Ironing, Sweeping, and Multivariate Majorization
Optimal Mechanisms for Mass-Produced Goods

Nicholas C. Bedard, Jacob K. Goeree, and Ningyi Sun*

August 31, 2020

Abstract

We study the sale of an excludable, non-rival good by a monopolist when buyers' values are interdependent, i.e. buyers' utilities from consuming the good depend on their own and others' information. Most mass-produced goods fit this framework. We develop a constructive approach to deal with incentive compatibility constraints and thereby characterize the profit-maximizing mechanism. In particular, we exploit the structure of the Kuhn-Tucker conditions resulting from the constrained optimization problem to define a novel multivariate majorization concept. Our majorization technique allows us to generalize Mussa and Rosen's (1978) "ironing" to settings with multidimensional information. We also relate our majorization approach to Rochet and Choné's (1998) "sweeping" method. We illustrate how discriminatory access rights lead to constrained majorization, resulting in higher seller profits as well as a more efficient production of mass-produced goods.

Keywords: mechanism design, majorization, ironing, sweeping, mass-produced goods
JEL Classification Numbers: D42, D47

*Bedard: Wilfrid Laurier University, 75 University Avenue W. Waterloo ON N2L 3C5; Email: nbedard@wlu.ca. Goeree: AGORA Center for Market Design and the School of Economics, UNSW Australia Business School, Sydney NSW 2052, Australia; Email: j.goeree@unsw.edu.au. Sun: AGORA Center for Market Design and the School of Economics, UNSW Australia Business School, Sydney NSW 2052, Australia; Email: ningyi.sun@student.unsw.edu.au. We are grateful to seminar audiences at the 2016 APIOS in Melbourne, the University of Western Ontario, the University of Guelph and the University of Waterloo for insightful comments and discussions. We gratefully acknowledge funding from the Australian Research Council.

1. Introduction

Any seller of a mass-market good faces the challenge of designing a single good to appeal to a large number of diverse consumers. The decision of interest is not how many copies of a good to make but how to best design the template used to generate copies. In this paper, we view the good not as the physical product sold but the design of the template. This makes the good non-rivalrous: the template can be replicated indefinitely, or enjoyed by many people, without diminishing others' ability to enjoy further copies. Just as two people can enjoy the same public park if they are given access, two people can enjoy an identically designed smart phone if given the hardware. We study the problem of how to optimally design and sell an excludable, non-rivalrous good also known as "club goods." In particular, the seller chooses a single quality level for the good to be enjoyed commonly by all consumers of the good. In addition, the seller may restrict access to the good and collects transfers.

Virtually any product sold as identical goods to more than one person will fit into this framework. For example, producers of popular movies, books, television programming and music rely on wide appeal to generate profits, rather than finding the ideal fit for each customer. Schools must contend with how to design lectures and deliver courses and, as technology relaxes physical constraints to the learning environment, whom to exclude, if anyone. Mass-produced goods such as furniture, electronics, and some food and drinks have the same characteristic. Most restaurants offer consistent menus over time as customers flow through, certainly in the case of national and global chains. Many non-profit outfits such as library, schools, museums, theatres and orchestras, and public parks similarly satisfy large groups with limited offerings.

An important feature of the type of goods we consider is the mixture of both private and common-value components for buyers, i.e. preferences are interdependent.[1] Informally, information one buyer receives regarding the product may be relevant to other buyers. For example, buyer 1 may hear rumours about a new phone's processing speed and buyer 2 may hear about its screen size. Both attributes determine the value of the product for each buyer, possibly in different ways since buyers need not agree on what makes a product valuable – buyer 1 may prefer a small screen while buyer 2 prefers a large one.

Interdependent preferences present a technical challenge to the characterization of the optimal mechanism. Buyers receive signals about their valuations for the good but do not observe the signals of other buyers. The seller's problem is to choose the quality of the good in order

[1] Milgrom and Weber (1982) first introduced the notion of interdependent preferences in an auction environment. Subsequently, a number of papers have incorporated interdependent preferences into models of mechanism design, e.g. Jehiel et al. (1999); Jehiel and Moldovanu (2001); Figueroa and Skreta (2011), Csapó and Müller (2013) and Roughgarden and Talgam-Cohen (2016). In each of these papers, the good being sold is rivalrous.

to maximize profits in expectation over the possible preferences of the buyers. Importantly, the seller knows only the distribution of possible signals of the buyers and does not observe the realized signals. In order to choose the optimal selling mechanism, the seller must elicit information from the buyers about their willingness to pay for the good. Using standard arguments in mechanism design theory, this process can be done by imposing constraints on the seller's choice of the mechanism that incentivize the buyers to truthfully reveal their information.[2] These incentive compatibility constraints ensure that buyers optimally reveal their information truthfully. After some manipulation, the relevant constraint on the mechanism is that each buyer's payoff, net of any transfers to the seller, must be increasing in her private information.

In a pure private-values setting without common-value components to buyers' preferences, fairly mild assumptions on the preference distributions ensure that these monotonicity constraints do not bind and can be ignored. This so called "regular" environment is appealing because the objective function is typically linear or concave in the seller's choice variables and the optimal allocation can be easily derived from the first-order conditions. With interdependent values, however, the assumptions needed to make the environment regular become too restrictive. Moreover, the classic ironing technique of Mussa and Rosen (1978) cannot be immediately applied to an interdependent-values environment, which is naturally multidimensional.[3] The complication is that preferences are functions of many variables (i.e. the signals of all buyers) whereas the ironing technique depends on the unidimensionality of preferences.[4]

Using tools from majorization theory, we develop a constructive approach to multidimensional ironing.[5] We generalize the "ironing" approach of Mussa and Rosen (1978) by manipulating buyers' interdependent preferences to generate an alternative problem whose unconstrained solution is the same as the solution to the original constrained problem. We show that, by defining the appropriate concept of multivariate majorization, we can find the needed ironed preferences through a simple quadratic minimization problem. Our definition for multivariate majorization is suggested by the structure of the Kuhn-Tucker conditions governing the constrained problem.

[2] A choice of mechanism involves choosing the quality of the good, a set of access rights for the buyers, and a set of transfers.

[3] An additional complication arises since we make no assumption on the additive separability of other buyers' signals in a buyer's preferences. If, for example, all buyers' valuations are sums of buyers' signals then a slight modification of the ironing technique can deal with the constraints: one irons in each dimension holding all other dimensions fixed, i.e. in dimension of buyer i's signal holding all others' signals constant. This works because with such preferences, the seller's problem is mathematically identical to one with private values. More generally, as in the environment we study, each dimension of preferences have to be manipulated simultaneously disallowing any use of the univariate ironing technique.

[4] Roughgarden and Talgam-Cohen (2016), for example, note that a naive generalization of ironing in a model with interdependent preferences fails to properly account for incentive-compatibility constraints.

[5] The connection between one-dimensional ironing and univariate majorization, in particular the work of Hardy et al. (1929), was first noted by Goeree and Kushnir (2011).

Rochet and Choné (1998) introduce the concept of "sweeping" to deal with multidimensional incentive-compatibility constraints.[6] Roughly, sweeping is an operation that redistributes the density of a measure in a way that preserves the original center-of-mass. Rochet and Choné (1998) do not use sweeping to construct a solution but instead its use is to verify whether a candidate incentive-compatible solution is optimal. Specifically, their results say that if the derivative of the first-order condition of the seller's unconstrained problem, evaluated at the candidate solution, is non-zero but can be "swept" to zero then the candidate solution is optimal.[7] We discuss how Rochet and Choné's (1998) sweeping method is related to our majorization approach.

Besides our interpretation of the seller choosing design quality, one can also view our model as the private provision of a public good with private information, with the seller choosing the quantity of the public good to provide. The analysis of this problem in a mechanism design framework goes back at least as far as Groves and Ledyard (1977). Classically, the decision is binary, i.e. to provide the good or not, and participation is compulsory, i.e. the good is non-excludable. The relevant question is whether the good is provided efficiently (see Güth and Hellwig, 1986; Malaith and Postlewaite, 1990; Csapó and Müller, 2013).[8] More recently, exclusions and continuous levels of the good have been allowed (see Cornelli, 1996; Ledyard and Palfrey, 1999; Hellwig, 2003; Norman, 2004; Ledyard and Palfrey, 2007). With the exception of Csapó and Müller (2013), preferences in the papers cited above are purely private valued.

This paper is organized as follows. Section 2 describes the environment and defines the seller's problem. Section 3 introduces our novel definition of multivariate majorization. Section 4 derives the optimal mechanism when all buyers have access (public goods) and when access is restricted (club goods). Section 5 discusses applications and extensions. Section 6 concludes. Proofs can be found in Appendix A.

[6]In their environment, a monopolist sells to a consumer with multidimensional preferences, each dimension corresponding to an attribute of the good designed by the seller.

[7]The intuition is that the first-order condition of the unconstrained problem evaluated at the candidate solution measures how valuable it is for the seller to deviate from the candidate solution. The fact that a mass-preserving redistribution of this value is zero means that any marginal profit the seller gains from a change from the candidate solution at some initial point will be exactly offset by a loss incurred through the change in the candidate solution at another point made necessary, via incentive compatibility, by the change at the initial point.

[8]Csapó and Müller (2013) study the private provision of a public good from a computer science perspective. They find that the solution can be solved in polynomial in the number of agents and type profiles.

2. Model

There is a set $N = \{1, \dots, n\}$ of buyers with preferences over a good to be provided by a seller. The seller chooses: (i) the quality $q \in \mathbb{R}_{\geq 0}$ of a product to supply; (ii) access rights $\boldsymbol{\eta} = \{\eta_1, \dots, \eta_n\} \in [0,1]^n$ where η_i is the probability that buyer i will be allowed to consume the good; and (iii) transfers $\mathbf{t} = \{t_1, \dots, t_n\} \in \mathbb{R}^n$ where t_i is the amount to collect from buyer i. Under the restriction $\eta_i = 1$ for all $i \in N$ the good is not excludable, i.e. it is a public good.

Buyers each receive a private signal regarding the value of the good. We refer to a buyer's signal as her type. Buyer i's type x_i is drawn from X_i according to distribution $F_i(x_i)$ with probability function (or density if X_i is continuous) $f_i(x_i)$. We make no assumption on X_i or F_i. Let $\underline{x}_i = \min X_i$ and $\overline{x}_i = \max X_i$. For a profile of types we write $\mathbf{x} = \{x_i, \mathbf{x}_{-i}\}$ with $\mathbf{x}_{-i} = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$. The set of all possible type profiles is $X = \prod_{i \in N} X_i$ with $X_{-i} = \prod_{i \in N \setminus \{i\}} X_i$. We assume types are drawn independently across bidders, i.e. $\mathbf{x}$ is drawn according to $f(\mathbf{x}) = \prod_{i \in N} f_i(x_i)$. For arbitrary $g : X \to \mathbb{R}$ we define the partial derivatives $\underline{\Delta}_i g(\mathbf{x}) = g(x_i, \mathbf{x}_{-i}) - g(x_i^-, \mathbf{x}_{-i})$ and $\overline{\Delta}_i g(\mathbf{x}) = g(x_i^+, \mathbf{x}_{-i}) - g(x_i, \mathbf{x}_{-i})$ where x_i^+ (x_i^-) is the type just above (below) x_i with the convention that $\underline{\Delta}_i g(\underline{x}_i, \mathbf{x}_{-i}) = g(\underline{x}_i, \mathbf{x}_{-i})$ and $\overline{\Delta}_i g(\bar{x}_i, \mathbf{x}_{-i}) = 0$. (In the case of continuous types, $\underline{\Delta}_i g(\mathbf{x}) = \overline{\Delta}_i g(\mathbf{x}) = \partial_{x_i} g(\mathbf{x})$.)

Buyer i's payoff is quasilinear. Her valuation for the good depends on her own signal as well as the signals of all other buyers, i.e. valuations are interdependent. In particular, her valuation for the good given type profile $\mathbf{x} \in X$ is $v_i(\mathbf{x})q(\mathbf{x})$, where $v_i : X \to \mathbb{R}$ is non-decreasing in x_i for all $\mathbf{x}_{-i} \in X_{-i}$. Buyer i's payoff, given choices $(q, \boldsymbol{\eta}, \boldsymbol{t})$ by the seller and type profile $\mathbf{x} \in X$ is $u_i(q, \boldsymbol{\eta}, \boldsymbol{t}; \mathbf{x}) = v_i(\mathbf{x})q(\mathbf{x})\eta_i(\mathbf{x}) - t_i(\mathbf{x})$.[9]

The seller's problem is to choose a mechanism to maximize the expected sum of transfers from buyers net of the expected cost. The cost of providing quantity q is $C(q)$, which is assumed to be an increasing convex (and differentiable) function with $C(0) = C'(0) = 0$. Due to the revelation principle, we can focus on (incentive compatible) direct mechanisms: $(q, \boldsymbol{\eta}, \boldsymbol{t})$ where (i) $q : X \to \mathbb{R}_{\geq 0}$ maps type profiles into quantity choices; (ii) $\boldsymbol{\eta} : X \to [0,1]^n$ maps type profiles into access rights; and (iii) $\mathbf{t} : X \to \mathbb{R}^n$ maps type profiles into transfers.

Let $\mathbb{E}[g(\mathbf{x})] = \sum_{x \in X} f(\mathbf{x})g(\mathbf{x})$. We sometimes write $||g||^2 = \mathbb{E}[g(\mathbf{x})^2]$ for the squared norm of g and $\langle g \,|\, h \rangle = \mathbb{E}[g(\mathbf{x})h(\mathbf{x})]$ for the inner-product between g and h. The seller's problem is therefore to choose a direct mechanism $(q, \boldsymbol{\eta}, \boldsymbol{t})$ to maximize

$$\mathbb{E}\Big[\sum_{i \in N} t_i(\mathbf{x}) - C(q(\mathbf{x}))\Big]$$

[9]More generally, payoffs could depend nonlinearly on q, for example $u_i(q, \boldsymbol{\eta}, \boldsymbol{t}; \mathbf{x}) = v_i(\mathbf{x})w(q)$ for some increasing, concave w. This is equivalent to our formulation up to a change in units for costs.

subject to (ex post) incentive compatibility, i.e. for all $i \in N$ and $\mathbf{x} \in X$

$$x_i \in \underset{\hat{x}_i \in X_i}{\operatorname{argmax}}\, u_i\left(q(\hat{x}_i, \mathbf{x}_{-i}), \eta_i(\hat{x}_i, \mathbf{x}_{-i}), t_i(\hat{x}_i, \mathbf{x}_{-i}); \mathbf{x}\right)$$

and individual rationality, i.e. $u_i(q(\mathbf{x}), \eta_i(\mathbf{x}), t_i(\mathbf{x}); \mathbf{x}) \geq 0$ for all $i \in N$ and $\mathbf{x} \in X$.

Proposition 1 *A direct mechanism* $(q, \boldsymbol{\eta}, \boldsymbol{t})$ *is incentive compatible and individually rational if and only if, for all* $i \in N$*,* $\mathbf{x} \in X$*,* $q(\mathbf{x})\eta_i(\mathbf{x})$ *non-decreasing in* x_i *and*

$$t_i(\mathbf{x}) \;=\; v_i(\mathbf{x})q(\mathbf{x})\eta_i(\mathbf{x}) - \sum_{s_i < x_i} \overline{\Delta}_i v_i(s_i, \mathbf{x}_{-i})q(s_i, \mathbf{x}_{-i})\eta_i(s_i, \mathbf{x}_{-i}) \tag{1}$$

The proof of this proposition is standard and is therefore omitted.[10]

Using equation (1), we can rewrite the seller's profit as

$$\Pi(\boldsymbol{MR}, q, \boldsymbol{\eta}) \;=\; \mathbb{E}\left[\, q(\mathbf{x}) \sum_{i \in N} MR_i(\mathbf{x})\eta_i(\mathbf{x}) - C(q(\mathbf{x}))\,\right] \tag{2}$$

where $\boldsymbol{MR} = \{MR_1, \ldots, MR_n\}$ and

$$MR_i(\mathbf{x}) \;=\; v_i(\mathbf{x}) - \frac{1 - F_i(x_i)}{f_i(x_i)}\, \overline{\Delta}_i v_i(\mathbf{x}) \tag{3}$$

Proposition 1 allows us to restate the seller's problem: choose the quality, $q(\mathbf{x})$, and access rights, $\boldsymbol{\eta}(\mathbf{x})$, to maximize (2) such that $\eta_i(\mathbf{x})q(\mathbf{x})$ is non-decreasing in x_i for all $i \in N$, $\mathbf{x} \in X$ and set transfers, $\boldsymbol{t}(\mathbf{x})$, according to (1). Let Π^* denote the seller's optimal profits when using the optimal mechanism $(q^*, \boldsymbol{\eta}^*, \mathbf{t}^*)$.

The unconstrained solution to the seller's problem is simple to characterize: set $\eta_i(\mathbf{x}) = 1$ if $MR_i(\mathbf{x}) > 0$ and $\eta_i(\mathbf{x}) = 0$ otherwise and choose $q(\mathbf{x})$ such that $C'(q(\mathbf{x})) = \sum_{i \in N} MR_i(\mathbf{x})\eta_i(\mathbf{x})$. Such a mechanism will be incentive compatible, i.e. $q(\mathbf{x})\eta_i(\mathbf{x})$ will be non-decreasing in x_i for all $i \in N$, $\mathbf{x} \in X$, as long as $\sum_{i \in N} MR_i(\cdot)\eta_i(\cdot)$ is sufficiently well behaved. We want to maintain the simple structure of the solution but do not want to be bound by the assumptions needed to make the parameters well behaved. Instead, we seek to manipulate the parameters $\boldsymbol{MR}$ in order to generate an equivalent unconstrained problem. That is, we are looking for $\overline{\boldsymbol{MR}}$ such that the solution $(\overline{q}, \overline{\boldsymbol{\eta}})$ to the unconstrained problem (ignoring incentive constraints)

$$(\overline{q}, \overline{\boldsymbol{\eta}}) \;=\; \underset{(q,\boldsymbol{\eta}) : X \to \mathbb{R}_{\geq 0} \times [0,1]^n}{\operatorname{argmax}} \Pi(\overline{\boldsymbol{MR}}; q, \boldsymbol{\eta})$$

is also the solution to the full problem: $\Pi(\boldsymbol{MR}, \overline{q}, \overline{\boldsymbol{\eta}}) \;=\; \Pi^*$. To construct the sought after $\overline{\boldsymbol{MR}}$, we expand on a concept in mathematics called majorization.

[10] The utility of the lowest type is set to zero as is standard in the context of a revenue-maximizing seller.

3. Multivariate Majorization

It is instructive to start with the case where all buyers have access to the public good, i.e. $\eta_i = 1$ for all $i \in N$, and costs are quadratic. The seller's problem

$$\Pi \;=\; \max_{\substack{q\,:\,X \to \mathbb{R}_{\geq 0} \\ q(\mathbf{x}) \text{ non-decreasing}}} \mathbb{E}\Big[q(\mathbf{x})MR(\mathbf{x}) - \frac{1}{2}q(\mathbf{x})^2\Big]$$

where $MR(\mathbf{x}) = \sum_{i \in N} MR_i(\mathbf{x})$, can be solved using standard Kuhn-Tucker methods. To deal with the constraint that $q(\mathbf{x})$ be non-decreasing in x_i for all $i \in N$ we add $\sum_{i \in N,\, \mathbf{x} \in X} \lambda_i(\mathbf{x})\overline{\Delta}_i q(\mathbf{x})$ to the objective where, for all $i \in N$, the $\lambda_i(\mathbf{x})$ are non-negative for all $\mathbf{x} \in X$ with $\lambda_i(\overline{x}_i, \mathbf{x}_{-i}) = 0$ for all $\mathbf{x}_{-i} \in X_{-i}$. This term can rewritten to obtain the saddle-point problem

$$\Pi \;=\; \min_{\boldsymbol{\lambda}\,:\,X \to \mathbb{R}_{\geq 0}} \; \max_{q\,:\,X \to \mathbb{R}_{\geq 0}} \mathbb{E}\Big[q(\mathbf{x})MR(\mathbf{x}, \boldsymbol{\lambda}(\mathbf{x})) - \frac{1}{2}q(\mathbf{x})^2\Big] \tag{4}$$

where $MR(\mathbf{x}, \boldsymbol{\lambda}(\mathbf{x})) = \sum_{i \in N} MR_i(\mathbf{x}, \lambda_i(\mathbf{x}))$ and

$$MR_i(\mathbf{x}, \lambda_i(\mathbf{x})) \;=\; MR_i(\mathbf{x}) - \underline{\Delta}_i \lambda_i(\mathbf{x})/f(\mathbf{x}) \tag{5}$$

For $\mathbf{x}, \mathbf{y} \in X$ we write $\mathbf{y} \leq \mathbf{x}$ ($\mathbf{y} < \mathbf{x}$) if $y_i \leq x_i$ ($y_i < x_i$) for $i \in N$.

To solve for the constraint coefficients, we derive a multivariate extension of majorization. We start with the multivariate version of a lower sum.

Definition 1 *A closed subset $X_- \subseteq X$ is a* ***lower set*** *of X if $\mathbf{x} \in X_-$ and $\mathbf{y} \in X$ with $\mathbf{y} \leq \mathbf{x}$ implies $\mathbf{y} \in X_-$.*

For instance, each of the subsets in the left panel of Figure 1 is an example of a lower set, while neither subset in the right panel is. Let $\overline{\partial} X_-$ denote the upper boundary of X_-: $\mathbf{x} \in \overline{\partial} X_-$ if there does not exist $\mathbf{y} \in X_-$ such that $\mathbf{y} > \mathbf{x}$. For instance, in the left panel of Figure 1, the upper boundaries of X_-, X'_-, and X''_- coincide with their borders in the interior of X plus the points where the interior border intersects the axes. Let $\overline{\partial}_i X_-$ denote the upper boundary for buyer $i \in N$: $\mathbf{x} \in \overline{\partial}_i X_-$ if $\mathbf{x} \in \overline{\partial} X_-$ and there does not exist $(x'_i, \mathbf{x}_{-i}) \in \overline{\partial} X_-$ with $x'_i > x_i$. For X_- and X'_-, buyers' upper boundaries coincide with the upper boundary, while for X''_-, buyer 1's (2's) upper boundary consists of all the vertical (horizontal) segments of the upper boundary.

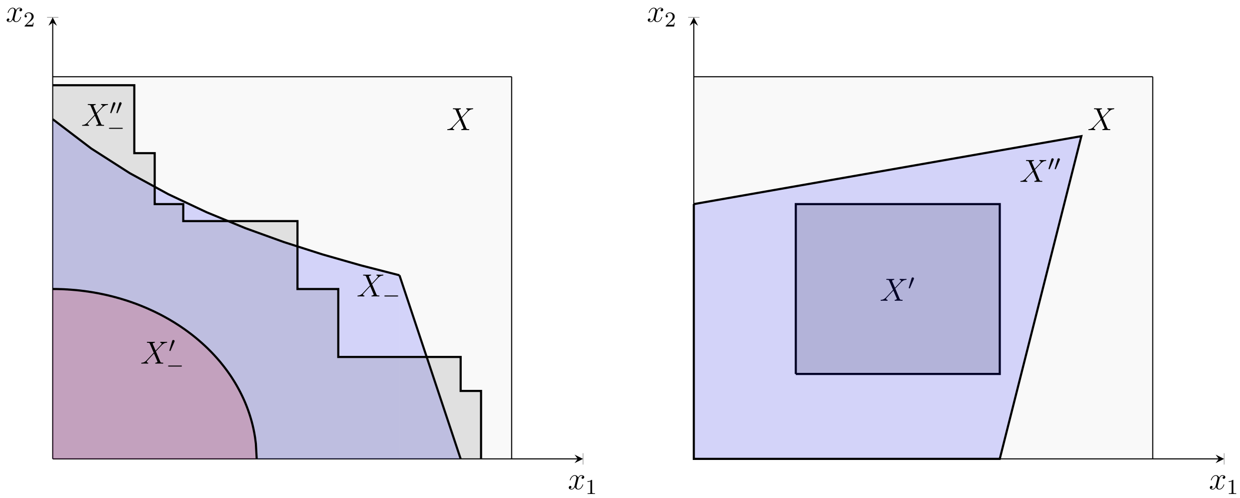

Figure 1: In the left panel X_-, X'_- and X''_- are all lower sets of X. In the right panel, X' and X'' are not lower sets of X.

Since $\sum_{s\leq x_i} \underline{\Delta}_i \lambda_i(s, \mathbf{x}_{-i}) = \lambda_i(x_i, \mathbf{x}_{-i})$ for any $\mathbf{x}_{-i} \in X_{-i}$, (5) implies, for any lower set X_-,

$$\begin{aligned}\mathbb{E}\big[MR(\mathbf{x}, \boldsymbol{\lambda}(\mathbf{x}))\,|\,\mathbf{x} \in X_-\big] &= \mathbb{E}\big[MR(\mathbf{x}) - \sum_{i\in N} \underline{\Delta}_i \lambda_i(\mathbf{x})/f(\mathbf{x})\,|\,\mathbf{x} \in X_-\big] \\ &= \mathbb{E}\big[MR(\mathbf{x})\,|\,\mathbf{x} \in X_-\big] - \frac{1}{F(X_-)} \sum_{i\in N} \sum_{\mathbf{x}\in \overline{\partial}_i X_-} \lambda_i(\mathbf{x}) \\ &\leq \mathbb{E}\big[MR(\mathbf{x})\,|\,\mathbf{x} \in X_-\big] \qquad (6)\end{aligned}$$

where $F(X_-) = \sum_{\mathbf{x}\in X_-} f(\mathbf{x})$. Equation (6) holds with equality if, for $i \in N$, $\lambda_i(\mathbf{x})$ vanishes on $\overline{\partial}_i X_-$. In particular, since $\lambda_i(\bar{x}_i, \mathbf{x}_{-i}) = 0$ for all $\mathbf{x}_{-i} \in X_{-i}$ and $i \in N$, we have

$$\mathbb{E}\big[MR(\mathbf{x}, \boldsymbol{\lambda}(\mathbf{x}))\big] = \mathbb{E}\big[MR(\mathbf{x})\big] \qquad (7)$$

Together (6) and (7) define the multivariate extension of majorization to n-dimensional functions.

Definition 2 *For $g : X \to \mathbb{R}$, $h : X \to \mathbb{R}$, g **majorizes** h, denoted $g \succ h$, if for any lower set $X_- \subset X$ we have $\mathbb{E}[g(\mathbf{x})|\mathbf{x} \in X_-] \leq \mathbb{E}[h(\mathbf{x})|\mathbf{x} \in X_-]$ and $\mathbb{E}[g(\mathbf{x})] = \mathbb{E}[h(\mathbf{x})]$.*

Remark 1 Alternative definitions of multivariate majorization exist in the mathematics literature, e.g. row-majorization, column-majorization, and linear-combinations majorization (see Marshall et al., 2010, Chapter 15). None of these alternative definitions are equivalent to Definition 2 nor do they originate from a Kuhn-Tucker like program. ■

We will also use a notion of univariate majorization that applies to multidimensional functions. For the univariate case, the lower sets are simply (discrete) intervals starting at the lowest type.

Definition 3 *For* $g : X \to \mathbb{R}$, $h : X \to \mathbb{R}$, g ***majorizes*** h ***in coordinate*** i*, denoted* $g \succ_i h$*, if for all* $\mathbf{s} \in X$, $\mathbb{E}[g(\mathbf{x}) \,|\, x_i \le s_i, \mathbf{x}_{-i} = \mathbf{s}_{-i}] \le \mathbb{E}[h(\mathbf{x}) \,|\, x_i \le s_i, \mathbf{x}_{-i} = \mathbf{s}_{-i}]$ *with equality if* $s_i = \overline{x}_i$*.*

We will use below that if $g_i \succ_i h_i$ for $i \in N$ then $\sum_{i\in N} g_i \succ \sum_{i\in N} h_i$.

Remark 2 The univariate-majorization literature typically considers only *non-decreasing* functions and defines the preorder $\succ$ in terms of *upper* sets, i.e. (discrete) intervals that end at the highest type. The latter difference is immaterial because of (7). To align with the literature we could reorder functions so they are non-decreasing. However, this would make comparison with Mussa and Rosen's (1978) ironing impossible. Hence, we take the (possibly non-monotonic) MR functions as exogenously given and define $\succ$ for all functions including non-monotonic ones. ■

3.1. Multivariate Majorization and Doubly Stochastic Matrices

In the majorization literature, the $-\underline{\Delta}_i \lambda_i(\mathbf{x})$ term in (5) is known as a "Dalton transfer." It involves taking $\lambda_i(\mathbf{x}) \ge 0$ from type $(x_i, \mathbf{x}_{-i})$ and passing it to a higher type $(x_i^+, \mathbf{x}_{-i})$, which is why it is also known as an *anti* "Robin Hood" transfer. A "Robin Hood" involves transferring $\lambda_i(\mathbf{x}) \ge 0$ from the higher type $(x_i^+, \mathbf{x}_{-i})$ to $(x_i, \mathbf{x}_{-i})$, which amounts to reversing the sign of the shift in (5) and defining $MR_i(\mathbf{x}, \boldsymbol{\lambda}(\mathbf{x})) = MR_i(\mathbf{x}) + \underline{\Delta}_i \lambda_i(\mathbf{x})/f(\mathbf{x})$.

Equivalently, we could transform the marginal revenues using proportional transfers. A well-known concept in the univariate-majorization literature is the T-transform, which is a doubly stochastic operator of the form $T = \alpha I + (1-\alpha)P$ where I is the identity, P a permutation that interchanges only two elements, and $0 \le \alpha \le 1$. The image of a univariate sequence $(g_1, \ldots, g_{|X|})$ under a T-transform is $(g_1, \ldots, \alpha g_i + (1-\alpha) g_j, \ldots, \alpha g_j + (1-\alpha) g_i, \ldots, g_{|X|})$, i.e. there is a transfer $(1-\alpha)(g_j - g_i)$ from type j to i. If g represents "wealth" and is non-decreasing in type then this amounts to a non-negative transfer from a "wealthier" to a "poorer" type (i.e. it is a "Robin Hood" transfer). Muirhead (1903) showed that for non-decreasing functions g and h, a necessary and sufficient condition for $g \succ h$ is that h can be obtained from g via a series of T-transforms.

A similar result can be obtained for functions defined over a multidimensional type set X if we restrict transfers to be between type profiles that differ only for a single buyer, i.e. between $(x_i, \mathbf{x}_{-i})$ and $(x_i', \mathbf{x}_{-i})$ for some $i \in N$, $x_i, x_i' \in X_i$, and $\mathbf{x}_{-i} \in X_{-i}$. We refer to T-transforms of this type as *orthogonal* T-transforms.

Proposition 2 *Let* $g(\mathbf{x})$ *and* $h(\mathbf{x})$ *be non-decreasing in each coordinate. Then* $g \succ h$ *if and only if* $h(\mathbf{x})$ *can be obtained from* $g(\mathbf{x})$ *via a series of orthogonal* T*-transforms.*

For the univariate case, Hardy et al. (1929) sharpened Muirhead's (1903) result to $g \succ h$ if and only if $h = S \cdot g$ for some doubly stochastic matrix S. But in the multivariate case, the orthogonality requirement restricts the possible doubly-stochastic transformations.[11]

3.2. Minimal Elements

Let us return to the seller's problem in (4). The (unconstrained) maximization over $q(\mathbf{x})$ yields $q(\mathbf{x}) = \max(0, \overline{MR}(\mathbf{x}))$ and $\Pi = \frac{1}{2}\,\mathbb{E}[q(\mathbf{x})^2]$ where

$$\overline{MR}(\mathbf{x}) \;=\; \underset{\substack{g\,:\,X\to\mathbb{R}\\ g\,\succ\,MR}}{\operatorname{argmin}}\;\mathbb{E}\big[g(\mathbf{x})^2\big] \tag{8}$$

The next lemma shows why we can take the objective to be $\mathbb{E}[g(\mathbf{x})^2]$ rather than $\mathbb{E}[\max(0, g(\mathbf{x}))^2]$.

Lemma 1 *If $\overline{MR}(\mathbf{x})$ satisfies (8) then*

$$\overline{MR}(\mathbf{x}) \;=\; \underset{\substack{g\,:\,X\to\mathbb{R}\\ g\,\succ\,MR}}{\operatorname{argmin}}\;\mathbb{E}\big[\phi(g(\mathbf{x}))\big]$$

for any convex function $\phi : \mathbb{R} \to \mathbb{R}$, e.g. $\phi(x) = \max(0, x)^2$.

An easy corollary is that if the quadratic cost is replaced by a general convex cost function, $C(q)$, the same solution for $\overline{MR}(\mathbf{x})$ applies. The optimal quality is $q(\mathbf{x}) = C'^{-1}(\max(0, \overline{MR}(\mathbf{x})))$ and $\Pi = \mathbb{E}\big[\phi(\max(0, \overline{MR}(\mathbf{x})))\big]$, where $\phi(y) = yC'^{-1}(y) - C(C'^{-1}(y))$ is convex if $C(\cdot)$ is.

From (8), $\overline{MR}$ is the smallest function, with respect to the norm $||g|| = \sqrt{\mathbb{E}[g(\mathbf{x})^2]}$, that majorizes MR. It is also "smallest" with respect to the pre-order $\prec$ because if there existed $MR \prec g \prec \overline{MR}$ then g can be obtained from $\overline{MR}$ via series of orthogonal T-transforms (see Proposition 2). But then g would yield a lower objective in (8). We say that $\overline{MR}$ is a *minimal element* with respect to $\prec$ if $MR \prec g \prec \overline{MR}$ implies $g = \overline{MR}$.

Proposition 3 *There exists a solution $\overline{MR}(\mathbf{x})$ to (8). This solution is unique, non-decreasing, and a minimal element with respect to $\prec$.*

Unlike the univariate case, there may exist multiple minimal elements in the multivariate case.[12]

[11] Consider, for instance, $g = \left(\begin{smallmatrix} 0 & 2 \\ 4 & 6 \end{smallmatrix}\right)$ where the rows (columns) correspond to buyer 1's (2's) types. If h is obtained from g by averaging along columns or rows then $g \succ h$ but not if we average along the diagonal, i.e. $\left(\begin{smallmatrix} 0 & 2 \\ 4 & 6 \end{smallmatrix}\right) \succ \left(\begin{smallmatrix} 1 & 1 \\ 4 & 6 \end{smallmatrix}\right)$ but $\left(\begin{smallmatrix} 0 & 2 \\ 4 & 6 \end{smallmatrix}\right) \not\succ \left(\begin{smallmatrix} 0 & 3 \\ 3 & 6 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 2 \\ 4 & 6 \end{smallmatrix}\right) \not\prec \left(\begin{smallmatrix} 0 & 3 \\ 3 & 6 \end{smallmatrix}\right)$.

[12] Goeree and Kushnir (2011) show for the univariate case that $\overline{MR}(\mathbf{x})$ is the unique minimum with respect to $\prec$.

Example 1 Suppose there are two agents with equally-likely types in $X_1 = X_2 = \{0, 1\}$ and value functions $v_i(x_i, x_{-i}) = 3 + 3x_i x_{-i} - 3x_{-i}$ for $i = 1, 2$. Then the marginal revenues are

$$MR_1 = \begin{pmatrix} 3 & -3 \\ 3 & 3 \end{pmatrix} \qquad \text{and} \qquad MR_2 = \begin{pmatrix} 3 & 3 \\ -3 & 3 \end{pmatrix}$$

where the rows (columns) correspond to buyer 1's (2's) types. The program in (8) entails minimizing the sum of the squared entries of

$$\begin{pmatrix} 6 - \lambda_1 - \mu_1 & -\lambda_2 + \mu_1 \\ \lambda_1 - \mu_2 & 6 + \lambda_2 + \mu_2 \end{pmatrix}$$

with respect to non-negative λs (for player 1) and μs (for player 2), which yields $\lambda_1 = \mu_1 = 2$ and $\lambda_2 = \mu_2 = 0$ so[13]

$$\overline{MR} = \begin{pmatrix} 2 & 2 \\ 2 & 6 \end{pmatrix}$$

For $0 \leq \alpha < 1$,

$$g(\alpha) = \begin{pmatrix} 2\alpha & 2\alpha \\ 6 - 4\alpha & 6 \end{pmatrix}$$

is non-decreasing and satisfies $g(\alpha) \succ MR$ but not $g(\alpha) \succ \overline{MR}$. The same is true for the transpose of $g(\alpha)$. It is readily verified that $g(\alpha)$ and its transpose are minimal elements with respect to $\succ$ for $\alpha \in [0, 1]$. (And they are extreme points only if $\alpha \in \{0, 1\}$.) ■

3.3. Characterizing the Ironed Marginal Values

In this section, we provide a general expression for the ironed marginal revenues $\overline{MR}(\mathbf{x})$.

Definition 4 *$S \subseteq X$ is ortho-convex if $(x_i, \mathbf{x}_{-i}) \in S$ and $(x'_i, \mathbf{x}_{-i}) \in S$ implies $(x, \mathbf{x}_{-i}) \in S$ for all $\min(x_i, x'_i) \leq x \leq \max(x_i, x'_i)$, $i \in N$, $\mathbf{x}_{-i} \in X_{-i}$. A partition $\mathscr{P}$ of X is ortho-convex if all its cells are ortho-convex.*

An alternative definition is that $S \subseteq X$ is ortho-convex if for any line parallel to one of the axes, the intersection with S is empty, a point, or a single segment. (Akin to the definition of convex sets for which the intersection with *any* line is empty, a point, or a single segment.) The reason for considering ortho-convex, rather than convex, partitions is that the incentive-compatibility constraints for player $i \in N$ impose restrictions only along the x_i coordinate.[14] As a result, the ironed sets are convex along each orthogonal but they are not necessarily convex subsets of the n-dimensional type space X (although they could be for certain parameter values).

[13]Note that one cannot simply iron the marginal revenues of each player separately and then add them back together, as pointed out by Roughgarden and Talgam-Cohen (2016). In this example, majorizing the MR_i separately would leave them unchanged and their sum would not be non-decreasing.

[14]And player i's equilibrium utility $u_i(q(\mathbf{x}), \eta_i(\mathbf{x}), t_i(\mathbf{x}); \mathbf{x})$ is convex only in x_i.

Proposition 4 *The level sets of the solution $\overline{MR}(\mathbf{x})$ to (8) form an ortho-convex partition $\mathscr{P}$ of X. For $\mathbf{x} \in X$, $P(\mathbf{x}) = \{\mathbf{y} \,|\, \overline{MR}(\mathbf{y}) = \overline{MR}(\mathbf{x})\}$ is the cell of $\mathscr{P}$ containing $\mathbf{x}$ and $\overline{MR}(\mathbf{x}) = \mathbb{E}[MR(\mathbf{y}) \,|\, \mathbf{y} \in P(\mathbf{x})]$. Moreover,*

$$S(\mathbf{x}, \mathbf{y}) \;=\; \frac{f(\mathbf{y})}{F(P(\mathbf{x}))}\, \delta_{\mathbf{y} \in P(\mathbf{x})} \tag{9}$$

is a stochastic operator that maps MR to $\overline{MR}$.

For Example 1, the partition is $\mathscr{P} = P_1 \sqcup P_2$ with $P_1 = \{(0,0),(1,0),(0,1)\}$ and $P_2 = \{(1,1)\}$, and the stochastic operator is $S(\mathbf{x},\mathbf{y}) = \frac{1}{3}\delta_{\mathbf{x},\mathbf{y} \in P_1} + \delta_{\mathbf{x},\mathbf{y} \in P_2}$.

3.4. Alternative Formulations

We end this section with an alternative to the program in (8). Recall that $||g|| = \sqrt{\mathbb{E}[g(\mathbf{x})^2]}$.

Proposition 5 *The solution $\overline{MR}(\mathbf{x})$ to (8) also follows from*

$$\overline{MR}(\mathbf{x}) \;=\; \underset{\substack{g \,:\, X \,\to\, \mathbb{R} \\ g(\mathbf{x}) \text{ non-decreasing}}}{\text{argmin}} \;\; ||MR(\mathbf{x}) - g(\mathbf{x})|| \tag{10}$$

or, equivalently,

$$\overline{MR}(\mathbf{x}) \;=\; MR(\mathbf{x}) - \underset{\substack{g \,:\, X \,\to\, \mathbb{R} \\ MR(\mathbf{x}) - g(\mathbf{x}) \text{ non-decreasing}}}{\text{argmin}} \;\; ||g(\mathbf{x})|| \tag{11}$$

The program in (10) establishes $\overline{MR}(\mathbf{x})$ as the non-decreasing function closest to $MR(\mathbf{x})$. The program in (11) shows that $\overline{MR}(\mathbf{x})$ is obtained from $MR(\mathbf{x})$ by subtracting the smallest function so that the result is non-decreasing.

For Example 1, the solution $\overline{g}$ to (11) is

$$\overline{g} \;=\; \begin{pmatrix} 4 & -2 \\ -2 & 0 \end{pmatrix}$$

Note that the sum of $\overline{g}$ over any lower set is non-negative. More generally, for $g : X \to R$, we could define the signed measure $\mu_g(X') = \mathbb{E}[g(\mathbf{x}) | \mathbf{x} \in X']$ for any $X' \subseteq X$, and say μ_g is *lower-set positive* if $\mu_g(X_-) \geq 0$ for any lower set $X_- \subset X$ and $\mu_g(X) = 0$. Since $\overline{MR}(\mathbf{x}) \succ MR(\mathbf{x})$, the solution $\overline{g}$ to (11) generally defines a lower-set positive measure $\mu_{\overline{g}}$.

4. Characterizing the Optimal Mechanism

The optimal mechanism for the public goods case ($\eta_i = 1$ for all $i \in N$) is readily obtained using the multivariate majorization results of the previous section.

Proposition 6 *The optimal mechanism for public goods is given by* $\{q^*(\mathbf{x}), t_i^*(\mathbf{x})\}$ *where*

$$q^*(\mathbf{x}) \;=\; C'^{-1}\big(\max(0, \overline{MR}(\mathbf{x}))\big)$$

and

$$t_i^*(\mathbf{x}) \;=\; v_i(\mathbf{x})q^*(\mathbf{x}) - \sum_{s_i < x_i} \overline{\Delta} v_i(s_i, \mathbf{x}_{-i})q^*(s_i, \mathbf{x}_{-i})$$

where $\overline{MR}(\mathbf{x}) = \sum_{i\in N} \overline{MR}_i(\mathbf{x})$ *follows from* (8).

The optimal quality coarsens the partition induced by the ironed marginal revenue $\overline{MR}(\mathbf{x})$ since the profiles for which $\overline{MR}(\mathbf{x}) < 0$ are pooled. Let $\mathcal{Q}$ denote the partition induced by $q^*(\mathbf{x})$ and let $Q(\mathbf{x}) = \{\mathbf{y} \in X \,|\, q^*(\mathbf{y}) = q^*(\mathbf{x})\}$ be the cell of $\mathcal{Q}$ that $\mathbf{x}$ belongs to. A singleton $Q(\mathbf{x})$ corresponds to a unique quality level while profiles are "pooled" or "bunched" when $|Q(\mathbf{x})| > 1$. Generally, the partition $\mathcal{Q}$ will have the structure

$$\mathcal{Q} \;=\; \mathcal{Q}_0 \bigcup \cup\, \mathcal{Q}_P \bigcup \cup\, \mathcal{Q}_S$$

where $\mathcal{Q}_0$ is the set on which $q^*(\mathbf{x})$ vanishes, $\mathcal{Q}_P$ are sets with more than one element on which $q^*(\mathbf{x})$ is constant, and $\mathcal{Q}_S$ are singleton sets. This partition is reminiscent of that in Rochet and Choné (1998) except that in their model there is a single non-empty $\mathcal{Q}_0$, $\mathcal{Q}_P$, and $\mathcal{Q}_S$, which is not necessarily the case here.

Example 2 Suppose costs are quadratic, types are uniform on $X_1 = X_2 = \{0, 1, 2\}$, and valuations are $v_i(x_i, x_{-i}) = 1 + \frac{1}{2}x_i + 9x_{-i}(2 - x_{-i})$ for $i = 1, 2$, then

$$MR \;=\; \begin{pmatrix} 0 & 10 & 2 \\ 10 & 20 & 12 \\ 2 & 12 & 4 \end{pmatrix}$$

The program in (8) yields

$$\overline{MR} \;=\; \begin{pmatrix} 0 & 6 & 6 \\ 6 & 12 & 12 \\ 6 & 12 & 12 \end{pmatrix}$$

so $\mathcal{Q}_0 = \{(0, 0)\}$, $\mathcal{Q}_P = \{(0, 1), (0, 2), (1, 0), (2, 0)\} \cup \{(2, 2), (2, 3), (3, 2), (3, 3)\}$, and $\mathcal{Q}_S = \emptyset$. The seller's profit is $\Pi^* = \frac{1}{2}\mathbb{E}[q^2]$ with $q = \overline{MR}$ so $\Pi^* = 360$. ■

4.1. Discriminatory Access Rights

Discriminatory access rights can take the form of all-or-nothing access rights that can be used to screen out negative marginal revenues. In addition, probabilistic access rights can be used to alleviate incentive constraints and allow the ironed marginal revenues to be closer to the original individual marginal revenues. As we will see, this type of constrained majorization results in higher profits in Example 2.

With access rights the seller's problem is

$$\Pi \;=\; \max_{\substack{(q,\boldsymbol{\eta}):X\to\mathbb{R}_{\geq 0}\times[0,1]^n \\ \eta_i(\mathbf{x})q(\mathbf{x})\text{ non-decreasing in } x_i}} \mathbb{E}\Big[q(\mathbf{x})\sum_{i\in N}\eta_i(\mathbf{x})MR_i(\mathbf{x}) - C(q(\mathbf{x}))\Big]$$

The constraint that $\eta_i(\mathbf{x})q(\mathbf{x})$ is non-decreasing in x_i for all $i\in N$ can be dealt with by adding $\sum_{i\in N,\mathbf{x}\in X}\lambda_i(\mathbf{x})\overline{\Delta}_i(\eta_i(\mathbf{x})q(\mathbf{x}))$ where the $\lambda_i(\mathbf{x})$ are non-negative for all $\mathbf{x}\in X$ with $\lambda_i(\overline{x}_i,\mathbf{x}_{-i})=0$ for all $i\in N$. This term can rewritten to yield the following saddle-point problem

$$\Pi \;=\; \min_{\lambda:X\to\mathbb{R}_+}\;\max_{(q,\boldsymbol{\eta}):X\to\mathbb{R}_{\geq 0}\times[0,1]^n}\mathbb{E}\Big[q(\mathbf{x})\sum_{i\in N}\eta_i(\mathbf{x})(MR_i(\mathbf{x})-\underline{\Delta}_i\lambda_i(\mathbf{x})/f(\mathbf{x}))-C(q(\mathbf{x}))\Big]$$

The $\eta_i(\mathbf{x})$ enter linearly so maximizing over them yields $\eta_i(\mathbf{x})\in\{0,1\}$ unless the shifted marginal revenue they multiply is 0 in which case $\eta_i(\mathbf{x})$ can be fractional. This turns the sum on the righthand side into $\sum_{i\in N}\max(0,MR_i(\mathbf{x})-\underline{\Delta}_i\lambda_i(\mathbf{x})/f(\mathbf{x}))$. Maximizing over $q(\mathbf{x})$ then yields $q(\mathbf{x})=C'^{-1}(\sum_{i\in N}\max(0,\widetilde{MR}_i(\mathbf{x}))$ and

$$\Pi \;=\; \mathbb{E}\big[\phi\big(\sum_{i\in N}\max(0,\widetilde{MR}_i(\mathbf{x}))\big)\big]$$

where $\phi(y)=yC'^{-1}(y)-C(C'^{-1}(y))$ is convex and the $\widetilde{MR}_i(\mathbf{x})$ follow from

$$\widetilde{\boldsymbol{MR}}(\mathbf{x}) \;=\; \operatorname*{argmin}_{\substack{g_i:X\to\mathbb{R}\\ g_i\succ MR_i}}\mathbb{E}\Big[\big(\sum_{i\in N}\max(0,g_i(\mathbf{x}))\big)^2\Big] \tag{12}$$

As in the previous section, we could replace the sum of squares by an arbitrary convex function. However, unlike the previous section, this does not allow us to replace the $\max(0,g_i(\mathbf{x}))$ with $g_i(\mathbf{x})$ since the $\max(0,\cdot)$ operation is applied to each term separately rather than to the sum $\sum_{i\in N}g_i(\mathbf{x})$. As a result, there is a trivial way in which the solution to (12) is no longer unique, unlike the solution to (8). If $\widetilde{MR}_i(\mathbf{x})<0$ for some $\mathbf{x}\in X$, $i\in N$ then we could replace it with any other negative number. This multiplicity poses no problem since these cases are screened out and the optimal mechanism is unique, see Lemma 2 in Appendix A.

A main difference with the previous section is that we cannot use level sets of $\widetilde{MR}(\mathbf{x})=\sum_{i\in N}\widetilde{MR}_i(\mathbf{x})$ to define the partition $\mathscr{P}$. To illustrate, consider Example 2 for which (12) yields

$$\widetilde{MR} \;=\; \begin{pmatrix} 0 & 9 & 3\\ 9 & 14 & 14\\ 3 & 14 & 6\end{pmatrix}$$

Based on level sets, $\mathscr{P}$ would consist of five cells: two singletons, two sets of size two, and a set of size three. However, this partition violates the majorization requirement that $\widetilde{MR}$ has the same expected value as MR on each cell.

The correct partition can be obtained from the program in (12), which can be executed by parameterizing $g(\mathbf{x}) = MR(\mathbf{x}) - \sum_i \underline{\Delta}_i \lambda_i(\mathbf{x})$, with $\lambda_i(\mathbf{x}) \geq 0$ and $\lambda_i(\overline{x}_i, \mathbf{x}_{-i}) = 0$ for $i = 1, 2$ and $\mathbf{x} \in X$, and then minimizing over the λ_is. This yields

$$\lambda_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

For buyer i we fix $\mathbf{x}_{-i}$ and find the smallest x_i (if any) for which $\lambda_i(x_i, x_{-i}) > 0$ as well as the smallest $x_i' > x_i$ for which $\lambda_i(x_i, x_{-i}) = 0$ to form the ortho-convex sets, $\{(x, x_{-i}) | x_i \leq x \leq x_i'\}$, which are highlighted above. We do this for all buyers and then "overlay" the individual ortho-convex sets to find the partition. Profiles that do not belong to any of the highlighted sets form singletons on which $\widetilde{MR}(\mathbf{x}) = MR(\mathbf{x})$. For the above example, the correct partition is: $X = P_1 \sqcup P_2 \sqcup P_3 \sqcup P_4$ where $P_1 = \{(0,0)\}$, $P_2 = \{(0,1),(0,2)\}$, $P_3 = \{(1,0),(2,0)\}$, and $P_4 = \{(2,2),(2,3),(3,2),(3,3)\}$. Now $\widetilde{MR}$ has the same expected values as MR on each of the cells. Note that $\widetilde{MR}$ is not necessarily constant on a cell, unlike $\overline{MR}$. Incentive compatibility is maintained by choosing appropriate access rights, i.e.

$$\eta_1 = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{3} & 1 & \frac{3}{7} \\ 1 & 1 & 1 \end{pmatrix}$$

and η_2 is the transpose of η_1. The seller's profit is $\Pi^* = \frac{1}{2}\mathbb{E}[q^2]$ with $q = \widetilde{MR}$ so $\Pi^* = 402$.

Proposition 7 *Let $\mathscr{P}$ denote the ortho-convex partition of X generated by (12). For $\mathbf{x} \in X$, let $P(\mathbf{x})$ be the cell of $\mathscr{P}$ containing $\mathbf{x}$ and let $P_i(\mathbf{x}) = \{\mathbf{y} \in P(\mathbf{x}) | \mathbf{y}_{-i} = \mathbf{x}_{-i}\}$. The optimal mechanism for excludable goods is given by $(q^*, \boldsymbol{\eta}^*, \mathbf{t}^*)$ where*

$$q^*(\mathbf{x}) = C'^{-1}\Big(\sum_{i \in N} \max(0, \widetilde{MR}_i(\mathbf{x}))\Big)$$

and, for $i \in N$,

$$\eta_i^*(\mathbf{x}) = \begin{cases} 0 & \text{if } \widetilde{MR}_i(\mathbf{x}) < 0 \\ \eta_i^0(\mathbf{x}) & \text{if } \widetilde{MR}_i(\mathbf{x}) = 0 \\ 1 & \text{if } \widetilde{MR}_i(\mathbf{x}) > 0 \end{cases}$$

where $\eta_i^0(\mathbf{x})$ *is such that* $q^*(\mathbf{x})\eta_i^0(\mathbf{x})$ *is constant on* $P_i(\mathbf{x})$ *and non-decreasing in* x_i,[15] *and*

$$t_i^*(\mathbf{x}) \;=\; v_i(\mathbf{x})q^*(\mathbf{x})\eta_i^*(\mathbf{x}) - \sum_{s_i < x_i} \overline{\Delta}_i v_i(s_i, \mathbf{x}_{-i})q^*(s_i, \mathbf{x}_{-i})\eta_i^*(s_i, \mathbf{x}_{-i})$$

Probabilistic access rights do not only arise as knife-edge cases. They allow the ironed marginal values and, hence, the optimal quality to more closely track the original marginal values when the latter are non-monotonic. As a result, the seller is better off while buyers may benefit or be worse off – see Section 5.4 for an example of the former and Example 1 for the latter. In both cases, however, access rights are welfare improving.

4.2. Welfare Maximizing Mechanisms

The welfare-maximizing mechanism can be derived similarly to Proposition 7. Simply replace the $\widetilde{MR}_i(\mathbf{x})$ in that proposition with $\widetilde{v}_i(\mathbf{x})$, where the $\widetilde{v}_i$ for $i \in N$ follow from

$$\widetilde{\boldsymbol{v}}(\mathbf{x}) \;=\; \underset{\substack{g_i : X \to \mathbb{R} \\ g_i \succ v_i}}{\operatorname{argmin}} \, \mathbb{E}\Big[\Big(\sum_{i \in N} \max(0, g_i(\mathbf{x}))\Big)^2\Big]$$

Even though $v_i(\mathbf{x})$ is non-decreasing in x_i for all $i \in N$, their sum may not be non-decreasing in each coordinate, hence the need to majorize. For Example 2, this yields

$$\widetilde{v} \;=\; \begin{pmatrix} 2 & 10 & \frac{9}{2} \\ 10 & \frac{43}{3} & \frac{43}{3} \\ \frac{9}{2} & \frac{43}{3} & 7 \end{pmatrix}$$

where $\widetilde{v}(\mathbf{x}) \;=\; \widetilde{v}_1(\mathbf{x}) + \widetilde{v}_2(\mathbf{x})$. Incentive compatibility is maintained using probabilistic access rights, i.e.

$$\eta_1 \;=\; \begin{pmatrix} 1 & 1 & 1 \\ \frac{9}{20} & 1 & \frac{21}{43} \\ 1 & 1 & 1 \end{pmatrix}$$

and η_2 is the transpose of η_1. This shows that probabilistic access rights can be socially optimal. The same is true for binary or "all-or-nothing" access rights used to exclude negative values, see Section 5.4.

5. Continuous Types

In this section, we extend our approach to continuous types, which allows for a more direct comparison to Mussa and Rosen's (1978) one-dimensional "ironing" technique. We also discuss

[15] In detail, (i) if there exists $\mathbf{y} \in P_i(\mathbf{x})$ such that $\widetilde{MR}_i(\mathbf{y}) \neq 0$ then $q^*(\mathbf{x})\eta_i^0(\mathbf{x}) = q^*(\mathbf{y})\eta_i^*(\mathbf{y})$, (ii) otherwise $q^*(\mathbf{x})\eta_i^0(\mathbf{x}) = q^*(\mathbf{y})\eta_i^0(\mathbf{y})$ for all $\mathbf{y} \in P_i(\mathbf{x})$ and $q^*(\underline{\mathbf{y}})\eta_i^*(\underline{\mathbf{y}}) \leq q^*(\mathbf{x})\eta_i^0(\mathbf{x}) \leq q^*(\overline{\mathbf{y}})\eta_i^*(\overline{\mathbf{y}})$ for any $\underline{y}_i \leq x_i \leq \overline{y}_i$ with $\widetilde{MR}_i(\underline{\mathbf{y}}) \neq 0$ and $\widetilde{MR}_i(\overline{\mathbf{y}}) \neq 0$.

connections with Rochet and Choné's (1998) "sweeping" method that applies to the multi-dimensional case. As we will show, continuous types can be dealt with in much the same manner as the discrete-type case studied above.

5.1. Majorization and the Gauss Divergence Theorem

Consider the distance minimization characterization in (10). To ensure the result will be non-decreasing in each coordinate, we add $\sum_{i\in N}\int_X \lambda_i(\mathbf{x})\partial_{x_i}g(\mathbf{x})$ to the (transformed) objective $\frac{1}{2}||MR(\mathbf{x})-g(\mathbf{x})||^2$ where $\lambda_i(\mathbf{x})\geq 0$ and $\lambda_i(\underline{x}_i,\mathbf{x}_{-i})=\lambda_i(\overline{x}_i,\mathbf{x}_{-i})=0$ for $i\in N$ and $\mathbf{x}\in X$. Let $\text{div}(\boldsymbol{\lambda}(\mathbf{x}))$ denote the divergence of $\boldsymbol{\lambda}(\mathbf{x})$, i.e. $\text{div}(\boldsymbol{\lambda}(\mathbf{x}))=\sum_{i\in N}\partial_{x_i}\lambda_i(\mathbf{x})$. Similar to the discrete case we define

$$MR(\mathbf{x},\boldsymbol{\lambda}(\mathbf{x})) \;=\; MR(\mathbf{x})-\text{div}(\boldsymbol{\lambda}(\mathbf{x}))/f(\mathbf{x})$$

then the first-order conditions imply $\partial_{x_i}MR(\mathbf{x},\boldsymbol{\lambda}(\mathbf{x}))\geq 0$ and $\lambda_i(\mathbf{x})\partial_{x_i}MR(\mathbf{x},\boldsymbol{\lambda}(\mathbf{x}))=0$ for $i\in N$, $\mathbf{x}\in X$.

Consider any lower set $X_-\subseteq X$ with boundary ∂X_- that is the union of the upper boundary $\overline{\partial}X_-$ and the lower boundary $\underline{\partial}X_-=\partial X_-\setminus\overline{\partial}X_-$. The geometry of lower sets is such that the normal $\boldsymbol{n}(\mathbf{x})$ to any point $\mathbf{x}$ on the upper boundary is positive, i.e. $\boldsymbol{n}(\mathbf{x})\geq\mathbf{0}$, whence $\boldsymbol{\lambda}(\mathbf{x})\cdot\boldsymbol{n}(\mathbf{x})\geq 0$. On the lower boundary we must have $x_i=\underline{x}_i$ for some $i\in N$, in which case the normal $\boldsymbol{n}(\mathbf{x})$ is minus the ith unit vector and $\boldsymbol{\lambda}(\mathbf{x})\cdot\boldsymbol{n}(\mathbf{x})=0$ since $\lambda_i(\underline{x}_i,\mathbf{x}_{-i})=0$. By the Gauss divergence theorem we have

$$\begin{aligned}
\mathbb{E}\big[MR(\mathbf{x},\boldsymbol{\lambda}(\mathbf{x}))\,|\,\mathbf{x}\in X_-\big] &= \mathbb{E}\big[MR(\mathbf{x})-\text{div}(\boldsymbol{\lambda}(\mathbf{x}))/f(\mathbf{x})\,|\,\mathbf{x}\in X_-\big]\\
&= \mathbb{E}\big[MR(\mathbf{x})\,|\,\mathbf{x}\in X_-\big]-\frac{1}{F(X_-)}\int_{\overline{\partial}X_-}\boldsymbol{\lambda}(\mathbf{x})\cdot\boldsymbol{n}(\mathbf{x})\\
&\leq \mathbb{E}\big[MR(\mathbf{x})\,|\,\mathbf{x}\in X_-\big]
\end{aligned}$$

with equality if $\boldsymbol{\lambda}(\mathbf{x})\cdot\boldsymbol{n}(\mathbf{x})=0$ on $\overline{\partial}X_-$. In particular, since $\lambda_i(\bar{x}_i,\mathbf{x}_{-i})=0$ for all $\mathbf{x}_{-i}\in X_{-i}$ and $i\in N$, we have

$$\mathbb{E}\big[MR(\mathbf{x},\boldsymbol{\lambda}(\mathbf{x}))\big]=\mathbb{E}\big[MR(\mathbf{x})\big]$$

To summarize, the majorization Definition 2 also applies to functions defined over continuous type spaces. Moreover, optimality dictates that $MR(\mathbf{x},\boldsymbol{\lambda}(\mathbf{x}))$ is non-decreasing in each coordinate and complementary slackness dictates that it is flat in the ith direction whenever $\lambda_i(\mathbf{x})>0$.

5.2. Ironing and Sweeping

The univariate case provides an opportunity to illustrate how majorization relates to Mussa and Rosen's (1978) "ironing" and Rochet and Choné's (1998) "sweeping." Suppose, for example, that types are uniform and marginal revenue are given by

$$MR(x) = \begin{cases} 2x & \text{if } x \leq \frac{1}{4} \\ \frac{1}{2} + 16(x - \frac{1}{4})(x - \frac{1}{2})(x - \frac{3}{4}) & \text{if } \frac{1}{4} \leq x \leq \frac{3}{4} \\ 2x - 1 & \text{if } x \geq \frac{3}{4} \end{cases}$$

Then Mussa and Rosen's (1978) "ironing" would flatten $MR(x)$ in the middle region to $MR(x) = \frac{1}{2}$. Proposition 4 shows that the ironed marginal revenues can be obtained via the stochastic operator in (9):

$$S(x, y) = \begin{cases} 2 & \text{if } \frac{1}{4} \leq x, y \leq \frac{3}{4} \\ \delta(x - y) & \text{otherwise} \end{cases}$$

with $\delta(x)$ the Dirac δ-function, i.e. $\overline{MR}(x) = \int_0^1 S(x, y)MR(y)dy$ or $\overline{MR} = S \cdot MR$ for short.

Rochet and Choné (1998) provide an alternative interpretation in terms of "sweeping." Seen as a measure, $\overline{MR}(x)$ can be obtained from $MR(x)$ by sweeping mass out of $[\frac{1}{4}, \frac{3}{4}]$ to mass-points at the endpoints of this interval – a "mean-preserving spread."[16] More generally, Rochet and Choné (1998) demonstrate that their sweeping method reproduces Mussa and Rosen's (1978) ironing in the univariate case (see their Appendix 4).

We next show how Rochet and Choné's (1998) "sweeping" operator is related to the "ironing" operator $S(x, y)$. One difference is that $S(x, y)$ maps the first-order condition $q(\mathbf{x}) - MR(\mathbf{x})$ evaluated at the candidate solution $q(\mathbf{x}) = \overline{MR}(\mathbf{x})$ to zero,[17] while Rochet and Choné's (1998) sweeping operator maps the derivative of the first-order condition evaluated at the candidate solution to zero. Hence, the sweeping operator, $T(x, y)$, follows from $S(x, y)$ by integrating over x and differentiating with respect to y:

$$T(x, y) = \begin{cases} (\frac{3}{2} - 2x)\delta(y - \frac{1}{4}) + (2x - \frac{1}{2})\delta(y - \frac{3}{4}) & \text{if } \frac{1}{4} \leq x, y \leq \frac{3}{4} \\ \delta(x - y) & \text{otherwise} \end{cases}$$

It is readily verified that $\int_0^1 T(x, y)dy = 1$, i.e. T is stochastic, and that $\int_0^1 yT(x, y)dy = x$, i.e. T preserves the center of mass and implements a mean-preserving spread. Finally, $(\overline{MR}'(x) - MR'(x)) \cdot T = 0$ since $\overline{MR}$ and MR differ only on the interval $(\frac{1}{4}, \frac{3}{4})$ where T vanishes.[18]

Rochet and Choné's (1998) sweeping method can be applied to multi-dimensional environments and, as such, provides an alternative to our majorization approach. However, an important difference is that Rochet and Choné's (1998) approach is non-constructive. A candidate solution

[16]See, in particular, footnote 23 in Rochet and Choné (1998). A recent paper by Kleiner et al. (2020) that characterizes extreme points for the (uniform) univariate case, also makes this observation.

[17]Note that $S(x, y)$ is idempotent so $\overline{MR} = S \cdot MR$ is equivalent to $S \cdot (\overline{MR} - MR) = 0$.

[18]As a further illustration of the sweeping operator, consider the uniform case $F(x) = x$. Then $F' \cdot T = \int_0^1 T(x, y)dx$ yields a continuous density, $f(y) = 1$ for $y \leq \frac{1}{4}$ and $y \geq \frac{3}{4}$, and two mass points, $\frac{1}{4}\delta(y - \frac{1}{4})$ and $\frac{1}{4}\delta(y - \frac{3}{4})$.

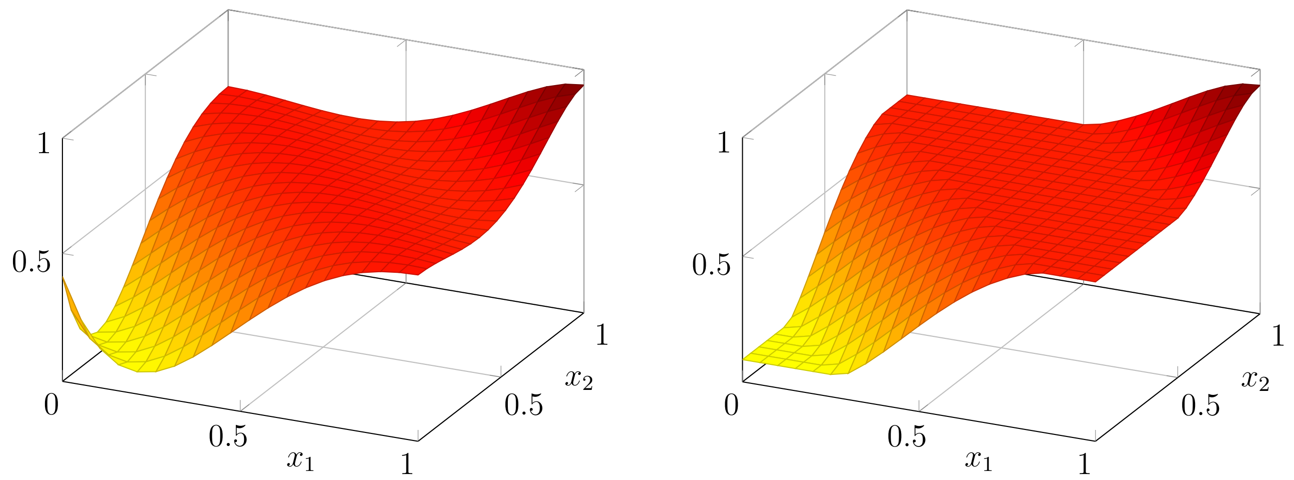

Figure 2: The left panel shows the non-monotonic marginal revenues $MR(\mathbf{x})$ of Example 3 and the right panel shows its ironed version $\overline{MR}(\mathbf{x})$.

is proposed and then a sweeping operator that maps the derivative of the first-order condition evaluated at this candidate solution to zero has to be found. In contrast, the multivariate majorization technique we propose constructs the optimal solution from the exogenously specified marginal revenues.

5.3. Characterizing the Ironed Marginal Revenues

The ironed marginal revenues for the continuous case can be obtained from (10) and the characterization in Proposition 4 applies: $\overline{MR}(\mathbf{x})$ is the non-decreasing function closest to $MR(\mathbf{x})$, its level sets form an ortho-convex partition $\mathscr{P}$ of X,[19] $\overline{MR}(\mathbf{x}) = \mathbb{E}[MR(\mathbf{y}) \,|\, \mathbf{y} \in P(\mathbf{x})]$ where $P(\mathbf{x}) = \{\mathbf{y} \,|\, \overline{MR}(\mathbf{y}) = \overline{MR}(\mathbf{x})\}$ is the partition cell that contains $\mathbf{x}$, and the stochastic operator in (9) maps $MR(\mathbf{x})$ to $\overline{MR}(\mathbf{x})$.

Example 3 Suppose types are uniformly distributed on $X = [0,1]^2$ and that value functions are $v_1(x_1, x_2) = v_2(x_1, x_2) = v(x_1 + x_2)$. It is readily verified that the marginal revenues will be functions of $x = x_1 + x_2$. The left panel of Figure 2 shows $MR(x) = MR_1(x) + MR_2(x)$ when $v(x) = \frac{1}{4} + x^2 - \frac{5}{3}x^3 + x^4 - \frac{1}{5}x^5$ and the right panel shows the ironed version $\overline{MR}(x)$.

There are two flat regions, i.e. $x_1 + x_2 \leq \alpha$ and $\beta \leq x_1 + x_2 \leq \gamma$, and two regions where $\overline{MR}(x) = MR(x)$, i.e. $\alpha \leq x_1 + x_2 \leq \beta$ and $x_1 + x_2 \geq \gamma$.[20] If we interpret $\overline{MR}(\mathbf{x})$ as a measure, it can be obtained from $MR(\mathbf{x})$ by sweeping all mass in the lower triangle $x_1 + x_2 \leq \alpha$ to a mass-point at 0. Likewise, for the region $\beta \leq x_1 + x_2 \leq \gamma$, the excess mass $MR(\mathbf{x}) - \overline{MR}(\mathbf{x}) \geq 0$

[19] If $\overline{MR}(x_i, \mathbf{x}_{-i}) = \overline{MR}(x_i', \mathbf{x}_{-i})$ for some $x_i < x_i'$ then $\overline{MR}(x, \mathbf{x}_{-i})$ is constant for all $x_i \leq x \leq x_i'$ since otherwise $\overline{MR}(\mathbf{x})$ is not non-decreasing.

[20] Where $\alpha \approx 0.29$, $\beta \approx 0.87$, and $\gamma \approx 1.51$.

of the lower part of this region is swept to cover the deficit $MR(\mathbf{x}) - \overline{MR}(\mathbf{x}) \leq 0$ of the upper part of this region. ■

5.4. The Optimal Mechanism

With the ironed marginal revenues $\overline{MR}(\mathbf{x})$ determined by (10), the optimal mechanism for public goods follows from Proposition 6. The only change is that the optimal payments become

$$t^*(\mathbf{x}) \;=\; v_i(\mathbf{x})q^*(\mathbf{x}) - \int_{\underline{x}_i}^{x_i} v_i'(s_i, \mathbf{x}_{-i})q^*(s_i, \mathbf{x}_{-i})ds_i$$

where $q^*(\mathbf{x}) = C'^{-1}(\max(0, \overline{MR}(\mathbf{x})))$ as before.

Access rights can be incorporated as well for continuous type spaces. To illustrate, suppose there are two buyers with uniformly distributed types on $X = [0,1]^2$ and value functions $v_i(x_i, x_{-i}) = x_i - 2x_{-i}$ for $i = 1, 2$. Then $MR(x_1, x_2) = -2$ for all $(x_1, x_2) \in X$ and, without access rights, the optimal quality is zero everywhere (as is the seller's revenue). In contrast, with access rights, $q^*(x_1, x_2) = \max(0, |x_1 - x_2| - \frac{1}{2})$ and $\eta_1^*(x_1, x_2) = \delta_{x_1 > x_2 + 1/2}$ and $\eta_2^*(x_1, x_2) = \delta_{x_2 > x_1 + 1/2}$. Note that the introduction of access rights benefits the seller as well as the buyers. Also, the welfare-maximizing mechanism has $q^*(x_1, x_2) = \max(0, x_1 - 2x_2, x_2 - 2x_1)$ and $\eta_1^*(x_1, x_2) = \delta_{x_1 > 2x_2}$ and $\eta_2^*(x_1, x_2) = \delta_{x_2 > 2x_1}$ showing that all-or-nothing access rights can be welfare improving.

Finally, the profitability of probabilistic access rights can be illustrated by creating a value non-monotonicity in the x_{-i} coordinate. Suppose we add 2 to $v_i(x_i, x_{-i})$ when $\frac{1}{3} \leq x_{-i} \leq \frac{2}{3}$ and add 20 outside this region. Then the sum of marginal revenues, $MR(x_1, x_2)$, is non-monotonic. Without access rights, the sum of the ironed marginal revenues is

$$\overline{MR}(x_1, x_2) \;=\; 20 + 9\,\delta_{x_1 > \frac{2}{3}} + 9\,\delta_{x_2 > \frac{2}{3}}$$

while, with access rights,

$$\widetilde{MR}(x_1, x_2) \;=\; \overline{MR}(x_1, x_2) + \frac{46}{9}\,(1 - 4\,\delta_{\min(x_1, x_2) > \frac{1}{3}})\,\delta_{\max(x_1, x_2) < \frac{2}{3}}$$

In other words, $\widetilde{MR}(x_1, x_2)$ drops in the region where both buyers' types are intermediate. Incentive compatibility is maintained, however, by setting appropriate access rights, i.e., for $i = 1, 2$, $\eta_i^*(x_i, x_{-i}) = \frac{15}{61}$ when $x_i \leq \frac{1}{3}$ and $\frac{1}{3} \leq x_{-i} \leq \frac{2}{3}$, and $\eta_i^*(x_i, x_{-i}) = 1$ otherwise.

6. Conclusion

We characterize the optimal mechanism for a monopolist to sell an excludable, non-rival good when buyers' values are interdependent. We show how the interdependency of buyers' values creates multidimensional incentive-compatibility constraints on the monopolist's problem. These constraints cannot be addressed using the standard (univariate) techniques of the literature. Instead, in the spirit of Mussa and Rosen (1978) and Myerson (1981), we define a concept of multivariate majorization that can be used to jointly "iron" buyers' preferences. The resulting preferences are sufficiently well behaved that the incentive constraints no longer strictly bind allowing the monopolist to effectively ignore them. Notably, our approach is constructive: we show how to find the needed preferences through simple quadratic optimization problems.

The problem we study is perhaps most easily analyzed in the case of pure public goods, i.e. when all buyers are guaranteed access regardless of their type. Indeed, this is the assumption under which we develop our concept of multivariate majorization. We show, however, that our approach extends directly to the case of excludable goods. In so doing, we demonstrate that exclusions and random access rights are important ways for the monopolist to raise profits. The former allows the monopolist to exclude buyers with negative marginal revenue and the latter allows the monopolist to fine tune incentives, both without manipulating the overall choice of quality for the market.

Though we focus mainly on buyers with discrete types, we demonstrate that our multivariate majorization techniques can also be applied to the continuous case. This allows us to compare our multivariate majorization approach to Rochet and Choné's (1998) "sweeping" method and demonstrate with an example how the stochastic operators that implement "ironing" and "sweeping" are related. Importantly, Rochet and Choné's (1998) approach is not constructive, i.e. a candidate solution has to be guessed and then a sweeping operator, which maps the derivative of the first-order condition evaluated at this solution to zero, has to be found. In contrast, our multivariate majorization approach constructs the optimal solution from the exogenously specified marginal revenues. An interesting avenue for future research is to examine whether the approach can be adapted to handle multidimensional types for individual buyers.

A. Proofs

Proof of Proposition 2. For notational convenience, we prove the proposition for the case of a uniform distribution over types for all players. We say that $\mathbf{x}$ is lexicographically lower than $\mathbf{y}$, written $\mathbf{y} \leq_L \mathbf{x}$, if there is a $k \in N$ such that $y_k > x_k$, then there exists $l \in N$, $l < k$ such that $y_l < x_l$; $\mathbf{x}$ is strictly lexicographically lower than $\mathbf{y}$, written $\mathbf{y} <_L \mathbf{x}$ if $\mathbf{y} \leq_L \mathbf{x}$ and $\mathbf{x} \neq \mathbf{y}$. In words, we first sort type profiles by player, then by type.

Suppose $g \succ h$. Choose $i \in N$ and $\mathbf{x}_{-i}$ such that there is $\hat{x}_i$ and $\tilde{x}_i$ with $\hat{x}_i < \tilde{x}_i$, $h(\hat{x}_i, \mathbf{x}_{-i}) > g(\hat{x}_i, \mathbf{x}_{-i})$, $h(\tilde{x}_i, \mathbf{x}_{-i}) < g(\tilde{x}_i, \mathbf{x}_{-i})$ and where $(\tilde{x}_i, \mathbf{x}_{-i})$ is the lexicographically lowest such profile and $(\hat{x}_i, \mathbf{x}_{-i})$ is the lexicographically highest such type (lexicographically below $(\tilde{x}_i, \mathbf{x}_{-i})$)). Let

$$\alpha = \frac{\delta}{g(\tilde{x}_i, \mathbf{x}_{-i}) - g(\hat{x}_i, \mathbf{x}_{-i})}$$

where $\delta = \min\big(g(\tilde{x}_i, \mathbf{x}_{-i}) - h(\tilde{x}_i, \mathbf{x}_{-i}), h(\hat{x}_i, \mathbf{x}_{-i}) - g(\hat{x}_i, \mathbf{x}_{-i})\big)$ and define the orthogonal T-transform

$$T(\mathbf{z}, \mathbf{y}) = \begin{cases} \alpha & \text{if } \mathbf{z}_{-i} = \mathbf{y}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = y_i = \hat{x}_i \text{ or } z_i = y_i = \tilde{x}_i \\ 1-\alpha & \text{if } \mathbf{z}_{-i} = \mathbf{y}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \hat{x}_i, y_i = \tilde{z}_i \text{ or } z_i = \tilde{x}_i, y_i = \hat{x}_i \\ 1 & \text{if } \mathbf{z} = \mathbf{y} \notin \{\hat{\mathbf{x}}, \tilde{\mathbf{x}}\} \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\begin{aligned} (Tg)(\mathbf{z}) &= \begin{cases} \alpha g(\hat{x}_i, \mathbf{x}_{-i}) + (1-\alpha) g(\tilde{x}_i, \mathbf{x}_{-i}) & \text{if } \mathbf{z}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \hat{x}_i \\ \alpha g(\tilde{x}_i, \mathbf{x}_{-i}) + (1-\alpha) g(\hat{x}_i, \mathbf{x}_{-i}) & \text{if } \mathbf{z}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \tilde{x}_i \\ g(\mathbf{z}) & \text{otherwise} \end{cases} \\ &= \begin{cases} g(\hat{x}_i, \mathbf{x}_{-i}) + \delta & \text{if } \mathbf{z}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \hat{x}_i \\ g(\tilde{x}_i, \mathbf{x}_{-i}) - \delta & \text{if } \mathbf{z}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \tilde{x}_i \\ g(\mathbf{z}) & \text{otherwise} \end{cases} \end{aligned}$$

It is clear that $Tg \prec g$. It is also true that $h \prec Tg$. To see this, suppose instead that there is a lower set $X_- \subset X$ such that $\mathbb{E}\big((Tg)(\mathbf{z}) | \mathbf{z} \in X_-\big) > \mathbb{E}\big(h(\mathbf{z}) | \mathbf{z} \in X_-\big)$. Then $(\hat{x}_i, \mathbf{x}_{-i}) \in X_-$ but $(\hat{x}_i, \mathbf{x}_{-i}) \notin X_-$. By our choice of i, x_{-i}, $\hat{x}_i$ and $\tilde{x}_i$, we may assume

$$\{\mathbf{x}' | \mathbf{x}' <_L (\tilde{x}_-, \mathbf{x}_{-i})\} \subseteq X_-.$$

But $(Tg)(\tilde{x}_i, \mathbf{x}_{-i}) \geq h(\tilde{x}_i, \mathbf{x}_{-i})$ so that

$$\mathbb{E}\big(Tg(\mathbf{z}) | \mathbf{z} \in X_- \cup \{(\tilde{x}_i, \mathbf{x}_{-i})\}\big) > \mathbb{E}\big(h(\mathbf{z}) | \mathbf{z} \in X_- \cup \{(\tilde{x}_i, \mathbf{x}_{-i})\}\big)$$

and $\mathbb{E}\big(Tg(\mathbf{z}) | \mathbf{z} \in X_- \cup \{(\tilde{x}_i, \mathbf{x}_{-i})\}\big) = \mathbb{E}\big(g(\mathbf{z}) | \mathbf{z} \in X_- \cup \{(\tilde{x}_i, \mathbf{x}_{-i})\}\big)$. Since $X_- \cup \{(\tilde{x}_i, \mathbf{x}_{-i})\}$ is a lower set, this contradicts our assumption that $g \succ h$ and we conclude that $Tg \succ h$.

Finally, for any two functions ϕ and ψ on X, let $d(\phi, \psi)$ denote the number of type profiles $\mathbf{x} \in X$ such that $\phi(\mathbf{x}) \neq \psi(\mathbf{x})$. Then $d(g, Tg) = d(g, h) - 1$. Therefore, repeating this step at most $|X| - 1$ times will result in h.

Now suppose that $h = Sg$ where S is the product of a sequence of orthogonal T-transforms $T^1, \ldots, T^K$ for some finite integer K. Suppose T^1 mixes between type profiles $(\hat{x}_i, \mathbf{x}_{-i})$ and $(\tilde{x}_i, \mathbf{x}_{-i})$, with probability $(\alpha, 1-\alpha)$ for $\alpha \in [0,1]$ and $\hat{x}_i < \tilde{x}_i$. Define

$$g^*(\mathbf{z}) \equiv (T^1 g)(\mathbf{z}) = \begin{cases} \alpha g(\hat{x}_i, \mathbf{x}_{-i}) + (1-\alpha) g(\tilde{x}_i, \mathbf{x}_{-i}) & \text{if } \mathbf{z}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \hat{x}_i \\ \alpha g(\tilde{x}_i, \mathbf{x}_{-i}) + (1-\alpha) g(\hat{x}_i, \mathbf{x}_{-i}) & \text{if } \mathbf{z}_{-i} = \mathbf{x}_{-i} \text{ and } z_i = \tilde{x}_i \\ g(\mathbf{z}) & \text{otherwise} \end{cases}.$$

Then

$$\begin{aligned} \mathbb{E}(g^*(\mathbf{x}) | x \in X_-) &= \begin{cases} \mathbb{E}(g(\mathbf{x}) | x \in X_-) \\ \quad +(1-\alpha)\big(g(\tilde{x}_i, \mathbf{x}_{-i}) - g(\hat{x}_i, \mathbf{x}_{-i})\big) f(\hat{x}_i, \mathbf{x}_{-i}) & \text{if } (\hat{x}_i, \mathbf{x}_{-i}) \in X_-, (\tilde{x}_i, \mathbf{x}_{-i}) \notin X_- \\ \mathbb{E}(g(\mathbf{x}) | x \in X_-) & \text{otherwise} \end{cases} \\ &\geq \quad \mathbb{E}(g(\mathbf{x}) | x \in X_-) \end{aligned}$$

So that $g^* \prec g$. Iterating in this way, we can conclude that $h \prec g$. ■

Proof of Lemma 1. To implement $g \succ MR$ in (8) we parameterize $g(\mathbf{x}) = MR(\mathbf{x}) - \sum_{i \in N} \underline{\Delta}_i \lambda_i^*(\mathbf{x}) / f(\mathbf{x})$ where the $\lambda_i^*(\mathbf{x})$ follow from a simple quadratic optimization program:

$$\boldsymbol{\lambda}^*(\mathbf{x}) \;=\; \underset{\boldsymbol{\lambda}: X \to \mathbb{R}^n_{\geq 0}}{\operatorname{argmin}} \; \mathbb{E}\big[(MR(x) - \sum_{i \in N} \underline{\Delta}_i \lambda_i(\mathbf{x}) / f(\mathbf{x}))^2\big] \tag{13}$$

Its solutions must satisfy, for $i \in N$,

$$\lambda_i^*(\mathbf{x}) \overline{\Delta}_i g(\mathbf{x}) \;=\; 0 \tag{14}$$

$$\overline{\Delta}_i g(\mathbf{x}) \;\geq\; 0 \tag{15}$$

which imply, for $i \in N$, $\lambda_i^*(\mathbf{x}) \overline{\Delta}_i \phi'(g(\mathbf{x})) = 0$ and $\overline{\Delta}_i \phi'(g(\mathbf{x})) \geq 0$ for any convex ϕ. ■

Proof of Proposition 3. The constraint $g \succ MR$ defines a convex polyhedron. Hence, the quadratic program has a unique minimizer that is non-decreasing by construction. Suppose $\overline{\Delta}_i \overline{MR}(\mathbf{x}) < 0$ for some $x \in X$, $i \in N$ then raising $\lambda_i(\mathbf{x})$ slightly lowers the objective in (13), a contradiction. Suppose, in contradiction, there exists a non-decreasing $g(\mathbf{x})$ such that $\overline{MR} \succ g \succ MR$. By Proposition 2 there exists a doubly stochastic operator $S(\mathbf{x}, \mathbf{y})$ such that $g(\mathbf{x}) = \sum_{\mathbf{y} \in X} S(\mathbf{x}, \mathbf{y}) \overline{MR}(\mathbf{y})$. But, convexity of the objective in (8) then implies that g yields a lower value than $\overline{MR}$, which contradicts $\overline{MR}$ being the unique minimizer of (8). ■

Proof of Proposition 4. Suppose, in contradiction, that level sets of $\overline{MR}(\mathbf{x})$ do not form an ortho-convex partition, i.e. there exists $i \in N$, $x_{-i} \in X_{-i}$, and $x_i, x, x_i' \in X_i$ with $x_i < x < x_i'$ such that (x_i, x_{-i}) and (x_i', x_{-i}) belong to the same level set but not (x, x_{-i}). This would violate $\overline{MR}(\mathbf{x})$ being non-decreasing in each coordinate. Next, suppose, again in contradiction, that

$\overline{MR}(\mathbf{x}) \neq \mathbb{E}[MR(\mathbf{y})|\mathbf{y} \in P(\mathbf{x})]$ for some level set $P(\mathbf{x})$ with $\mathbf{x}$ such that there is no $\mathbf{z} \leq \mathbf{x}$ for which $\overline{MR}(\mathbf{z}) \neq \mathbb{E}[MR(\mathbf{y})|\mathbf{y} \in P(\mathbf{z})]$. Since $\overline{MR} \succ MR$ we must have $\overline{MR}(\mathbf{x}) < \mathbb{E}[MR(\mathbf{y})|\mathbf{y} \in P(\mathbf{x})]$ and, since $\mathbb{E}[\overline{MR}(\mathbf{x})] = \mathbb{E}[MR(\mathbf{x})]$, there must exist $\mathbf{x}'$ such that $\overline{MR}(\mathbf{x}') > \mathbb{E}[MR(\mathbf{y})|\mathbf{y} \in P(\mathbf{x}')]$ with $\mathbf{x}'$ such that there is no $\mathbf{z} \leq \mathbf{x}'$ for which $\overline{MR}(\mathbf{z}) > \mathbb{E}[MR(\mathbf{y})|\mathbf{y} \in P(\mathbf{z})]$. The lower set of $\mathbf{x}'$ must include $\mathbf{x}$, i.e. $\mathbf{x}' > \mathbf{x}$, since otherwise $\overline{MR} \succ MR$ is violated. Lowering $\overline{MR}$ on $P(\mathbf{x}')$ by $\varepsilon/|P(\mathbf{x}')|$ while raising $\overline{MR}$ on $P(\mathbf{x})$ by $\varepsilon/|P(\mathbf{x})|$ for some small $\varepsilon > 0$ lowers the objective in (8), contradicting the fact that $\overline{MR}$ minimizes (8).

For the final statement of the proposition, note that $\mathbf{y} \in P(\mathbf{x})$ iff $\mathbf{x} \in P(\mathbf{y})$ so $\delta_{\mathbf{y} \in P(\mathbf{x})} = \delta_{\mathbf{x} \in P(\mathbf{y})}$. Moreover, if $\mathbf{z} \in P(\mathbf{x})$ and $\mathbf{y} \in P(\mathbf{x})$ then $\mathbf{z} \in P(\mathbf{y})$ so $\delta_{\mathbf{y}\in P(\mathbf{x})}/\sum_{\mathbf{z}\in P(\mathbf{x})} f(\mathbf{z}) = \delta_{\mathbf{x}\in P(\mathbf{y})}/\sum_{\mathbf{z}\in P(\mathbf{y})} f(\mathbf{z})$. Hence, $\sum_{\mathbf{x}\in X} S(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{x}\in X} \delta_{\mathbf{x}\in P(\mathbf{y})} f(\mathbf{x})/\sum_{\mathbf{z}\in P(\mathbf{y})} f(\mathbf{z}) = 1$. And $\sum_{\mathbf{y}\in X} S(\mathbf{x},\mathbf{y}) MR(\mathbf{y}) f(\mathbf{y}) = \mathbb{E}[MR(\mathbf{z})|\mathbf{z} \in P(\mathbf{x})] f(\mathbf{x}) = \overline{MR}(\mathbf{x}) f(\mathbf{x})$. ■

Proof of Proposition 5. If a non-decreasing $g(\mathbf{x})$ minimizes $||MR(\mathbf{x}) - g(\mathbf{x})||$ then it also minimizes $\frac{1}{2}||MR(\mathbf{x}) - g(\mathbf{x})||^2$. Take this to be the objective to which we add $\sum_{i\in N, \mathbf{x}\in N} \lambda_i(\mathbf{x}) \overline{\Delta}_i g(\mathbf{x})$ to deal with the constraint that $g(x)$ has to be non-decreasing (where $\lambda(\mathbf{x}) \geq 0$ and $\lambda(\overline{x}, \mathbf{x}_{-i}) = 0$). This term can be rearranged to yield the first-order condition $g(\mathbf{x}) = MR(\mathbf{x}) - \sum_{i\in N} \underline{\Delta}_i \lambda_i^*(\mathbf{x})/f(\mathbf{x})$ where

$$\lambda^*(\mathbf{x}) \;=\; \underset{\boldsymbol{\lambda}\,:\,X \to \mathbb{R}^n_{\geq 0}}{\operatorname{argmax}} \; \frac{1}{2}\,\mathbb{E}\Big[MR(\mathbf{x})^2 - (MR(\mathbf{x}) - \sum_{i\in N} \underline{\Delta}_i \lambda_i(\mathbf{x})/f(\mathbf{x}))^2\Big]$$

which yields the same $\lambda^*(\mathbf{x})$ as (13). Hence, $g(\mathbf{x}) = \overline{MR}(\mathbf{x})$. ■

Proof of Proposition 6. From 1, the mechanism $(q^*, \mathbf{t}^*)$ satisfies incentive compatibility constraints as $\overline{MR}(\mathbf{x})$ is non-decreasing in each coordinate by Proposition 3.

It is clear that weak duality holds for the saddle-point problem. To show that strong duality holds, we need to show that the value gap under $\overline{MR}(\mathbf{x})$ and $q^*(\mathbf{x})$ is zero. By Proposition 4, there exists a partition $\mathscr{P}$ such that when $MR(\mathbf{x}) \neq \overline{MR}(\mathbf{x})$ on $P_i \in \mathscr{P}$, $\overline{MR}(\mathbf{x})$ is constant on P_i. Thus $q^*(\mathbf{x})$ is also partition-wise constant given $MR(\mathbf{x}) \neq \overline{MR}(\mathbf{x})$. We calculate the value gap as follows:

$$\begin{aligned}
\pi(MR, q^*) - \pi(\overline{MR}, q^*) &= \mathbb{E}\left[q^*(\mathbf{x})(MR(\mathbf{x}) - \overline{MR}(\mathbf{x}))\right] \\
&= \underset{P_i\in\mathscr{P}}{\mathbb{E}}\left[\mathbb{E}\left[q^*(\mathbf{x})(MR(\mathbf{x}) - \overline{MR}(\mathbf{x}))|\mathbf{x} \in P_i\right]\right] \\
&= \underset{P_i\in\mathscr{P}}{\mathbb{E}}\left[q^*(\mathbf{x})\,\mathbb{E}\left[(MR(\mathbf{x}) - \overline{MR}(\mathbf{x}))|\mathbf{x} \in P_i\right]\right] \\
&= \underset{P_i\in\mathscr{P}}{\mathbb{E}}\left[q^*(\mathbf{x})\left[\mathbb{E}[MR(\mathbf{x})|\mathbf{x} \in P_i] - \mathbb{E}[\overline{MR}(\mathbf{x}))|\mathbf{x} \in P_i]\right]\right] \\
&= 0
\end{aligned}$$

The last equation is due to Proposition 4. We thus have

$$\pi(q, MR) \;\leq\; \pi(q, \overline{MR}) \;\leq\; \pi(q^*, \overline{MR}) \;=\; \pi(q^*, MR)$$

where the first inequality follows because $\overline{MR} \succ MR$ and q non-decreasing implies $\pi(q, MR) - \pi(q, \overline{MR}) = \mathbb{E}[q(\mathbf{x})(MR(\mathbf{x}) - \overline{MR}(\mathbf{x}))] \leq 0$, the second inequality follows from optimality of q^* for $\overline{MR}$, and the final equality follows from the zero value gap. Hence, q^* is optimal for MR. ■

Proof of Proposition 7. We first construct the partition used in the proposition and prove it is ortho-convex. First note that for any player, or any dimension, i, while fixing $\mathbf{x}_{-i}$, i's type space is partitioned into intervals P such that (i) $\lambda_i(x_i, \mathbf{x}_{-i}) = 0$ for $x_i < \min\{x'_i | (x'_i, \mathbf{x}_{-i}) \in P\}$; (ii) $\lambda_i(x_i, \mathbf{x}_{-i}) = 0$ for $x = \max\{x'_i | (x_i, \mathbf{x}_{-i}) \in P\}$; and (iii) if $|P| > 1$ then $\lambda_i(x_i, \mathbf{x}_{-i}) > 0$ for $\min\{P\} \leq x < \max\{P\}$. Denote this partition by $\mathscr{P}_i$.

The following algorithm constructs the ironed or flattened subset of X containing $P \in \mathscr{P}_i$. Let $P^1 = P$. By the Kuhn-Tucker conditions (14) and (15), any feasible mechanisms requires that $\eta_i(x_i, \mathbf{x}_{-i})q(x_i, \mathbf{x}_{-i})$ be constant in i for all $\mathbf{x} \in P$. Moreover, $\mathbb{E}(\underline{\Delta}_i \lambda_i(\mathbf{x}) | \mathbf{x} \in P^1) = 0$.

Given P^{k-1}, define

$$P^k = P^{k-1} \bigcup_{j \in N} \{\hat{P} \in \mathscr{P}_j \big| \hat{P} \not\subseteq P^{k-1}, \hat{P} \cap P^{k-1} \neq \emptyset\}.$$

The set P^k adds to P^{k-1} all intervals with positive multipliers in all dimensions that intersect with P^{k-1} (but are not already contained in P^{k-1}).

The solution to (12) must be such that $\eta_j(\cdot)q(\cdot)$ is constant on P^k. To see this, suppose $\eta_j(\cdot)q(\cdot) = c$ is constant on P^{k-1} and that there is some $j \in N$ and $\hat{P} \in \mathscr{P}_j$ such that $\hat{P} \not\subseteq P^{k-1}$ and $\hat{P} \cap P^{k-1} \neq \emptyset$ – i.e. an interval in dimension j that intersects with and is not inside P^{k-1}. Then $\eta_j(\cdot)q(\cdot) = c$ on $\hat{P} \cap P^{k-1}$ by assumption. Using conditions (14) and (15) in dimension j, this extends to the entire set $\hat{P}$.

Each iterative set is ortho-convex, after possibly adding some knife-edge type profiles where the monotonicity constraint just binds: take $(x_j, \mathbf{x}_{-j}) \in P^k$ and $(x'_j, \mathbf{x}_{-j}) \in P^k$ with $x'_j > x_j$. Since any incentive compatible mechanism requires that $\eta_j(\cdot)q_j(\cdot)$ be non-decreasing in dimension j, $\eta_j(x'_j, \mathbf{x}_{-j})q(x'_j, \mathbf{x}_{-j}) \geq \eta_j(\hat{x}_j, \mathbf{x}_{-j})q(\hat{x}_j, \mathbf{x}_{-j}) \geq \eta_j(x_j, \mathbf{x}_{-j})q(x_j, \mathbf{x}_{-j})$ for any $\hat{x}_j \in (x_j, x'_j)$. But $\eta_j(x'_j, \mathbf{x}_{-j})q(x'_j, \mathbf{x}_{-j}) = \eta_j(x_j, \mathbf{x}_{-j})q(x_j, \mathbf{x}_{-j})$ so $\eta_i(\hat{x}_j, \mathbf{x}_{-j})q(\hat{x}_j, \mathbf{x}_{-j}) = \eta_i(x_j, \mathbf{x}_{-j})q(x_j, \mathbf{x}_{-j})$ for any $\hat{x}_j \in (x_j, x'_j)$. If the interval $[x_j, x'_j] \times \{\mathbf{x}_{-j}\} \not\subseteq P^k$, add it to the set before moving to the next step.

The algorithm ends when the set $\{\hat{P} \in \mathscr{P}_j \big| \hat{P} \not\subseteq P^k, \hat{P} \cap P^k \neq \emptyset\}$ is empty for all players (i.e. when no intervals not inside the iterative set intersect with the iterative set). Call the resulting set P. Then P is ortho-covex and $\mathbb{E}(\sum_{j \in N} \underline{\Delta}_j \lambda_j(\mathbf{x}) | \mathbf{x} \in P) = 0$. To see the latter, suppose k was the final step in the algorithm and that $\{\hat{P} \in \mathscr{P}_l \big| \hat{P} \not\subseteq P^k, \hat{P} \cap P^k \neq \emptyset\}$ is empty for all players $l \in N$ but there exists $j \in N$ such that $\mathbb{E}(\underline{\Delta}_j \lambda_j(\mathbf{x}) | \mathbf{x} \in P^k) \neq 0$. Then, there must be $\mathbf{x} \in P^k \cap \hat{P}$ for some $\hat{P} \in \mathscr{P}_j$ by definition of $\mathscr{P}_j$. But, since $\mathbb{E}(\underline{\Delta}_j \lambda_j(\mathbf{x}) | \mathbf{x} \in P') = 0$ for any $P' \in \mathscr{P}_j$, it must be that $\hat{P} \not\subseteq P^k$. Therefore $\hat{P} \cap P \neq \emptyset$ and $\hat{P} \not\subseteq P$, contradicting the assumption that $\{\hat{P} \in \mathscr{P}_j \big| \hat{P} \not\subseteq P^{k-1}, \hat{P} \cap P^{k-1} \neq \emptyset\}$ is empty.

Repeating this algorithm for all $P \in \mathscr{P}_i$ and all $i \in N$ results in a partition of the type space; denote this partition by $\mathscr{P}$. Define $P(\mathbf{x}) = \{\mathbf{y} \in P | \mathbf{x} \in P, P \in \mathscr{P}\}$.

Next, we show the saddle-point problem for restricted rights has the strong duality property under the choice of $q^*(\mathbf{x})$ and $\eta_i^*(\mathbf{x})$. From above, we have that for any cell $P_i \in \mathscr{P}_i$, P_i is contained in one cell of $\mathscr{P}$, and, for fixed $\mathbf{x}_{-i}$, $q^* \eta_i^*$ is constant on P_i. Because on each ironed interval P_i, $\mathbb{E}[\underline{\Delta}_i \lambda_i^*(x_i, \mathbf{x}_{-\mathbf{i}}) | (x_i, \mathbf{x}_{-\mathbf{i}}) \in P_i] = 0$, we have

$$\mathbb{E}[\widetilde{MR}_i(x_i, \mathbf{x}_{-\mathbf{i}}) | (x_i, \mathbf{x}_{-\mathbf{i}}) \in P_i] = \mathbb{E}[MR_i(x_i, \mathbf{x}_{-\mathbf{i}}) | (x_i, \mathbf{x}_{-\mathbf{i}}) \in P_i].$$

We calculate the value gap between the saddle-point problem and the primal problem under $(q^*, \boldsymbol{\eta}^*, \boldsymbol{\lambda}^*)$:

$$\begin{aligned}
&\pi(\widetilde{\mathbf{MR}}, q^*, \boldsymbol{\eta}^*) - \pi(\widetilde{\mathbf{MR}}, q^*, \boldsymbol{\eta}^*) \\
&\quad= \mathbb{E}\left[q^*(\mathbf{x})\sum_i \widetilde{MR}_i(\mathbf{x})\eta_i^*(\mathbf{x})\right] - \mathbb{E}\left[q^*(\mathbf{x})\sum_i MR_i(\mathbf{x})\eta_i^*(\mathbf{x})\right] \\
&\quad= \mathbb{E}\left[\sum_i q^*(\mathbf{x})\eta_i^*(\mathbf{x})(\widetilde{MR}_i(\mathbf{x}) - MR_i(\mathbf{x}))\right] \\
&\quad= \sum_i \mathbb{E}\left[q^*(\mathbf{x})\eta_i^*(\mathbf{x})(\widetilde{MR}_i(\mathbf{x}) - MR_i(\mathbf{x}))\right] \\
&\quad= \sum_i \mathbb{E}\left[\mathbb{E}[q^*(x_i, \mathbf{x}_{-i})\eta_i^*(x_i, \mathbf{x}_{-\mathbf{i}})(\widetilde{MR}_i(x_i, \mathbf{x}_{-\mathbf{i}}) - MR_i(x_i, \mathbf{x}_{-i}))|(x_i, \mathbf{x}_{-i}) \in P_i]\right] \\
&\quad= \sum_i \mathbb{E}\left[q^*(x_i, \mathbf{x}_{-i})\eta_i^*(x_i, \mathbf{x}_{-i})\,\mathbb{E}[\widetilde{MR}_i(x_i, \mathbf{x}_{-i}) - MR_i(x_i, \mathbf{x}_{-i})|(x_i, \mathbf{x}_{-i}) \in P_i]\right] = 0
\end{aligned}$$

We thus have

$$\pi(\mathbf{MR}, q, \boldsymbol{\eta}) \;\le\; \pi(\widetilde{\boldsymbol{MR}}, q, \boldsymbol{\eta}) \;\le\; \pi(\widetilde{\boldsymbol{MR}}, q^*, \boldsymbol{\eta}^*) \;=\; \pi(\boldsymbol{MR}q^*, \boldsymbol{\eta}^*)$$

where the first inequality follows because, for every $i \in N$, $\widetilde{MR}_i \succ_i MR_i$ and $q(\cdot)\eta_i(\cdot)$ is non-decreasing implies $\mathbb{E}[q(\mathbf{x})\eta_i(\mathbf{x})(MR_i(\mathbf{x}) - \widetilde{MR}_i(\mathbf{x}))] \le 0$, the second inequality follows from optimality of $(q^*, \boldsymbol{\eta}^*)$ for $\widetilde{\boldsymbol{MR}}$, and the final equality follows from the zero value gap. Hence, $(q^*, \boldsymbol{\eta}^*)$ is optimal for $\boldsymbol{MR}$. Therefore, strong duality holds for the saddle-point problem with restricted access rights, $(q^*, \boldsymbol{\eta}^*, \boldsymbol{\lambda}^*)$ is a saddle-point for the Lagrangian, and $(q^*, \boldsymbol{\eta}^*)$ solves the primal problem. ■

Lemma 2 *If there exists two solutions, $\widetilde{\boldsymbol{MR}}$ and $\widetilde{\boldsymbol{MR}}'$, to (12) then the associated optimal mechanisms $(q^*, \boldsymbol{\eta}^*, \mathbf{t}^*)$ and $(\hat{q}^*, \hat{\boldsymbol{\eta}}^*, \hat{\mathbf{t}}^*)$ are identical.*

Proof of Lemma 2. Since both $\widetilde{\boldsymbol{MR}}$ and $\widetilde{\boldsymbol{MR}}'$ are assumed to achieve the maximum in problem (12) and since $\sum_{i\in N} \max\{0, g_i(\mathbf{x})\}^2$ is strictly convex over the positive range of g_i, we can assume that $\widetilde{MR}_i(\mathbf{x}) \le 0$ and $\widetilde{MR}_i'(\mathbf{x}) \le 0$ whenever $\widetilde{MR}_i(\mathbf{x}) \ne \widetilde{MR}_i'(\mathbf{x})$. Otherwise, for the convex combination $\widetilde{\boldsymbol{MR}}'' = \frac{1}{2}(\widetilde{\boldsymbol{MR}} + \widetilde{\boldsymbol{MR}}')$,

$$\mathbb{E}\Big[\big(\sum_{i\in N} \max(0, \widetilde{MR}_i(\cdot))\big)^2\Big] > \mathbb{E}\Big[\big(\sum_{i\in N} \max(0, \widetilde{MR}_i''(\cdot))\big)^2\Big]$$

and $\widetilde{MR}_i'' \succ MR_i$. Moreover, if $\widetilde{MR}_i(\mathbf{x}) < 0$ and $\widetilde{MR}_i'(\mathbf{x}) < 0$ whenever $\widetilde{MR}_i(\mathbf{x}) \ne \widetilde{MR}_i'(\mathbf{x})$, then $\eta_i(\mathbf{x}) = \hat{\eta}_i(\mathbf{x}) = 0$.

Suppose then that there exists $(x_i, \mathbf{x}_{-i}) \in X$ such that $\widetilde{MR}_i'(x_i, \mathbf{x}_{-i}) < \widetilde{MR}_i(x_i, \mathbf{x}_{-i}) \le 0$ and without loss of generality suppose it is the largest type with $\widetilde{MR}_i'(x_i, \mathbf{x}_{-i}) \ne \widetilde{MR}_i(x_i, \mathbf{x}_{-i})$. Then

$\hat{\eta}_i^*(\mathbf{x}) = 0$ and $\hat{q}^*(x_i', \mathbf{x}_{-i})\hat{\eta}_i^*(x_i', \mathbf{x}_{-i}) = 0$ for any $x_i' < x_i$. But this implies that $\widetilde{MR}_i'(x_i, \mathbf{x}_{-i}) \leq 0$ for all $x_i' < x_i$ (since otherwise Proposition 7 would require that $\hat{\eta}_i^*(x_i', \mathbf{x}_{-i}) > 0$). Similarly, if $\widetilde{MR}_i(x_i, \mathbf{x}_{-i}) < 0$ then $q^*(x_i', \mathbf{x}_{-i})\eta_i^*(x_i', \mathbf{x}_{-i})$ and we are done.

Suppose instead that $\widetilde{MR}_i(x_i, \mathbf{x}_{-i}) = 0$. There exists x_i' such that $\widetilde{MR}(x_i', \mathbf{x}_{-i}) < \widetilde{MR}_i'(x_i', \mathbf{x}_{-i}) \leq 0$ since $\widetilde{MR}_i$ and $\widetilde{MR}_i'$ must have the same expectation below x_i. Since x_i is the largest type for which the two solutions differ, $\lambda_i(x_i, \mathbf{x}_{-i}) = \hat{\lambda}_i(x_i, \mathbf{x}_{-i})$. Further, since $\widetilde{MR}_i'(x_i, \mathbf{x}_{-i}) < \widetilde{MR}_i(x_i, \mathbf{x}_{-i})$, $\lambda_i(x_i^-, \mathbf{x}_{-i}) > \hat{\lambda}_i(x_i^-, \mathbf{x}_{-i}) \geq 0$; that is, the monotonicity constraints binds between for x_i^- and x_i. Furthermore, either $\lambda_i((x_i^-)^-, \mathbf{x}_{-i}) > \hat{\lambda}_i((x_i^-)^-, \mathbf{x}_{-i}) \geq 0$ or $\widetilde{MR}_i(x_i^-, \mathbf{x}_{-i}) < \widetilde{MR}_i'(x_i^-, \mathbf{x}_{-i}) \leq 0$. In words, either the $\widetilde{MR}_i$ is negative for the type just below x_i or the monotonicity constraints binds between for $(x_i^-)^-$, x_i^- and x_i. Iterating in this way, we can conclude that $\lambda_i(\tilde{x}_i, \mathbf{x}_{-i}) > 0$ for all $\tilde{x}_i \in [x_i', x_i^-]$ which implies that $\eta_i^*(\tilde{x}_i, \mathbf{x}_{-i})q^*(\tilde{x}_i, \mathbf{x}_{-i})$ is constant over this interval. Since $\widetilde{MR}_i(x_i^-, \mathbf{x}_{-i}) < 0$, Proposition 7 requires $\eta_i^*(\tilde{x}_i, \mathbf{x}_{-i})q^*(\tilde{x}_i, \mathbf{x}_{-i}) = 0$. ■

References

Cornelli, F.: 1996, 'Optimal Selling Procedures with Fixed Costs'. *Journal of Economic Theory* **71**(0106), 1–30.

Csapó, G. and R. Müller: 2013, 'Optimal mechanism design for the private supply of a public good'. *Games and Economic Behavior* **80**, 229–242.

Figueroa, N. and V. Skreta: 2011, 'Optimal allocation mechanisms with single-dimensional private information'. *Review of Economic Design* **15**(3), 213–243.

Goeree, J. K. and A. Kushnir: 2011, 'A geometric approach to mechanism design'. SSRN working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1974922.

Groves, T. and J. O. Ledyard: 1977, 'Optimal allocation of public goods: a solution to the "free rider" problem'. *Econometrica* **45**(4), 783–810.

Güth, W. and M. Hellwig: 1986, 'The private suply of a public good'. *Journal of Economics* pp. 121–159.

Hardy, G. H., J. E. Littlewood, and G. Pólya: 1929, 'Some simple inequalities satisfied by convex functions'. *Messenger of Mathematics* **58**, 145–152.

Hellwig, M.: 2003, 'Public-good provision with many participants'. *The Review of Economic Studies* **70**(3), 589–614.

Jehiel, P. and B. Moldovanu: 2001, 'Efficient design with interdependent valuations'. *Econometrica* **69**(5), 1237–1259.

Jehiel, P., B. Moldovanu, and E. Stacchetti: 1999, 'Multidimensional mechanism design for auctions with externalities'. *Journal of Economic Theory* **85**, 258–293.

Kleiner, A., B. Moldovanu, and P. Strack: 2020, 'Extreme Points and Majorization: Economic Applications'. SSRN working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3551258.

Ledyard, J. O. and T. R. Palfrey: 1999, 'A characterization of intermin efficiency with public goods'. *Econometrica* **67**(2), 435–448.

Ledyard, J. O. and T. R. Palfrey: 2007, 'A general characterization of interim efficient mechanisms for independent linear environments'. *Journal of Economic Theory* **133**, 441–466.

Malaith, G. J. and A. Postlewaite: 1990, 'Asymmetric information bargaining problems with many agents'. *The Review of Economic Studies* **57**(3), 351–367.

Marshall, A. W., I. Olkin, and B. Arnold: 2010, *Inequalities: Theory of majorization and its applications*. Springer Science & Business Media.

Milgrom, P. R. and R. J. Weber: 1982, 'A theory of auctions and competitive bidding'. *Econometrica* **50**(5), 1089–1122.

Muirhead, R. F.: 1903, 'Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters'. *Proc. Edinburgh Math. Soc.* **21**, 144–157.

Mussa, M. and S. Rosen: 1978, 'Monopoly and Product Quality'. *Journal of Economic Theory* **18**, 301–317.

Myerson, R. B.: 1981, 'Optimal Auction Design'. *Mathematics of Operations Research* **6**(1), 58–73.

Norman, P.: 2004, 'Efficient mechanisms for public good with exclusions'. *The Review of Economic Studies* **71**(4), 1163–1188.

Rochet, J.-C. and P. Choné: 1998, 'Ironing, sweeping, and multidimensional screening'. *Econometrica* **66**(4), 783–826.

Roughgarden, T. and I. Talgam-Cohen: 2016, 'Optimal and robust mechanism design with interdependent values'. *ACM Transactions on Economics and Computation* **4**(18), 18:1–18:34.