{i(r^*)} \;\geq\; 0 \end{cases}$$

The constrained support function can be written as

$$\mathcal{S}_{constrained}(\mathbf{w}) \;=\; \sum_{\mathbf{x}\in X}\sum_{r=1}^{r^*} K_{i(r)} w_{i(r)}(\mathbf{x}) + (K - \sum_{r=1}^{r^*} K_{i(r)}) \sum_{\mathbf{x}\in X} \max(0, w_{i(r^*+1)}(\mathbf{x}))$$

where $w_{i(r^*+1)} = 0$ if $r^* = I$. The revenue-maximizing allocation rule follows from Hotelling's lemma, i.e. from $\nabla\mathcal{S}_{constrained}(\overline{\mathbf{MR}}^+)$.

Proposition 6. *The revenue-maximizing mechanism assigns to the bidders with the highest positive marginal revenues their capacities until the total quantity K is exhausted or all bidders with positive marginal revenues have been served.*

This result was previously derived by Maskin and Riley (1989) for the case where each bidder has a capacity of 1 (see also Ausubel and Cramton, 1998).

4.2. Public Goods Provision with Ex Post Budget Balancedness

Consider $I \geq 1$ agents whose valuations for a public good are high, $\overline{x}$, with probability p and low, $\underline{x}$, with probability $1-p$. There are only $K = 2$ alternatives, i.e. either the public good is produced or not. Let $q(\mathbf{x})$ denote the probability that the public good is produced when the realized type profile is $\mathbf{x}$. If the marginal cost of producing the public good production is C, the ex post budget balance constraint can be written as

$$\sum_{i \in \mathcal{I}} t_i(\mathbf{x}) \;\geq\; Cq(\mathbf{x}),$$

for each $\mathbf{x} \in X$, and the public good is produced if and only if this condition is met. In the absence of the budget balance constraint, the support function is

$$\mathcal{S}(\mathbf{w}) \;=\; \sum_{\mathbf{x} \in X} \max(0, w(\mathbf{x}))$$

We focus on symmetric mechanisms and denote by $q(m)$, $m = 0, \ldots, I$, the probability that the public good is produced when m agents are of high type and $I - m$ agents are of low type. Similarly, $t(\underline{x}, m)$ and $t(\overline{x}, m)$ are the payments of a low-type agent and a high-type agent in this event. The individual rationality and ex post incentive compatibility constraints imply the following upper bound on the agents' payments:

$$\begin{aligned} t(\underline{x}, m) \;&=\; q(m)\underline{x} \\ t(\overline{x}, m) \;&=\; (q(m) - q(m-1))\overline{x} + q(m-1)\underline{x} \end{aligned}$$

For convenience, let us parameterize the cost of the public good as $C = \gamma\overline{x} + (I - \gamma)\underline{x}$ where $\gamma \in \{0, \ldots, I\}$. The ex post budget balance constraint becomes $\boldsymbol{\beta}_m \cdot \mathbf{q} \geq 0$ for $m = 0, \ldots, I$, where

$$\boldsymbol{\beta}_m \;=\; (0, \ldots, 0, \underbrace{m(\underline{x} - \overline{x})}_{m-1}, \underbrace{(m - \gamma)(\overline{x} - \underline{x})}_{m}, 0, \ldots, 0)$$

The constrained support function can be computed using Fact 3

$$S_{constrained}(\mathbf{w}) \;=\; \inf_{0\,\leq\,\lambda_m}\;\sum_{m\,=\,0}^{I}\max\bigl(0, w_m + (\overline{x} - \underline{x})(\lambda_m(m-\gamma) - \lambda_{m+1}(m+1)\bigr) \tag{10}$$

with $\lambda_{I+1} = 0$. For $m < \gamma$ the coefficients multiplying λ_m are non-positive so the infimum is achieved for $\lambda_m = \infty$. Hence, for these profiles the public good is not provided. For $m \geq \gamma$, the λ_m can be solved recursively by pushing the second argument of the max function to 0, except for the final term of the sum (i.e. when $m = I$).

Proposition 7. *The support function that includes ex post budget balance is given by*

$$\mathcal{S}_{constrained}(\mathbf{w}) \;=\; \sum_{m\,=\,\gamma}^{I} w_m \binom{m}{\gamma}\Big/\binom{I}{\gamma}$$

The optimal allocation rule follows from $\mathbf{q} = \nabla\mathcal{S}_{constrained}$

$$q(m) \;=\; \begin{cases} \binom{m}{\gamma}\Big/\binom{I}{\gamma} & \text{if } m \,\geq\, \gamma \\ 0 & \text{otherwise} \end{cases}$$

The ex ante probability that the public good is produced is equal to

$$\sum_{m\,=\,\gamma}^{I} p^m(1-p)^{I-m}q(m) \;=\; p^\gamma$$

Mailath and Postlewaite (1990) consider the limit when the number of people grows large, $I \to \infty$, and the per-capita cost of producing the public good is constant. This implies that γ is proportional to I and also diverges, so the probability that the public good is produced tends to 0. This is true even when the per-capita cost is low, say $\frac{9}{10}\underline{x} + \frac{1}{10}\overline{x}$, and the probability of a high type is high, say $p = \frac{9}{10}$, so that the per-capita value of the public good, $\frac{1}{10}\underline{x} + \frac{9}{10}\overline{x}$, far exceeds its cost. In other words, the public good is not produced even when it is common knowledge that it would be efficient to do so.

In contrast, Hellwig (2003) assumes that the total cost of producing the public good is constant, which means that γ will tend to 0 as the number of people diverges. Now the probability of efficient public goods provision goes to 1 for all $p > 0$.

4.3. Bargaining with Ex Ante Budget Balancedness

Following Myerson and Satterthwaite (1983) we consider a simple bargaining setting with a single seller and a single buyer. Seller and buyer values are discrete, i.e. $x_b, x_s \in X = \{x^1, \ldots, x^N\}$ for $N \geq 1$, with probability distributions $f_b(x)$ and $f_s(x)$ that are not necessarily the same. Let $q(x_b, x_s)$ denote the probability of trade given profile $\mathbf{x} = (x_b, x_s)$, then the value gain to the buyer is her type times $q(x_b, x_s)$ and the gain to the seller is her type times $-q(x_b, x_s)$. The ex post support function, ignoring budget balance, is therefore

$$\mathcal{S}_{ex\ post}(\mathbf{w}) \;=\; \sum_{\mathbf{x} \in X} \max(0, w_b(\mathbf{x}) - w_s(\mathbf{x}))$$

The interim expected probability of trade from the buyer's point of view is given by $Q_b(x_b) = E_{x_s}(q(x_b, x_s))$ and from the seller's point of view it is $Q_s(x_s) = E_{x_b}(q(x_b, x_s))$. The implied interim support function is

$$\mathcal{S}_{interim}(\mathbf{W}) \;=\; E_{\mathbf{x}}\big(\max(0, W_b(x_b) - W_s(x_s))\big)$$

We consider ex ante budget balance, which can be stated as $\mathbf{Q} \circ \mathbf{MR} \geq 0$ where $\mathbf{Q} = (Q_b(x_b), -Q_s(x_s))$ and the buyer and seller marginal revenues are

$$\begin{aligned} MR_b(x^j) &= x^j - (x^{j+1} - x^j)\frac{1 - F_b(x^j)}{f_b(x^j)} \\ MR_s(x^j) &= x^j + (x^j - x^{j-1})\frac{F_s(x^{j-1})}{f_s(x^j)} \end{aligned}$$

for $j = 1, \ldots, N$ with $x^0 = 0$ and $x^{N+1} = x^N$. Using Fact 3, the interim support function that satisfies ex ante budget balance is

$$\mathcal{S}_{constrained}(\mathbf{W}) \;=\; \inf_{0 \leq \lambda}\; E_{\mathbf{x}}\big(\max(0, W_b(x_b) - W_s(x_s) + \lambda(MR_b(x_b) - MR_s(x_s)))\big)$$

When the constrained support function is evaluated at $W_b(x_b) = x_b$ and $W_s(x_s) = x_s$, the argument of the max function is positive if and only if $MR_b^\lambda(x_b) > MR_s^\lambda(x_s)$ where we defined the generalized marginal revenues

$$\begin{aligned} MR_b^\lambda(x^j) &= x^j - (x^{j+1} - x^j)\frac{\lambda}{1+\lambda}\frac{1 - F_b(x^j)}{f_b(x^j)} \\ MR_s^\lambda(x^j) &= x^j + (x^j - x^{j-1})\frac{\lambda}{1+\lambda}\frac{F_s(x^{j-1})}{f_s(x^j)} \end{aligned}$$

The surplus-maximizing allocation rule can be read off from the gradient $\nabla\mathcal{S}_{constrained}$.

Proposition 8. *The trading rule that maximizes social surplus is given by*

$$q(x_b, x_s) \;=\; \begin{cases} 1 & \text{if } MR_b^\lambda(x_b) \;>\; MR_s^\lambda(x_s) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

with $0 < \lambda < \infty$ *determined by*

$$E_{MR_b^\lambda(x_b) > MR_s^\lambda(x_s)}\big(MR_b(x_b) - MR_s(x_s)\big) \;=\; 0$$

In particular, in the second-best outcome, neither "no trade" $(\lambda = \infty)$ *nor "fully efficient trade"* $(\lambda = 0)$ *occur.*

5. Extensions

When we relax one of the assumptions of the main model, i.e. that values are linear in types and types are private, independent, and one-dimensional, the equivalence between Bayesian and dominant strategy implementation breaks down. Importantly, however, this does not mean that our geometric approach cannot be applied. Let $v_i^k(\mathbf{x})$ denote agent i's value when alternative k is selected and the profile of types is $\mathbf{x} = (x_1^1, \ldots, x_1^{T_1}, \ldots, x_I^1, \ldots, x_I^{T_I})$, i.e. agent i's type is T_i dimensional. Let $\mathbf{X}_i = \prod_{j=1}^{T_i} X_{ij}$ where each $X_{ij} = \{x_{ij}^1, \ldots, x_{ij}^{N_{ij}}\}$ and let $\mathbf{X} = \prod_i \mathbf{X}_i$. We allow for correlation in types and denote the joint probability distribution by $f(\mathbf{x})$.

Note that this setup relaxes all assumptions of the main model: the values v_i^k can be non-linear functions of the types, the values are not private since they depend on others' types, and types are correlated and multi-dimensional. While the setup is much more general, the derivation of the ex post and interim support functions parallels that of Proposition 1.

Proposition 9. *The support function* $\mathcal{S}_{ex\ post} : \mathbb{R}^{I|\mathbf{X}|} \to \mathbb{R}$ *for the feasible ex post values is*

$$\mathcal{S}_{ex\ post}(\mathbf{w}) \;=\; \sum_{\mathbf{x} \in X} \max_{k \in \mathcal{K}} \big(0, \sum_{i \in \mathcal{I}} v_i^k(\mathbf{x}) w_i(\mathbf{x})\big)$$

and the support function $\mathcal{S}_{interim} : \mathbb{R}^{\sum_i |\mathbf{X}_i|} \to \mathbb{R}$ *for the feasible interim values is*

$$\mathcal{S}_{interim}(\mathbf{W}) \;=\; E_{\mathbf{x}}\big(\max_{k \in \mathcal{K}} \big(0, \sum_{i \in \mathcal{I}} v_i^k(\mathbf{x}) W_i(\mathbf{x}_i)\big)\big)$$

These support functions determine the set of feasible values without any incentive constraints imposed. For the general model, determining the consequences of Bayesian or dominant strategy incentive compatibility is complicated and requires more than comparing adjacent types only. We leave a complete analysis to future research but illustrate how our methodology applies to the case of linear value interdependencies for which adjacent comparisons are sufficient.

5.1. Interdependent Values

Consider a single-unit auction with two bidders and two equally likely and independent types, $\underline{x} = 1$ and $\overline{x} = 10$. Bidders' values depend on both their types in a simple linear way

$$v_i(x_i, x_j) \;=\; x_i + \alpha x_j$$

for $i \neq j = 1, 2$. A continuous-type version of this example was first studied by Maskin (1992), who showed that when $\alpha > 1$ the first-best efficient outcome is not Bayesian implementable (and, hence, not dominant strategy implementable). Hernando-Veciana and Michelucci (2011) show that, with two bidders, the second-best outcome can be implemented via an English auction, although its equilibrium is not in dominant strategies. In other words, when $\alpha > 1$, the second-best outcome is Bayesian but not dominant strategy implementable and BIC-DIC equivalence fails.

We show these results are neatly explained by comparing the sets of feasible outcomes that satisfy Bayesian and dominant strategy incentive compatible respectively. Since the bidders are ex ante symmetric, the allocation rule has no player specific subscript and can be represented by a four-dimensional vector

$$\mathbf{q}(\mathbf{x}) \;=\; (q(\underline{x},\underline{x}), q(\overline{x},\underline{x}), q(\underline{x},\overline{x}), q(\overline{x},\overline{x}))$$

The dominant strategy incentive compatibility constraints can be written as $\boldsymbol{\beta}_m \cdot \mathbf{q} \geq 0$ for $m = 1, 2$ where

$$\begin{aligned} \boldsymbol{\beta}_1 &= (-1, 1, 0, 0) \\ \boldsymbol{\beta}_2 &= (0, 0, -1, 1) \end{aligned}$$

The VPS, i.e. the set of interim expected values, consists of two-dimensional vectors

$$\mathbf{V} \;=\; (V(\underline{x}), V(\overline{x}))$$

that follow from applying the linear transformation

$$P \;=\; \frac{1}{2}\begin{pmatrix} 1+\alpha & 0 & 1+10\alpha & 0 \\ 0 & 10+\alpha & 0 & 10+10\alpha \end{pmatrix}$$

to the ex post allocation probabilities $\mathbf{q}(\mathbf{x})$. The Bayesian incentive compatibility constraint is most easily expressed in terms of the ex post allocation probabilities, i.e. $\boldsymbol{\mathcal{B}} \cdot \mathbf{q} \geq 0$ where

$$\boldsymbol{\mathcal{B}} \;=\; (-1, 1, -1, 1)$$

Note that for $\alpha > 0$ it is no longer the case that $P\boldsymbol{\beta}_1$ and $P\boldsymbol{\beta}_2$ are proportional to $P\boldsymbol{\mathcal{B}}$. This is the simple reason that BIC-DIC equivalence fails.

To make this precise we next compare the support functions for the feasible interim expected values that satisfy Bayesian and dominant strategy incentive compatibility respectively. Using Proposition 9 and Fact 3 we have

$$\begin{aligned}\mathcal{S}^{DIC}_{interim} \;&=\; \min_{0\,\leq\,\underline{\lambda},\overline{\lambda}}\; \frac{1}{4}\max(0,(1+\alpha)\underline{W}-\underline{\lambda}) + \frac{1}{2}\max(0,(1+10\alpha)\underline{W}-\overline{\lambda},(10+\alpha)\overline{W}+\underline{\lambda}) \\ &\qquad + \frac{1}{4}\max(0,(10+10\alpha)\overline{W}+\overline{\lambda})\end{aligned}$$

and

$$\begin{aligned}\mathcal{S}^{BIC}_{interim} \;&=\; \min_{0\,\leq\,\Lambda}\; \frac{1}{4}\max(0,(1+\alpha)\underline{W}-\Lambda) + \frac{1}{2}\max(0,(1+10\alpha)\underline{W}-\Lambda,(10+\alpha)\overline{W}+\Lambda) \\ &\qquad + \frac{1}{4}\max(0,(10+10\alpha)\overline{W}+\Lambda)\end{aligned}$$

For $\alpha = 0$ the support functions reduce to those in Section 3 and BIC-DIC equivalence holds. However, for $\alpha > 0$ the DIC and BIC minimization problems will in general yield different values for $\underline{\lambda}$, $\overline{\lambda}$, and Λ, resulting in different sets.[10]

Figure 5 shows the sets of interim feasible values that result when $\alpha = 0$ (left panel), $\alpha = \frac{1}{2}$ (middle panel), and $\alpha = 2$ (right panel). In each of the panels, the light area corresponds to the set of feasible outcomes without any incentive constraints imposed, the medium dark area to the Bayesian implementable outcomes, and the dark area to the dominant strategy implementable outcomes. When $\alpha = 0$ the latter two sets coincide as shown by the left panel, but BIC-DIC equivalence generally fails when $\alpha > 0$ as shown by the middle and right panels.

The easiest way to describe the different sets is by their vertices,[11] which correspond to certain specific allocation rules. For instance, the set of feasible outcomes that are dominant

[10] For $(1+10\alpha)\underline{W} \leq (10+\alpha)\overline{W}$ the solutions are $\underline{\lambda} = \overline{\lambda} = \Lambda = 0$, while for $(1+10\alpha)\underline{W} \geq (10+\alpha)\overline{W}$ possible solutions are $\Lambda = \frac{1}{2}((1+10\alpha)\underline{W} - (10+\alpha)\overline{W})$, $\underline{\lambda} = (1+\alpha)\underline{W}$, and $\overline{\lambda} = 9\alpha\underline{W} + \min((1+\alpha)\underline{W}, -(10+\alpha)\overline{W})$.

[11] The vertices follow from the gradient of the support function at points of differentiability. Using the solutions in footnote 10 yields the five DIC vertices $(0,0)$, $(0,15{+}6\alpha)$, $(\frac{1}{2}{+}\frac{1}{2}\alpha, 15{+}6\alpha)$, $(1{+}\frac{11}{2}\alpha, 10{+}\frac{11}{2}\alpha)$, $(\frac{1}{2}{+}5\alpha, 5{+}5\alpha)$. The first four plus $(\frac{3}{4} + \frac{15}{2}\alpha, \frac{15}{2} + \frac{21}{4}\alpha)$, $(\frac{1}{2} + 5\alpha, 5 + \frac{1}{2}\alpha)$ constitute the six BIC vertices.

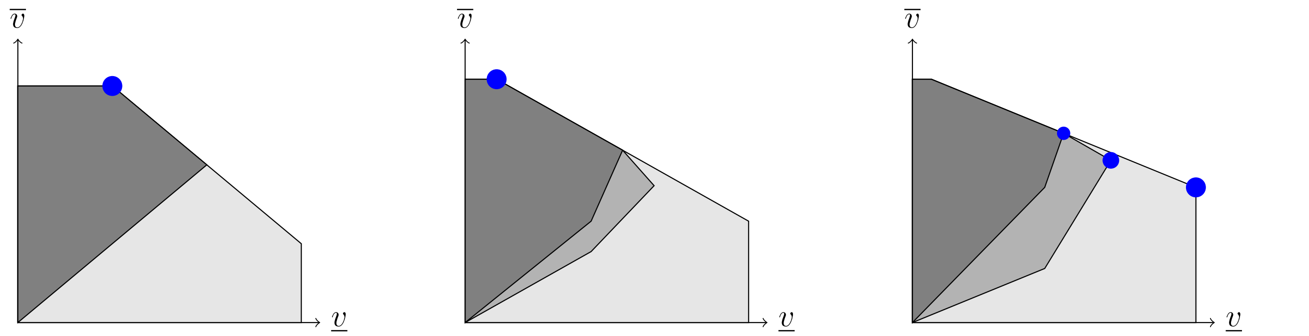

Figure 5. Illustration of BIC-DIC equivalence and its failure. Shown are the feasible outcomes with no incentive constraints imposed (light), Bayesian implementable outcomes (medium dark), and dominant strategy implementable outcomes (dark) for $\alpha = 0$ (left panel), $\alpha = \frac{1}{2}$ (middle panel), and $\alpha = 2$ (right panel). The largest blue dot indicates the first-best outcome, the medium-sized blue dot the second-best outcome under BIC, and the smallest blue dot the second-best outcome under DIC.

strategy incentive compatible can be described by five vertices, which (clockwise starting at the origin) correspond to the following allocation rules

$$\mathbf{q}^{DIC} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$$

where we organize the four-dimensional vector representing the allocation rule by a matrix with entries $q_{11} = q(\underline{x}, \underline{x})$, $q_{12} = q(\underline{x}, \overline{x})$, $q_{21} = q(\overline{x}, \underline{x})$, and $q_{22} = q(\overline{x}, \overline{x})$. Likewise, for the Bayesian implementable outcomes the six vertices correspond to the allocation rules

$$\mathbf{q}^{BIC} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 0 & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$$

Bayesian incentive compatibility requires that the sum of entries in the top row does not exceed the sum of entries in the bottom row. In contrast, dominant strategy incentive compatibility requires that the entries in the top row do not exceed the entries in the bottom row for both columns. Notice that the final two BIC matrices violate this more stringent condition.

The blue dots in Figure 5 indicate first and second-best outcomes. For $\alpha \le 1$, the first-best outcomes under BIC and DIC are the same and correspond to the third DIC or BIC matrix. When $\alpha > 1$, the penultimate DIC matrix, which implies complete randomization, yields the second-best outcome under dominant strategy implementation. Similarly, the penultimate BIC matrix yields the second-best outcome under Bayesian implementation. BIC implementation now leads to more social surplus than DIC, although it is no longer first best.[12]

[12]For $\alpha = 2$, the second-best outcomes under DIC and BIC are $\mathbf{V} = (12, 21)$ and $\mathbf{V} = (\frac{63}{4}, 18)$ respectively, while the first-best outcome is $\mathbf{V} = (\frac{45}{2}, 15)$.

6. Conclusions

This paper introduces a novel approach to mechanism design, one that brings it closer to standard micro-economic analyses of consumer and producer choice. The main insight is to characterize the entire set of feasible and incentive compatible outcomes so that the optimal mechanism can be derived using standard tools such as Hotelling's lemma. We do so by providing a geometric interpretation of the incentive compatibility constraints, which puts them on an equal footing with basic resource constraints. Employing techniques from convex analysis we are able to characterize the resulting high-dimensional set via its support function.

We demonstrate the usefulness of our geometric approach in several ways. First, we characterize the set of feasible outcomes for a general class of social choice problems. The geometric interpretation of the incentive compatibility constraints translates into a well-defined minimization process for the associated support function. Using arguments from majorization theory, we show that this minimization problem naturally generates the "ironing" procedure first described by Mussa and Rosen (1978) and Myerson (1981). Second, we provide a simple geometric proof that when values are linear in types and types are independent, private, and one-dimensional, the feasible sets that remain after imposing Bayesian or dominant strategy incentive compatibility coincide. Third, we derive the optimal mechanism for *any* social choice problem and *any* linear objective, including revenue and surplus maximization. Fourth, we show how to incorporate other types of constraints, e.g. capacity constraints and budget balancedness. Finally, when the equivalence between Bayesian and dominant strategy implementation breaks down our approach naturally produces second-best outcomes for both types of incentive constraints.

Our geometric approach can be extended to continuous type spaces in a straightforward manner, either by starting with support functions for infinite dimensional sets or by considering limit results of our discrete setting. For instance, the optimal allocation rule of Proposition 4 already applies to continuous types and the payment rule extends trivially by replacing the sum with an integral. Throughout the paper, the interim support functions are expressed as expectations over type profiles, reflecting either sums over discrete types or integrals over continuous types. Indeed, we consider it a benefit that our methodology produces parallel results, e.g. optimal mechanisms and revenue equivalence, for continuous and discrete types.

Importantly, our geometric approach applies beyond the main assumptions of linear value functions and independent, private, and one-dimensional types, see Proposition 9. As such it may provide a powerful tool to study mechanism design problems that have hitherto resisted thorough analysis because of analytical intractability, e.g. when type spaces are multi-dimensional. We leave this exciting prospect as a topic for future research.

A. Appendix

Proof of Lemma 1. We first show that, without loss of generality, we can restrict attention to sequences $\boldsymbol{\varsigma}$ that are non-decreasing. Suppose not and $\varsigma_j > \varsigma_k$ for $j < k$. Then define the sequence $\tilde{\boldsymbol{\varsigma}}$ with elements $\tilde{\varsigma}_j = \varsigma_j - \varepsilon(\varsigma_j - \varsigma_k)/p_j$ and $\tilde{\varsigma}_k = \varsigma_k + \varepsilon(\varsigma_j - \varsigma_k)/p_k$ while $\tilde{\varsigma}_i = \varsigma_i$ for $i \neq j, k$. The sequence $\tilde{\boldsymbol{\varsigma}}$ also satisfies $\sigma \succeq_p \tilde{\boldsymbol{\varsigma}}$. Since $g(\cdot)$ is convex we have

$$p_j g(\tilde{\varsigma}_j) + p_k g(\tilde{\varsigma}_k) \leq p_j g(\varsigma_j) + p_k g(\varsigma_k)$$

and, hence, $\sum_{i=1}^{n} p_i g(\tilde{\varsigma}_i) \leq \sum_{i=1}^{n} p_i g(\varsigma_i)$. Repeatedly applying this procedure results in a non-decreasing sequence $\tilde{\boldsymbol{\varsigma}}$ that satisfies $\sigma \succeq_p \tilde{\boldsymbol{\varsigma}}$. But any such sequence is p-majorized by σ^+ and Fact 6 proves the claim. ■

Proof of Proposition 1. The proof follows from the basic definition of the support function

$$\begin{aligned}
\mathcal{S}_{interim}(\mathbf{W}) &= \max\Big\{ \sum\nolimits_i E_{x_i}(V_i(x_i) W_i(x_i)) \,|\, \mathbf{V} \in \text{VPS} \Big\} \\
&= \max\Big\{ \sum\nolimits_k E_{\mathbf{x}}\big(q^k(\mathbf{x}) \sum\nolimits_i a_i^k W_i(x_i)\big) \,|\, \mathbf{q} \text{ is feasible} \Big\} \\
&= E_{\mathbf{x}}\big(\max\nolimits_k (0, \sum\nolimits_i a_i^k W_i(x_i))\big)
\end{aligned}$$

Alternatively, we can derive the interim support function directly from the ex post support function by applying Fact 4 to the linear transformation that corresponds to going from the ex post to the interim stage (i.e. taking expectations over others' types). We have $f_i(x_i)V_i(x_i) = \sum_{\mathbf{x}_{-i}} f(\mathbf{x}) v_i(\mathbf{x})$ where the interim expected value is multiplied by the probability to reflect that the inner product at the interim stage is probability weighted. Using Fact 4, the interim support function follows by evaluating the ex post support function at $w_i(\mathbf{x}) = f(\mathbf{x}) W_i(x_i)$. ■

Proof of Proposition 2. The implications of the DIC constraints for the ex post support function in (2) follows from Fact 3:

$$\mathcal{S}_{ex\ post}^{DIC}(\mathbf{w}) = \min_{0 \leq \lambda_i(\mathbf{x})} \sum_{\mathbf{x} \in X} \max_{k \in \mathcal{K}} \big(0, \sum_{i \in \mathcal{I}} a_i^k (w_i(\mathbf{x}) - \Delta\lambda_i(\mathbf{x}))\big)$$

with $\Delta\lambda_i(x_i^j, \mathbf{x}_{-i}) \equiv \lambda_i(x_i^j, \mathbf{x}_{-i}) - \lambda_i(x_i^{j-1}, \mathbf{x}_{-i})$ for $j = 1, \ldots, N_i$, and $\lambda_i(x_i^0, \mathbf{x}_{-i}) = \lambda_i(x_i^{N_i}, \mathbf{x}_{-i}) = 0$. As in the proof of Proposition 1, the interim support function simply follows by evaluating the ex post support function at $w_i(\mathbf{x}) = f(\mathbf{x}) W_i(x_i)$. If we also replace $\lambda_i(x_i, \mathbf{x}_{-i})$ with $\lambda_i(x_i, \mathbf{x}_{-i}) f_{-i}(\mathbf{x}_{-i})$ we can write the result as

$$\mathcal{S}_{interim}^{DIC}(\mathbf{W}) = \min_{0 \leq \lambda_i(\mathbf{x})} E_{\mathbf{x}} \Big(\max_{k \in \mathcal{K}} \big(0, \sum_{i \in \mathcal{I}} a_i^k (W_i(x_i) - \frac{\Delta\lambda_i(\mathbf{x})}{f_i(x_i)})\big)\Big)$$

Similarly, the interim support function that incorporates BIC constraints can be written as

$$\mathcal{S}_{interim}^{BIC}(\mathbf{W}) = \min_{0 \leq \Lambda_i(x_i)} E_{\mathbf{x}} \Big(\max_{k \in \mathcal{K}} \big(0, \sum_{i \in \mathcal{I}} a_i^k (W_i(x_i) - \frac{\Delta\Lambda_i(x_i)}{f_i(x_i)})\big)\Big)$$

with $\Delta\Lambda_i(x_i^j) \equiv \Lambda_i(x_i^j) - \Lambda_i(x_i^{j-1})$ for $j = 1, \ldots, N_i$, and $\Lambda_i(x_i^0) = \Lambda_i(x_i^{N_i}) = 0$. ■

Proof of Proposition 3. Consider the minimization of the interim BIC support function (7) with respect to shifted weights $\widetilde{W}_i(x_i)$ for given weights $\widetilde{W}_j(x_j)$ of others $j \neq i$:

$$\min_{\mathbf{W}_i \succeq_{f_i} \widetilde{\mathbf{W}}_i} \sum_{x_i} f_i(x_i) G_i(\widetilde{W}_i(x_i))$$

where $G_i(y) = \sum_{\mathbf{x}_{-i}} f_{-i}(\mathbf{x}_{-i}) \max_{k \in \mathcal{K}} \big(0, a_i^k y + \sum_{j \neq i} a_j^k \widetilde{W}_j(x_j)\big)$ is a convex function of y. Recall from Lemma 1 that $\mathbf{W}_i^+$ solves this minimization problem. Repeating this argument for each agent $i = 1, \ldots, I$ yields $S^{BIC}_{interim}(\mathbf{W}) = S_{interim}(\mathbf{W}^+)$.

Next consider the minimization of the interim DIC support function (6) with respect to the shifted weights $\widetilde{W}_i(\mathbf{x}) = W_i(x_i) - \Delta\lambda_i(\mathbf{x})/f_i(x_i)$. Assume $\widetilde{W}_j(\mathbf{x}) = \widetilde{W}_j^+(x_j)$ for $\mathbf{x} \in \mathbb{R}^{|X|}$ and $j \neq i$, and consider the minimization problem with respect to agent i's shifted weights only. In other words, for each $\mathbf{x}_{-i}$, consider the minimization problem with respect to the vector $\widetilde{\mathbf{W}}_i(\cdot, \mathbf{x}_{-i})$, which satisfies $\mathbf{W}_i \succeq_{f_i} \widetilde{\mathbf{W}}_i(\cdot, \mathbf{x}_{-i})$:

$$\sum_{\mathbf{x}_{-i}} \min_{\mathbf{W}_i \succeq_{f_i} \widetilde{\mathbf{W}}_i(\cdot, \mathbf{x}_{-i})} \sum_{x_i} f_i(x_i) g_i(\widetilde{W}_i(x_i, \mathbf{x}_{-i}))$$

where $g_i(y) = f_{-i}(\mathbf{x}_{-i}) \max_{k \in \mathcal{K}} \big(0, a_i^k y + \sum_{j \neq i} a_j^k \widetilde{W}_j(x_j)\big)$ is a convex function of y. For each $\mathbf{x}_{-i}$, $\widetilde{\mathbf{W}}_i(\cdot, \mathbf{x}_{-i}) = \mathbf{W}_i^+$ solves the minimization problem. Therefore, one cannot lower the value by changing $\widetilde{W}_i(\mathbf{x})$ for agent i only when $\widetilde{W}_j(\mathbf{x}) = W_j^+(x_j)$ for $j \neq i$ are fixed. This implies that $W_i^+(x_i)$ is a local minimum of the interim DIC support function (6) and since the support function is convex it is also the global minimum (e.g. Rockafellar, 1997). Hence, the interim support functions coincide: $\mathcal{S}^{BIC}_{interim}(\mathbf{W}) = \mathcal{S}^{DIC}_{interim}(\mathbf{W}) = \mathcal{S}_{interim}(\mathbf{W}^+)$. ■

Proof of Lemma 3. Expected social surplus is given by

$$S \;=\; \sum_{i=1}^{I} E_{\mathbf{x}}(x_i v_i(\mathbf{x})) \;=\; \sum_{i=1}^{I} E_{x_i}(x_i V_i(x_i)) \;\equiv\; \mathbf{V} \circ \mathbf{x}$$

Expected revenue is

$$\begin{aligned}
R \;&=\; \sum_{i=1}^{I} \sum_{t=1}^{N_i} f_i(x_i^t) T_i(x_i^t) \\
&=\; \sum_{i=1}^{I} \sum_{t=1}^{N_i} f_i(x_i^t) V_i(x_i^t) x_i^t(\alpha_i^t) - \sum_{i=1}^{I} \sum_{t=1}^{N_i} \sum_{j=1}^{t-1} f_i(x_i^t) V_i(x_i^j)(x_i^{j+1}(\alpha_i^{j+1}) - x_i^j(\alpha_i^j)) \\
&=\; \sum_{i=1}^{I} \sum_{t=1}^{N_i} f_i(x_i^t) V_i(x_i^t) x_i^t(\alpha_i^t) - \sum_{i=1}^{I} \sum_{j=1}^{N_i} \sum_{t=j+1}^{N_i} f_i(x_i^t) V_i(x_i^j)(x_i^{j+1}(\alpha_i^{j+1}) - x_i^j(\alpha_i^j)) \\
&=\; \sum_{i=1}^{I} \sum_{t=1}^{N_i} f_i(x_i^t) V_i(x_i^t) \big(x_i^t(\alpha_i^t) - (x_i^{t+1}(\alpha_i^{t+1}) - x_i^t(\alpha_i^t)) \frac{1 - F_i(x_i^t)}{f_i(x_i^t)}\big) \;=\; \mathbf{V} \circ \mathbf{MR}(\boldsymbol{\alpha})
\end{aligned}$$

where $x_i^{N_i+1} = x_i^{N_i}$. ■

Proof of Proposition 4. From Proposition 3

$$\mathcal{S}^{BIC}_{interim}(\boldsymbol{\omega}) \;=\; \mathcal{S}_{interim}(\boldsymbol{\omega}^+) \;=\; E_{\mathbf{x}}\big(\max_k\,(0, \textstyle\sum_i a_i^k \omega_i^+(x_i))\big)$$

and from the basic definition of the interim support function

$$\mathcal{S}_{interim}(\boldsymbol{\omega}^+) \;=\; \max\big\{\textstyle\sum_k E_{\mathbf{x}}\big(q^k(\mathbf{x})\textstyle\sum_i a_i^k\omega_i^+(x_i)\big)\,|\,\mathbf{q}\text{ is feasible}\big\}$$

which establishes optimality of the allocation rule in equation (8). The payment rule in equation (9) follows from Lemma 2. ■

Proof of Proposition 5. The ex post support function is given by

$$\begin{aligned}\mathcal{S}_{ex\ post}(\mathbf{w}) &= \sum_{\mathbf{x}\in X}\max\big\{\textstyle\sum_i q_i(\mathbf{x})^{1-\gamma}w_i(\mathbf{x})\,|\,\textstyle\sum_i q_i(\mathbf{x}) \;=\; K\big\}\\ &= v(K)\sum_{\mathbf{x}\in X}\big(\textstyle\sum_i \max(0,w_i(\mathbf{x}))^{1/\gamma}\big)^{\gamma}\end{aligned}$$

The optimal allocation rule follows from Hotelling's lemma, i.e. $\mathbf{q} = \nabla\mathcal{S}_{ex\ post}(\mathbf{w})$:

$$q_i(\mathbf{x}) \;=\; K\,\frac{\max(0,w_i(\mathbf{x}))^{1/\gamma}}{\sum_j \max(0,w_j(\mathbf{x}))^{1/\gamma}},$$

where we interpret 0/0 as 0. The payment rule follows from Lemma 2. ■

Proof of Proposition 7. A solution to (10) must satisfy

$$\lambda_{m+1}(m+1)(\overline{x}-\underline{x}) \;=\; w_m + \lambda_m m(1-\frac{\gamma}{m})(\overline{x}-\underline{x})$$

for $m = \gamma-1,\ldots,I-1$. Define $\hat\lambda_m = m(\overline{x}-\underline{x})\lambda_m$ to obtain the recursive equation

$$\hat\lambda_{m+1} \;=\; w_m + \hat\lambda_m(1-\frac{\gamma}{m})$$

The support function can now be calculated as

$$\begin{aligned}S_{constrained}(\mathbf{w}) &= w_I + \hat\lambda_I(1-\frac{\gamma}{I})\\ &= w_I + \textstyle\sum_{m=1}^{I-\gamma} w_{I-m}\textstyle\prod_{j=0}^{m-1}(1-\dfrac{\gamma}{I-j})\\ &= \textstyle\sum_{m=\gamma}^{I-1} w_m \textstyle\prod_{j=m+1}^{I}(1-\dfrac{\gamma}{j}) + w_I\\ &= \textstyle\sum_{m=\gamma}^{I} w_m \binom{m}{\gamma}/\binom{I}{\gamma}\end{aligned}$$

The other statements follow from Fact 5 and direct calculations. ■

Proof of Proposition 8. The constrained support function is given by

$$\mathcal{S}_{constrained}(\mathbf{W}) \;=\; \inf_{0\le\lambda}\; E_{\mathbf{x}}\big(\max(0, W_b(x_b) - W_s(x_s) + \lambda(MR_b(x_b) - MR_s(x_s)))\big)$$

Hotelling's lemma $\mathbf{Q} = \nabla\mathcal{S}_{constrained}(\mathbf{W})$ evaluated at $W_b(x_b) = x_b$ and $W_s(x_s) = x_s$ yields[13]

$$\begin{aligned} Q_b(x_b) \;&=\; E_{x_s}(\mathbf{1}(MR_b^\lambda(x_b) > MR_s^\lambda(x_s)) \\ Q_s(x_s) \;&=\; E_{x_b}(\mathbf{1}(MR_b^\lambda(x_b) > MR_s^\lambda(x_s)) \end{aligned}$$

where $\mathbf{1}(\cdot)$ denotes an indicator function. The ex post allocation rule can now easily be read off since $Q_b(x_b) = E_{x_s}(q(x_b, x_s))$ and $Q_s(x_s) = E_{x_b}(q(x_b, x_s))$.

The optimal λ can be found by equating the derivative of the constrained support function with respect to λ to 0, which yields the condition

$$E_{MR_b^\lambda(x_b) > MR_s^\lambda(x_s)}\big(MR_b(x_b) - MR_s(x_s)\big) \;=\; 0$$

When $\lambda = 0$ the left side is equal to

$$-\sum_{j=1}^{N}(x^j - x^{j-1})F_s(x^{j-1})(1 - F_b(x^{j-1})) \;<\; 0,$$

Hence, fully efficient trade cannot occur. Likewise, no trade requires $\lambda = \infty$, but then the left side is strictly positive since $MR_b^\infty(x_b) = MR_b(x_b)$ and $MR_s^\infty(x_s) = MR_s(x_s)$. ■

Proof of Proposition 9. The support function for the $q^k(\mathbf{x})$ is

$$\mathcal{S}(\mathbf{w}) \;=\; \sum_{x\in X}\max_{k\in\mathcal{K}}(0, w_k(\mathbf{x}))$$

and applying Fact 4 yields the ex post support function for $v_i(\mathbf{x}) = \sum_k v_i^k(\mathbf{x})q^k(\mathbf{x})$:

$$\mathcal{S}_{ex\ post}(\mathbf{w}) \;=\; \sum_{\mathbf{x}\in X}\max_{k\in\mathcal{K}}(0, \sum_{i\in\mathcal{I}} v_i^k(\mathbf{x})w_i(\mathbf{x}))$$

The interim expected values are given by $V_i(\mathbf{x}_i) = \sum_{\mathbf{x}_{-i}} f(\mathbf{x}|\mathbf{x}_i)v_i(\mathbf{x})$ and applying Fact 4 once more yields

$$\mathcal{S}_{interim}(\mathbf{W}) \;=\; E_{\mathbf{x}}\big(\max_{k\in\mathcal{K}}(0, \sum_{i\in\mathcal{I}} v_i^k(\mathbf{x})W_i(\mathbf{x}_i))\big)$$

where, as before, we multiplied by $f_i(\mathbf{x}_i)$, i.e. we used the probability weighted inner product $\circ$ to define the set of interim expected values. ■

[13]Since the inner product $\circ$ is probability weighted, the components $\partial_{W_i(x_i)}$ of ∇ are multiplied by $1/f_i(x_i)$.

References

Ausubel, Lawrence M. and Peter Cramton (1998) "Demand Reduction and Inefficiency in Multi-Unit Auctions," Working paper, University of Maryland.

Bapat, R. B. (1991) "Majorization and Singular Values III, " *Linear Algebra and its Applications*, 145, 59–70.

Dasgupta, Patha S., Peter Hammond, and Eric Maskin (1979) "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility," *Review of Economic Studies*, 46, 181-216.

Fuchs, L. (1947) "A New Proof of an Inequality of Hardy-Littlewood-Polya," *Mat. Tidsskr. B*, 53–54.

Gershkov, Alex, Benny Moldovanu, and Xianwen Shi (2011) "Bayesian and Dominant Strategy Implementation Revisited," Working Paper, University of Bonn.

Gibbard, Allan (1973) "Manipulation of Voting Schemes: a General Result," *Econometrica*, 41, 587–601.

Goeree, Jacob K. and Alexey Kushnir (2011) "On the Equivalence of Bayesian and Dominant-Strategy Implementation in a General Class of Social Choice Problems," Working paper, University of Zürich.

Gutmann, Sam, J. H. B. Kemperman, J. A. Reeds, and L. A. Shepp (1991) "Existence of Probability Measures with Given Marginals," *Annals of Probability*, 19, 1781–1797.

Hayek, Friedrich A. (1945) "The Use of Knowledge in Society," *American Economic Review*, 35, 519–530.

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

Hernando-Veciana, Angel and Fabio Michelucci (2011) "Second-Best Efficiency and the English Auction," *Games and Economic Behavior*, 73, 496–506.

Hurwicz, Leonard (1960) "Optimality and Informational Efficiency in Resource Allocation Processes," in *Mathematical Methods in the Social Sciences*, K. J. Arrow et al., eds., Stanford University Press, Stanford, 27–46.

Hurwicz, Leonard (1972) "On Informationally Decentralized Systems," in *Decision and Organization*, C. B. McGuire and R. Radner, eds., North-Holland, Amsterdam, 297–336.

Mailath, George J. and Andrew Postlewaite (1990) "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57, 351–367.

Manelli, Alejandro M. and Daniel R. Vincent (2010) "Bayesian and Dominant-Strategy Implementation in the Independent Private Values Model," *Econometrica*, 78, 1905–1938.

Marshall, A. W., I. Olkin, I. and B. C. Arnold (2011) *Inequalities: Theory of Majorization and its Applications*, Springer, New York.

Maskin, Eric (1992) "Auctions and Privatization," in *Privatization: Symposium in honour of Herbert Giersh*, ed. by H. Siebert, Institute fur Weltwirtshaft an der Univarsitat Kiel.

Maskin, Eric and John Riley (1989) "Optimal Multi-Unit Auctions," in Frank Hahn, Editor, *The Economics of Missing Markets, Information, and Games*, Oxford: Oxford University Press, 312–335.

Mussa, Michael and Sherwin Rosen (1978) "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

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

Myerson, Roger B. and Mark A. Satterthwaite (1983) "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–281.

Rockafellar, R. Tyrrell (1997) *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press.