Necessity of Auctions for Redistributive Optimality*

Mingshi Kang† Charles Z. Zheng‡

August 28, 2019

Abstract

Two items, one commonly desirable, the other commonly disliked, may be assigned to n asymmetrically distributed players, whose types determine their marginal rates of substitution for money. This paper characterizes the set of all interim Pareto optimal mechanisms. Despite the infinite dimensions of interim Pareto optimality, any such optimum is in the form of auctions, with the winner-selection rule adjusted to the particular optimum. All surplus from the auctions is rebated to the players and goes only to those most favored by the welfare ranking associated with the Pareto optimum, who need not be the stochastically high- or low-type players. A player's equilibrium payoff is roughly U-shape in his type, hence type-inequality does not beget payoff-inequality. Optimal matching also obtains because there is zero probability for both items to go to the same player. In characterizing the optimal mechanisms we develop a new kind of operators to incorporate every player's endogenously countervailing incentive.

JEL Classification: C61, D44, D82

Keywords: mechanism design, optimal auction, redistribution, interim Pareto optimal mechanisms, countervailing incentive, ironing

---

*We thank Victor Aguiar, Roy Allen, Yi Chen, Scott Kominers, Alexey Kushnir, Greg Pavlov, Edward Schlee, and the seminar participants in the 2019 North American Econometric Society Summer Meetings, the 2019 Canadian Economic Theory Conference and the 2019 International Conference of Game Theory at Stony Brook, for their insights and comments. Zheng acknowledges financial support from the Social Science and Humanities Research Council of Canada.

†Department of Economics, University of Western Ontario, London, ON, Canada, mkang94@uwo.ca.

‡Department of Economics, University of Western Ontario, London, ON, Canada, charles.zheng@uwo.ca, https://sites.google.com/site/charleszhenggametheorist/.

1 Introduction

This paper is motivated by the question how to induce Pareto improving wealth transfers across individuals. In order for wealth transfer to be Pareto improving, let us consider an environment where individuals may have different marginal rates of substitution for money. To induce voluntary wealth transfers, suppose that the social planner has two items, one good, the other bad, to assign to n players. For example, she needs to locate among n cities a high-tech giant's headquarter and an oil pipeline terminal. If the social planner sells the good to a player who values money less and uses the revenue to pay another who values money more to take the bad, a Pareto improving wealth transfer is induced. Such a transfer, however, is only one instance among a large variety of redistribution schemes that a social planner may deem Pareto improving. Depending on her value judgement, the social planner may favor one player against another, or favor one type of a player against another type of the same player, whether or not the former values money intrinsically more than the latter does. Thus, we assume no stand on interpersonal comparison, as one dollar for one type of a player may be deemed by the social planner more valuable than one dollar for another type of the same or a different player. Rather we consider the entire set of interim Pareto optimal mechanisms, without assuming the existence of any rule according to which a social planner assigns welfare weights across players and across types of a player. That is, we shall find out the common features of all the Pareto optima not only among all players but also among all types of each player. The latter aspect makes this study relevant not only to mechanism design but also to macro settings where types in a continuum are interpreted as atomless individuals and players interpreted as regions or other collections of individuals.

The model has n players, whose types are independently drawn from possibly different distributions. The positive values of the good, and the negative values of the bad, are commonly known. A player's type determines his marginal rate of substitution of money. Any mechanism committed to by the social planner is subject to the standard constraints: incentive compatibility (IC), (interim) individual rationality (IR) and (ex post) budget balance (BB). Interim Pareto improvement means an IC, IR and BB mechanism that makes a positive measure of some player's types better-off, and zero measure of every player's types worse-off, than the status quo. Interim Pareto optimality means IC, IR, BB and immunity to interim Pareto dominance. The problem is to characterize the set of all interim Pareto optimal mechanisms and identify their common features.

This problem is novel to the mechanism design literature because our design objective, interim Pareto optimality, has infinite dimensions, whereas the objective in mechanism design has only one dimension such as an auctioneer's expected revenue or a planner's social welfare function aggregating individual preferences through exogenous welfare weights.[1] Another feature of the problem is each player's endogenously countervailing incentive: Depending on what the mechanism entails contingent on his realized type, a player may act as a buyer of the good sometimes, and as a recipient of the bad some other times. He would underreport his willingness to pay in the former event, and exaggerate his cost in the latter. By contrast, in the literatures of optimal auctions (Myerson [11]), optimal taxation (Mirrlees [10]) and bilateral trade (Myerson and Satterthwaite [12]), the role of a player is exogenous.

Our solution to this problem says that any interim Pareto optimal mechanism is necessarily in the form of auctions, with the winner-selection rule adjusted to the particularity of the optimum. First, for any interim Pareto optimum there is an associated *welfare ranking*, as a profile of type-distributions across players, that aggregates individual preferences into a unidimensional social welfare function which the Pareto optimum maximizes subject to IC, IR and BB (Theorem 1). Second, the associated welfare ranking, coupled with the Lagrange multiplier for the joint constraint of IC, IR and BB, determines a rule to select the winner of the good, and another rule to select the "winner" of the bad, that the given Pareto optimum entails. These winner-selection rules together determine each player's expected value of money transfers in the Pareto optimal mechanism up to a constant, and the constant is determined by the expectation of the player's marginal value of money measured by the associated welfare-ranking distribution (Theorem 2).

This general characterization has several implications. The first is about redistribution across players. An interim Pareto optimal mechanism entails lump sum transfers across players, but the transfer need not go to those whose types are stochastically low (or stochastically high). Rather, the lump sum transfer goes only to those who are most favored ex ante by the endogenous welfare ranking associated with the particular Pareto optimum, and

[1] Should each player have only finitely many types, it would be trivial to reduce the finite-dimensional Pareto optimality to a unidimensional objective, but the finite-type assumption would restrict the model, undermine its relevance to macro considerations of a continuum of agents, and make it hard to relate to much of the mechanism design literature, where most of the elegant results is based on continuum types.

the exogenous type-distributions have no bearing on the welfare ranking (Section 3.3.1).

The second implication is about redistribution across types. Any interim Pareto optimal mechanism that assigns the bad with a positive probability entails non-monotone interim expected payoff functions for all players, with types at the high and low ends enjoying larger surpluses than those in the middle (Figure 4). Thus, in the symmetric case where all players are treated identically so that they differ only in types, there is no way to reorder the types for the ranking on the players in terms of their types to propagate to a ranking on them in terms of their payoffs. That is, an interim Pareto optimal mechanism breaks the linkage from type-inequality to payoff-inequality (Corollary 1).

Third, in any interim Pareto optimal mechanism, a player-type that has a positive probability to get the good has zero probability to get the bad, and vice versa. In other words, any interim Pareto optimum in our model is also an interim Pareto optimum in the matching environment that disallows a player to have both items (Corollary 2). This result complements the new literature of matching with transfers (Chiappori [1]) by showing that a particular auction mechanism achieves optimal matching.[2]

Fourth, the winner-selection rule in an interim Pareto optimal mechanism often involves ironing when players' realized types belong to some intervals, regardless of the functional form of his type-distribution (Section 3.3.4). Consequently, efficient allocation, a main goal in mechanism design, is in general suboptimal. When the players are i.i.d. and treated equally by the social planner's welfare ranking, ironing takes the form of lotteries. Such prevalence of ironing stands in contrast to the optimal auction literature, where ironing can be avoided by assuming regular type-distributions.

Our method to obtain these results has two novel aspects. The first is to reduce the infinite-dimensional objective, interim Pareto optimality, to a unidimensional objective that allows for the calculation of optimal mechanisms (Section 4.1). Although this step is essentially an application of the Hahn-Banach theorem, we need to resolve a dilemma, present in infinite-horizon macro models, between ensuring existence of a separating hyperplane and guaranteeing that the hyperplane can be properly represented (cf. Stokey, Lucas and Prescott [14, §15.4, §16.6]). This dilemma is resolved in our model because the separating hyperplane here needs only to be represented as a distribution rather than an inner product operator (as in the macro models), and the hyperplane can be represented as a distribution

[2] According to Herodotus [4], auctions were used in ancient Babylon marriage matching markets.

because of a basic continuity observation in mechanism design.

The second novelty in our method is a new kind of operators that calculate every player's information rent bifurcated by his endogenously countervailing incentive (Section 4.2.1). With such operators, we reduce any interim Pareto optimal mechanism to a solution of a maximization problem subject to only two constraints, one being the standard monotonicity condition, the other a combined constraint of IC, IR and BB (Lemma 4). A useful property of such operators guarantees that any solution of this constrained optimization problem satisfies the saddle point condition (Lemma 5). This property also provides the basis for us to solve the Lagrange problem associated with the saddle point condition, despite nonlinearity of the Lagrangian due to endogenously countervailing incentives (Section 4.3.2). Our solution of that Lagrange problem is an extension of the ironing technique in mechanism design to handle countervailing incentives and multiplicity of binding constraints.

Our modeling choice of individual preferences and types follows the idea in Dworczak, Kominers and Akbarpour [3] that captures wealth inequality by heterogeneous marginal rates of substitution (MRS) for money and meanwhile maintains quasilinear preferences for tractability. Dworczak et al. suggest that quasilinearity at the presence of wealth inequality is an appropriate local approximation when one's valuation of money is a smooth function of his wealth. Even when wealth imposes a hard budget constraint on an individual, a recent finding by Khan and Schlee [5] suggests that quasilinearity might still obtain through a saddle point reformulation.

Considering a bilateral trade environment, Dworczak et al. [3] characterize the set of mechanisms that maximize the sum of the integrals across agents' utilities given exogenous welfare density functions of the agents (same as types in their model) such that the welfare density functions can be arbitrary. They use a novel technique and observe that the optimal design uses tax-like pricing mechanisms, with a wedge between the price for the buyers and that for the sellers, and lump sum transfers to the poorer agents. This paper builds upon their idea of capturing wealth inequality by the notion of heterogeneous MRS in a quasilinear setting. Our model differs from theirs in four aspects. First, we do not assume a unidimensional, utilitarian design objective such as a sum or an integral of utilities across types or agents; rather, the associated welfare ranking that aggregates preferences, across types and across players, is a consequence of the Pareto optimum under consideration (and the welfare ranking need not be representable as an inner product operator with its densities).

Second, we have n players whose types are drawn from possibly different distributions; in their model, there is a continuum of i.i.d. buyers, and a continuum of i.i.d. sellers. Third, in our model a player's role—whether to be a seller or to be a buyer—is endogenous and hence has countervailing incentives, whereas in their model an agent's role is exogenously assumed. Fourth, the items in our model need not be assigned and hence the probability of assigning the good need not be equal to the probability of assigning the bad; in their model, market clearance requires that the aggregate probability of sales be equal to that of purchases.

Because of the first difference, this paper complements Dworczak et al. with our Theorem 1, which suggests that with a similar separating hyperplane argument their assumption of exogenous welfare densities might be relaxable. Because of the other differences, Pareto optima in our model are all auction-like mechanisms rather than the tax-like mechanisms in their model. A player's bid in our model affects the type-cutoffs for other players to receive an item, whereas in their model an agent, atomless, has no influence on others. Although our optimal mechanisms also entail lump sum transfers, the beneficiaries of the transfers in our case need not be the stochastically poorer players but rather are those who are most favored by the endogenous welfare ranking associated with the Pareto optimum. Because of the third difference, a player's surplus in our model is a non-monotone function of his type, while in theirs it is monotone. Applied to symmetric cases, this non-monotonicity implication says that our Pareto optimal mechanisms breaks the type-generated hierarchy through giving higher surpluses to types near the high and low ends than to those in the middle.

Countervailing incentives have been considered in the partnership dissolution literature, initiated by Cramton, Gibbons and Klemperer [2]. The focus of that literature is implementability of one particular winner-selection rule, the efficient allocation, which would be optimal if implementable and if the objective is the simple sum of surpluses across players. Loertscher and Wasser [7], differently, consider a design objective that is a convex combination between the auctioneer's expected revenue and the expected utility of the good for its final owner. Since the total money transfer from the players to the auctioneer is a plus rather than a negative in that objective,[3] their optimal mechanism does not rebate surplus to players but rather squeezes the lowest surplus for each player down to the player's exogenous

[3]The expected utility of the good for its final owner (called social surplus by Loertscher and Wasser) is not equal to the total surplus among all players. That is because the total money transfer from the players to the auctioneer is not subtracted from the expected utility of the good for its final owner.

outside option. This outside option equation is crucial in Loertscher and Wasser's solution of the countervailing incentive problem. Our paper differs from the partnership dissolution literature by charactering the entire set of interim Pareto optimal mechanisms. With players heterogenous in MRS for money, our counterpart of the efficient allocation is suboptimal even if it is implementable (Section 3.3.4). The paper differs from Loertscher and Wasser also in the first and fourth aspects in which we differ from Dworczak et al. Consequently, our optimal mechanisms do rebate surplus back to players and do not squeeze the lowest surplus for each player down to the player's exogenous outside option. Hence Loertscher and Wasser's outside option equation is unavailable to us. Conversely, however, our new operator is applicable to their model, with or without their assumption of regular type-distributions, and delivers their Lagrangian, though our bisection technique is inapplicable to maximization of their Lagrangian because of the fourth difference listed previously.

The following Section 2 defines the model and the design problem. Section 3 then presents the main results and implications. Section 4 presents the proofs, with some details relegated to the Appendix. The new operator is introduced in Section 4.2.1. Our extension of the ironing technique is in Section 4.3.2 and Appendix B.6. Section 5 concludes.

2 The Model

2.1 The Good, the Bad, and n Players

Two items, named A and B, each indivisible, are to be allocated among n players ($n \geq 2$), each of whom can get one or both or none of the items. A social planner commits to a mechanism that may allocate the items to some players and may mandate money transfers among the players. The outcome to any player i takes the form (x_{iA}, x_{iB}, y_i), where x_{ij} ($\forall j \in \{A, B\}$) denotes the probability with which player i gets item j, and y_i the net money transfer from player i to others (with negative y_i signifying the transfer from others to i). After the mechanism is announced and before it is operated, each player, given his own private information, can opt out of the mechanism thereby getting the outcome $(0, 0, 0)$ for himself. Each player i's preference relation is represented by a vNM utility function

$$(x_{iA}, x_{iB}, y_i) \mapsto x_{iA} - c x_{iB} - \frac{y_i}{t_i}, \tag{1}$$

with $c \geq 0$ a constant across all players, and t_i player i's realized *type*. Thus, item A is interpreted as a good, and item B a bad, to all players; $1/t_i$ may be interpreted as player i's marginal rate of substitution for money.

Assume that each player i's type t_i is independently drawn from a commonly known cumulative distribution function (CDF) F_i such that its support is $T_i := [a_i, b_i]$, its density f_i is positive on the support, and $a_i > 0$.

Remark 1 Without changing the results, one can generalize (1) to

$$(x_{iA}, x_{iB}, y_i) \mapsto (v_i x_{iA} - c_i x_{iB})\, \omega_i - y_i \theta_i,$$

where v_i and c_i are commonly known, positive constants, and $(\omega_i, \theta_i) \in \mathbb{R}^2_{++}$ is drawn from a joint distribution whose realization is privately known to player i. That is, the intensities of the good and the bad may vary across players, and each player has two-dimensional types. Dworczak, Kominers and Akbarpour [3] have shown that there is no loss of generality to restrict attention to mechanisms whose message space for player i consists only of the realized values of $t_i := \omega_i/\theta_i$. Thus, the more general vNM utility function is simplified to

$$(x_{iA}, x_{iB}, y_i) \mapsto v_i x_{iA} - c_i x_{iB} - y_i/t_i.$$

Furthermore, since we want a general characterization of Pareto optima and hence take no stand on interpersonal comparison of utilities, there is no loss of generality to simplify the above utility function further by replacing the player-specific coefficients v_i and c_i with constants v and c common across i, and normalizing the constant v to one. Thus (1) obtains.

2.2 Allocations and Mechanisms

For each player i, denote $T_{-i} := \prod_{j \neq i} T_j$, and let F_{-i} be the product measure on T_{-i} generated by $(F_j)_{j \neq i}$. An *ex post allocation* means a list $(q_{iA}, q_{iB})_{i=1}^n$ of functions such that $q_{iA}, q_{iB} : \prod_{j=1}^n T_j \to [0,1]$ for each i and, for each $t \in \prod_{j=1}^n T_j$,

$$\sum_i q_{iA}(t) \leq 1 \quad \text{and} \quad \sum_i q_{iB}(t) \leq 1.$$

An *ex post payment rule* means a list $(p_i)_{i=1}^n$ of functions such that $p_i : \prod_j T_j \to \mathbb{R}$ for each i. By the revelation principle, any equilibrium-feasible mechanism corresponds to a pair of ex post allocation $(q_{iA}, q_{iB})_{i=1}^n$ and ex post payment rule $(p_i)_{i=1}^n$, with $q_{ij}(t)$ interpreted as the probability with which item j ($j \in \{A, B\}$) is assigned to player i, and $p_i(t)$ the net money transfer from player i to others, when t is the profile of alleged types across players.

A list $(Q_i)_{i=1}^n$ of functions $Q_i : T_i \to \mathbb{R}$ ($\forall i$) is said *generated* by an ex post allocation $(q_{iA}, q_{iB})_{i=1}^n$ iff, for each $i = 1, \dots, n$ and each $t_i \in T_i$,

$$Q_i(t_i) = \int_{T_{-i}} q_{iA}(t_i, t_{-i}) dF_{-i}(t_{-i}) - c \int_{T_{-i}} q_{iB}(t_i, t_{-i}) dF_{-i}(t_{-i}). \tag{2}$$

Likewise, a list $(P_i)_{i=1}^n$ of functions $P_i : T_i \to \mathbb{R}$ ($\forall i$) is said *generated* by an ex post payment rule $(p_i)_{i=1}^n$ iff, for each $i = 1, \dots, n$ and each $t_i \in T_i$, $P_i(t_i) = \int_{T_{-i}} p_i(t_i, t_{-i}) dF_{-i}(t_{-i})$. Any list $(Q_i, P_i)_{i=1}^n$ such that $(Q_i)_{i=1}^n$ is generated by some ex post allocation $(q_{iA}, q_{iB})_{i=1}^n$, and $(P_i)_{i=1}^n$ generated by some ex post payment rule $(p_i)_{i=1}^n$, is called *reduced-form mechanism*, or *mechanism* for short.

2.3 Constraints

Given any (reduced-form) mechanism $(Q_i, P_i)_{i=1}^n$, it follows from the vNM utility function (1), and the shorthand (2), that the interim expected utility for any type $t_i \in T_i$ of player i to act type $\hat{t}_i$, given truthtelling from others, is equal to $Q_i(\hat{t}_i) - P_i(\hat{t}_i)/t_i$. Denote

$$U_i(t_i \mid Q, P) := \max_{\hat{t}_i \in T_i} Q_i(\hat{t}_i) - P_i(\hat{t}_i)/t_i. \tag{3}$$

Since $\inf T_i > 0$ by assumption, the maximization problem in (3) is equivalent to

$$\tilde{U}_i(t_i \mid Q, P) := \max_{\hat{t}_i \in T_i} t_i Q_i(\hat{t}_i) - P_i(\hat{t}_i). \tag{4}$$

Thus, as is routine in auction theory, *incentive compatibility* (IC) of $(Q_i, P_i)_{i=1}^n$ is equivalent to simultaneous satisfaction of two conditions for each player i: (i) Q_i is weakly increasing on T_i; (ii) for any $t_i, t_i^0 \in T_i$,

$$P_i(t_i) - P_i(t_i^0) = \int_{t_i^0}^{t_i} s dQ_i(s). \tag{5}$$

Since each player can opt out of a mechanism before it operates, his outside payoff is zero, hence $(Q_i, P_i)_{i=1}^n$ is said *individually rational* (IR) iff $U_i(t_i|Q, P) \geq 0$ for all i and all $t_i \in T_i$. By (3) and (4), $U_i(t_i|Q, P) = \tilde{U}_i(t_i \mid Q, P)/t_i$ for any $t_i \in T_i$, and it is routine to show that $\tilde{U}_i(\cdot \mid Q, P)$ is convex, with derivative almost everywhere equal to Q_i, which is weakly increasing by IC. Thus, $\tilde{U}_i(\cdot \mid Q, P)$ attains its minimum at

$$\tau(Q_i) := \inf \{t_i \in T_i : Q_i(t_i) \geq 0 \text{ or } t_i = b_i\}. \tag{6}$$

Consequently, $\tilde{U}_i(\tau(Q_i) \mid Q, P) \geq 0$ iff "$\tilde{U}_i(t_i \mid Q, P) \geq 0$ for all $t_i \in T_i$" iff "$U_i(t_i \mid Q, P) \geq 0$ for all $t_i \in T_i$." Thus, IR is equivalent to $\tilde{U}_i\left(\tau(Q_i) \mid Q, P\right) \geq 0$ for all players i.

For the society consisting of the n players to transfer wealth among themselves without relying on outside subsides, we require that a mechanism be always budget-balanced: $(Q_i, P_i)_{i=1}^n$ satisfies *budget balance* (BB) iff $(P_i)_{i=1}^n$ is generated by some ex post payment rule $(p_i)_{i=1}^n$ such that $\sum_i p_i(t) \geq 0$ for all $t \in \prod_i T_i$.

2.4 The Problem

To characterize a large class of Pareto optimal mechanisms, we use a strong notion of Pareto dominance based on interim, rather than ex ante, expected payoffs. A mechanism (Q^*, P^*) is *interim Pareto optimal* iff (i) (Q^*, P^*) is IC, IR and BB, and (ii) there does not exist any IC, IR and BB mechanism (Q, P) such that $u_i(\cdot \mid Q, P) \geq u_i(\cdot \mid Q^*, P^*)$ a.e. on T_i for all $i \in \{1, \ldots, n\}$ and, for some i, $u_i(\cdot \mid Q, P) > u_i(\cdot \mid Q^*, P^*)$ on a positive-measure subset of T_i. The problem is to characterize the set of all interim Pareto optimal mechanisms.

With interim Pareto optimality the welfare criterion, not only do we take no stand a priori regarding interpersonal comparison, we also allow for any inter-type comparison for the same player. That is, regardless of the *cardinal* interpretation of (1), the social planner may want to subsidize one player against another, or rank one type of a player higher than another type of the same player. Without even assuming the existence of such ranking rules, we shall find out the common feature of all interim Pareto optimal mechanisms.

3 The Solution

The result, roughly speaking, is that any interim Pareto optimum, no matter how it ranks across players and across types, is necessarily in the form of auctions. To state the result, we need only to formalize the notation for some standard concepts in auction design.

3.1 Notation

Ironing For any integrable function $\psi_i : T_i \to \mathbb{R}$, define $H_i(\psi_i) : [0, 1] \to \mathbb{R}$ by

$$\left(H_i(\psi_i)\right)(s) := \int_0^s \psi_i\left(F_i^{-1}(r)\right) dr \tag{7}$$

for any $s \in [0,1]$. Denote $\widehat{H}_i(\psi_i)$ for the convex hull of $H_i(\psi_i)$. Then $\widehat{H}_i(\psi_i)$ is differentiable almost everywhere on $[0,1]$. Define the *ironed copy* $\overline{\psi}_i$ of ψ_i by

$$\overline{\psi}_i(t_i) := \frac{d}{ds}\left(\widehat{H}_i(\psi_i)\right)(s)\Big|_{s=F_i(t_i)} \tag{8}$$

for any $t_i \in T_i$ such that $\widehat{H}_i(\psi_i)$ is differentiable at $F_i(t_i)$. Thus $\overline{\psi}_i$ is weakly increasing on the set of such t_i's. Extend $\overline{\psi}_i$ to other points in T_i to maintain its monotonicity.

The ϕ^+ and ϕ^- of function ϕ For any function $\phi : \mathbb{R} \to \mathbb{R}$, denote ϕ^+ and ϕ^- by $\phi^+(x) := \max\{\phi(x), 0\}$ and $\phi^-(x) := \max\{-\phi(x), 0\}$ for all $x \in \mathbb{R}$. Thus $\phi = \phi^+ - \phi^-$.

Allocations by Ranks For any profile $(\varphi_i)_{i=1}^n$ of integrable functions $\varphi_i : T_i \to \mathbb{R}$ $(\forall i)$ and any $(t_i)_{i=1}^n \in \prod_i T_i$, denote

$$\mathscr{A}\left((\varphi_i)_{i=1}^n\right) := \arg\max_{(\pi_i)_{i=1}^n \in \mathscr{S}} \int_{\prod_i T_i} \sum_{i=1}^n \varphi_i(t_i)\pi_i\left((t_k)_{k=1}^n\right) dF_1(t_1)\cdots dF_n(t_n), \tag{9}$$

where $\mathscr{S}$ denotes the set of all profiles $(\pi_i)_{i=1}^n$ of functions $\pi_i : \prod_k T_k \to [0,1]$ $(\forall i)$ such that $\sum_i \pi_i \leq 1$ on $\prod_k T_k$. Clearly, $(\pi_i)_{i=1}^n \in \mathscr{A}\left((\varphi_i)_{i=1}^n\right)$ if and only if, for almost every $(t_k)_{k=1}^n \in \prod_k T_k$ and for any i, $\varphi_i(t_i) > \max_{j\neq i}\varphi_j^+(t_j)$ implies $\pi_i\left((t_k)_{k=1}^n\right) = 1$, and $\varphi_i(t_i) < \max_{j\neq i}\varphi_j^+(t_j)$ implies $\pi_i\left((t_k)_{k=1}^n\right) = 0$. Note that $\mathscr{A}\left((\varphi_i)_{i=1}^n\right)$ contains an element $(\pi_i)_{i=1}^n$ such that $\pi_i(t_i,\cdot) = 0$ on T_{-i} whenever $\varphi_i(t_i) \leq 0$, as the previous sentence, combined with equal-probability random assignment whenever $\varphi_i(t_i) = \max_{j\neq i}\varphi_j^+(t_j) > 0$ and $\pi_i(t_i,\cdot) = 0$ on T_{-i} whenever $\varphi_i(t_i) = 0$, defines an element of this set.

3.2 The Result: Common Features of All Interim Pareto Optima

By *distribution* on an interval $[a,b]$, we mean a real function on $\mathbb{R}$ that is weakly increasing on $\mathbb{R}$, right-continuous on (a,b), constant on (b,∞), and equal to zero on $(-\infty,a)$. Our first theorem, proved in Section 4.1, reduces interim Pareto optimality, an objective with infinite dimensions, to a mechanism design problem with a unidimensional objective.

Theorem 1 *For any interim Pareto optimal mechanism $(Q_i,P_i)_{i=1}^n$, there exists a profile $(\lambda_i)_{i=1}^n$ such that λ_i is a distribution on T_i for each i, $\lambda_i > 0$ on some positive-measure subset of T_i for some i, and $(Q_i,P_i)_{i=1}^n$ maximizes*

$$\sum_i \int_{T_i} U_i(\cdot \mid \tilde{Q},\tilde{P})d\lambda_i \tag{10}$$

among all IC, IR and BB mechanisms $(\tilde{Q}, \tilde{P})$.

The distribution profile $(\lambda_i)_{i=1}^n$ in Theorem 1 can be interpreted as the *welfare ranking*, across players and across types, that supports the given Pareto optimum as a maximum of the unidimensional social welfare (10) among all IC, IR and BB mechanisms $(\tilde{Q}, \tilde{P})$. Note that λ_i need not be absolutely continuous with respect to the prior distribution F_i, hence it need not have a derivative λ_i' for which the integral in (10) equals an inner product $\int_{T_i} U_i(\cdot \mid \tilde{Q}, \tilde{P})\lambda_i' dF_i$. Our characterization of optimal mechanisms does not need such inner product representation of λ_i, hence we make no assumption to force its absolute continuity. The next theorem, proved in Sections 4.2–4.3, characterizes all constrained optima of (10).

Theorem 2 *For any profile* $(\lambda_i)_{i=1}^n$ *of distributions* λ_i *on* T_i*, if* $(Q_i, P_i)_{i=1}^n$ *maximizes (10) among all IC, IR and BB mechanisms, then there exists a profile* $(Z_{i,+}, Z_{i,-})_{i=1}^n$ *of integrable functions* $Z_{i,+}, Z_{i,-} : T_i \to \mathbb{R}$ *such that* $Z_{i,+} \leq Z_{i,-}$ *and:*

a. for each i*,* $Q_i^+ = \int_{T_i} q_{iA}(\cdot, t_{-i}) dF_{-i}$ *and* $Q_i^- = c\int_{T_i} q_{iB}(\cdot, t_{-i}) dF_{-i}$ *on* T_i *for some* $(q_{iA})_{i=1}^n \in \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$ *and* $(q_{iB})_{i=1}^n \in \mathscr{A}\left((-\overline{Z}_{i,-})_{i=1}^n\right)$*;*

b. $(P_i)_{i=1}^n$ *is determined by* $(Q_i)_{i=1}^n$ *according to:*

i. Eq. (5) for any i *and any* $t_i, t_i^0 \in T_i$*;*

ii. if $\int_{T_i}(1/s)d\lambda_i(s) < \int_{T_k}(1/s)d\lambda_k(s)$ *for some* $k \neq i$*, then* $\min_{t_i \in T_i} U_i(t_i \mid Q, P) = 0$*;*

iii. $\sum_{i=1}^n \int_{T_i} P_i(t_i) dF_i(t_i) = 0$.

Theorems 1 and 2 combined, any interim Pareto optimal mechanism is necessarily in form of auctions, encapsulated in a profile $(Z_{i,+}, Z_{i,-})_{i=1}^n$ of functions. These functions are jointly determined by the welfare ranking $(\lambda_i)_{i=1}^n$ supporting the Pareto optimum and the Lagrange multiplier for the constraint combining IR, IC and BB. Here $Z_{i,+}$ corresponds to the virtual surplus to the society contributed by player i who acts as a buyer of the good, and $Z_{i,-}$ the virtual surplus contributed by i who acts as a seller of the service of receiving the bad (Figure 1). A crucial feature, asserted by Theorem 2, is that $Z_{i,-}$ is above $Z_{i,+}$.

From $(Z_{i,+}, Z_{i,-})_{i=1}^n$, one finds out how the Pareto optimal mechanism allocates the two items according to Claim (a) of Theorem 2. First, obtain the ironed copy $\overline{Z}_{i,+}$ of $Z_{i,+}$ and the ironed copy $\overline{Z}_{i,-}$ of $Z_{i,-}$, for each i (Figure 2). Second, assign the good (item A) by the rank of $\left(\overline{Z}_{i,+}\right)_{i=1}^n$ à la Myerson [11]: Score each player i's alleged type t_i according

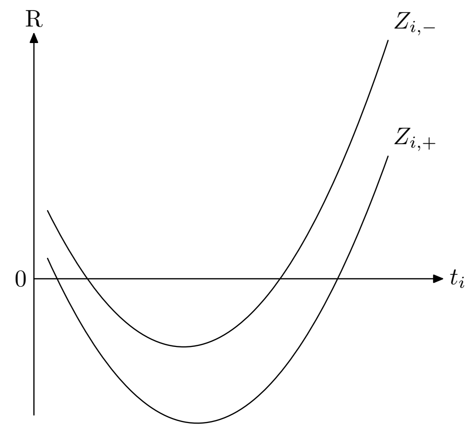

Figure 1: The bifurcated Z-values

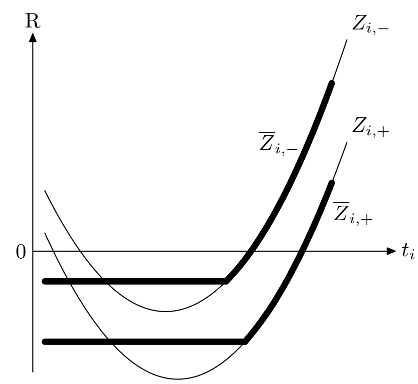

Figure 2: The thick curves: The ironed Z-values

to the ironed function $\overline{Z}_{i,+}$, and assign the good to a player with the highest ironed Z score provided that it is nonnegative. Likewise, assign the bad (item B) by the rank of $\left(-\overline{Z}_{i,-}\right)_{i=1}^{n}$: Score each player i's alleged type t_i according to the ironed function $\overline{Z}_{i,-}$, and assign the bad to a player with the lowest ironed Z score provided that it is nonpositive.

These two assignments together generate the reduced form allocation $(Q_i)_{i=1}^n$ of the mechanism. Then Claim (b.i) of the theorem says that from $(Q_i)_{i=1}^n$ one can pin down the money transfer rule P_i for each player i via the envelope equation up to a constant and pin down the constant according to (b.ii) and (b.iii). There, $\int_{T_i}(1/s)d\lambda_i(s)$ stands for player i's average weight in the social welfare that incorporates both the welfare ranking λ_i on the various types of his and his marginal valuations of money given these types. Claim (b.ii) says that anyone whose average weight is less than someone else's has zero as his minimum surplus in the mechanism. Claim (b.iii) states the obvious fact that the auctioneer retains no money surplus at any Pareto optimum.

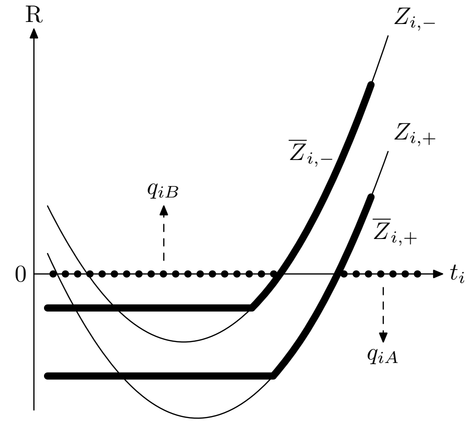

Figure 3: The optimal allocation

Claim (a) of the theorem implies a surprising property of an interim Pareto optimal mechanism (Q, P). The claim says that the positive part Q_i^+ is the marginal of q_{iA}, and the negative part Q_i^- the marginal of q_{iB}. Thus one can show that $q_{iA}(t)q_{iB}(t) = 0$ for almost every profile t of types across players. That is, a player-type that has a positive probability to get the good has zero probability to get the bad, and vice versa: In Figure 3, since $Z_{i,-}$ is above $Z_{i,+}$, any type with a nonngegative ironed $\overline{Z}_{i,+}$ score is larger than any type with a negative $\overline{Z}_{i,-}$ score.

Ensuring that the characterization in Theorem 2 is not vacuous, the next theorem asserts existence of Pareto optima. Specifically, it asserts existence of the interim Pareto optimum that maximizes the utilitarian social welfare given any welfare ranking across players and across types. The proof is deferred to Appendix C.

Theorem 3 *For any profile $(\lambda_i)_{i=i}^n$ of distributions λ_i on T_i there exists a mechanism that maximizes (10) among all IC, IR and BB mechanisms $(\tilde{Q}, \tilde{P})$.*

3.3 Implications

For any interim Pareto optimal mechanism supported by a profile $(\lambda_i)_{i=1}^n$ of distributions, let

$$\Lambda_i(t_i) := \int_{a_i}^{t_i} \frac{1}{s} d\lambda_i(s), \tag{11}$$

$$\alpha_\lambda := \max_{i=1,\ldots,n} \Lambda_i(b_i). \tag{12}$$

3.3.1 Redistribution across Players: Lump Sum Transfers

According to Claim (b.iii) of Theorem 2, any interim Pareto optimal mechanism rebates all the surplus thereby generated back to the players. Subtract from the transfer received by a player the amount necessitated by the envelope equation (Claim (b.i)) and we obtain the lump sum transfer the player receives. According to Claim (b.ii), the lump sum transfer goes only to those players i whose $\Lambda_i(b_i)$'s are equal to α_λ (defined in (11) and (12)). See the proof of Lemma 4 for the reasoning and explicit formulas for the transfer. Importantly, note from (11) that the direction of lump sum transfers depends purely on the welfare ranking $(\lambda_i)_{i=1}^n$ associated with the Pareto optimal mechanism, and not tied to prior distributions $(F_i)_{i=1}^n$ of types. Thus, an interim Pareto optimal mechanism need not take money from the (ex ante) stochastically high-type players ("the rich") to subsidize the stochastically low-type ones ("the poor"). Only when the endogenous welfare ranking $(\lambda_i)_{i=1}^n$ is identical to the exogenous type-distribution profile $(F_i)_{i=1}^n$ would the optimal mechanism necessarily direct lump sum transfers from the stochastically rich to the stochastically poor.

3.3.2 Redistribution across Types: Non-monotone Equilibrium Surplus

Consider, within this subsection, the symmetric case where types are drawn from the same distribution and the symmetric welfare ranking $(\lambda_i)_{i=1}^n$ such that $\lambda_i = \lambda_j$ for any players i and j. Then the only inequality among players is their realized types. In most mechanism design models, a player's role, whether to be a buyer or to be a seller, is assumed a priori, hence incentive compatibility implies that his equilibrium surplus from any mechanism is necessarily monotone (weakly increasing or weakly decreasing). Thus, the player's possible types can be reordered so that the mechanism gives higher types higher expected payoffs. That is, inequality in types begets inequality in payoffs.

In our model, by contrast, a player's role is endogenous, and incentive compatibility implies that he acts as a seller when his realized type is sufficiently low, and a buyer when his realized type is sufficiently high. Then an envelope-theorem argument gives the next corollary, saying that a player's interim expected payoff function in any Pareto optimal mechanism is like the roughly U-shape curve in Figure 4. With equilibrium surplus non-monotone in types, it is impossible to reorder the types so that a Pareto optimal mechanism would let inequality in types propagate to inequality in payoffs.

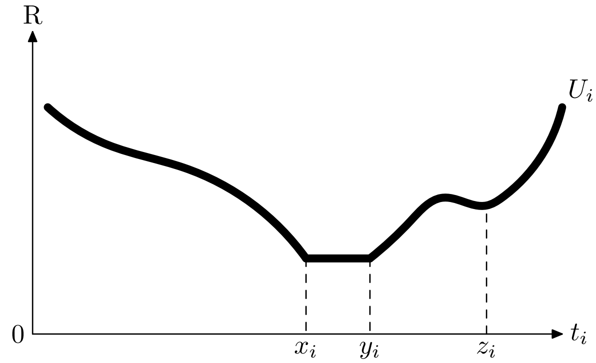

Figure 4: Non-monotone Interim Payoff Functions

Corollary 1 *Suppose: (i) $F_i = F_j$ and $\lambda_i = \lambda_j$ for all i and j, (ii) (Q, P) maximizes (10) subject to IC, IR and BB, and (iii) the ex ante probability with which Q assigns the bad is strictly positive. Then for each i there exists (x_i, y_i, z_i) such that $a_i < x_i < y_i < z_i < b_i$ and $U_i(\cdot \mid Q, P)$ is strictly decreasing on $[a_i, x_i)$, constant on (x_i, y_i), and strictly increasing on $(z_i, b_i]$.*

Proved in Appendix D.1, this corollary demonstrates the crucial role of having not only a good but also a bad to assign. Had there been only one item, be it good or bad, any IC allocation Q_i would be either nonnegative on T_i or nonpositive on T_i. In either case any player's equilibrium surplus would be monotone.

3.3.3 Exclusive Assignments

While we do not impose the condition that the good and the bad be assigned to different players, any interim Pareto optimal mechanism also satisfies that condition. More precisely, a mechanism satisfies *assignment exclusivity* iff, for the underlying ex post allocation rule $(q_{iA}, q_{iB})_{i=1}^n$, $q_i^A(t)q_i^B(t) = 0$ for almost every $t \in \prod_k T_k$ and all i.

Corollary 2 *If $c > 0$ then any interim Pareto optimal mechanism is also an interim Pareto optimal mechanism subject to not only IC, IR and BB but also assignment exclusivity.*

Proved in Appendix D.2, this corollary is a consequence of Theorem 2.a, which says that the positive part Q_i^+ of an optimal allocation Q_i is the marginal of q_{iA}, and the negative part Q_i^- the marginal of q_{iB}.

3.3.4 Prevalence of Ironing

In the literature, one can guarantee monotonicity of the virtual surplus functions, thereby avoiding ironing, by imposing regular conditions on the prior distributions F_i. By contrast, our counterpart of the virtual surplus, the $(Z_{i,+}, Z_{i,-})_{i=1}^n$ illustrated in Figure 1, cannot be guaranteed monotone even with such restrictions of the priors. That is because these Z functions depend not only on the priors but also on two endogenous objects: the welfare ranking $(\lambda_i)_{i=1}^n$ associated with the Pareto optimum and the Lagrange multiplier of the joint constraint of IC, IR and BB. For instance, when the Lagrange multiplier is zero, one can show that both a_i and b_i are local maxima of $Z_{i,-}$ on $[a_i, z_i]$. Even when the welfare ranking is well-behaved, the next corollary, proved in Appendix D.3, says that ironing is still prevalent.

Corollary 3 *If f_i is continuously differentiable at a_i for each i, then in any interim Pareto optimum $(Q_i, P_i)_{i=1}^n$ for which the supporting welfare ranking $(\lambda_i)_{i=1}^n$ is absolutely continuous in $(F_i)_{i=1}^n$ and, with λ_i' the Radon-Nikodym derivative of λ_i with respect to F_i ($\forall i$), satisfies*

$$\forall i \in \{1, \ldots, n\} : \liminf_{t_i \downarrow a_i} \lambda_i'(t_i) > 2\alpha_\lambda a_i, \tag{13}$$

there exists a player i for whom Q_i is constant on a neighborhood of a_i.

The absolute continuity assumption in Corollary 3 means that the welfare-ranking distributions λ_i's are nonsingular. In other words, the welfare ranking does not weigh any single type more than a continuum of types (unlike the Dirac measure), nor does it ignore almost all types by assigning zero density to each of them (unlike the Cantor measure). Thus, the assumption allows the simple case, often assumed in mechanism design, where the Radon-Nikodym derivative of λ_i with respect to F_i is constantly equal to one so that the social welfare function is a simple sum of players' surpluses.

A main goal in the mechanism design literature is efficient allocation, which in our model means allocating the good to the highest-type player and the bad to the lowest one. With ironing prevalent, however, efficient allocation in general does not belong to the Pareto frontier. Furthermore, as a direct consequence of Corollary 3, efficient allocation is suboptimal even in the simple case where each player weighs equally in the social welfare function.

As another implication of Corollary 3, if players' types are i.i.d. and the welfare ranking treats the players equally, the optimal mechanism entails ironing for all realized types belonging to a lower truncation of the type support. That is, the bad is allocated through an egalitarian lottery when the realized types eligible for the bad (i.e., below the cutoff x_i in Figure 4) belong to that lower truncation.[4]

4 The Method

The proof of Theorems 1 and 2 has three main steps. First is to quantify the infinite-dimensional design objective, interim Pareto optimality, into a one-dimension, utilitarian, social welfare function. The second step is to incorporate part of the IC constraint and optimality condition into the social welfare function thereby obtaining a tractable optimization problem. The third is to solve this optimization problem through bisecting the associated Lagrange problem into two linear programmings. The first and third steps are novel to the mechanism design literature. The second step, albeit stemming from the envelope theorem and integration-by-parts routines in the literature, develops a new operator—two-part operator—to calculate a player's endogenously countervailing information rents. Properties of this operator are important to the third step.

4.1 Quantifying the Objective: Proof of Theorem 1

This step uses the Hahn-Banach theorem to obtain a welfare ranking $(\lambda_i)_{i=1}^n$ that supports a Pareto optimal mechanism under consideration as a maximum of the unidimensional objective (10). Since a mechanism corresponds to functions defined on a continuum, the choice of the function space requires care. On one hand, the space needs to satisfy the nonempty interior condition for existence of a linear functional on the space that supports the Pareto optimum. On the other hand, the space needs to guarantee that the linear functional be representable by a profile of distributions in the form $(\lambda_i)_{i=1}^n$. Our choice of the function space stems from an observation that the interim expected surplus generated by any IC mechanism is a continuous function of types. Hence our function space consists of continuous functions defined on compact intervals, conducive to the Hahn-Banach and representation theorems.

First, we specify the function space. For each $i \in \{1, \ldots, n\}$ denote $C(T_i)$ for the space of continuous real functions defined on the closed, bounded interval T_i, with the maximum

[4] Dworczak, Kominers and Akbarpour [3] also obtain such an implication. Interestingly, our sufficient condition (13) for the implication looks similar to their sufficient condition, despite differences in our models.

norm $\|\cdot\|_{\max}$. Let

$$\mathscr{C} := \prod_{i=1}^{n} C(T_i)$$

and endow $\mathscr{C}$ with the maximum norm such that $\|(\varphi_i)_{i=1}^n\|_{\max} := \max_i \|\varphi_i\|_{\max}$ for all $(\varphi_i)_{i=1}^n \in \mathscr{C}$. Thus, $\mathscr{C}$ is a normed linear space. Define the utility possibility set

$$\mathbb{U} := \{(W_i)_{i=1}^n \in \mathscr{C} : \exists \text{ IC, IR \& BB } (Q,P)\, [\forall i\, \forall t_i \in T_i\, [W_i(t_i) \leq U_i(t_i \mid Q,P))]]\}\,. \tag{14}$$

As noted above, $(U_i(\cdot \mid Q,P))_{i=1}^n \in \mathscr{C}$ for any IC mechanism (Q,P).

To prove Theorem 1, pick any interim Pareto optimal mechanism (Q^*,P^*). Denote $u_i^* := U_i(\cdot \mid Q^*,P^*)$ for each i. Then $(u_i^*)_{i=1}^n \in \mathbb{U}$. Denote

$$\mathbb{V}((u_i^*)_{i=1}^n) := \{(u_i)_{i=1}^n \in \mathscr{C} \; : \; \forall i\, [u_i \geq u_i^* \text{ a.e. } T_i]\,; \exists i\, [u_i > u_i^* \text{ on a positive-measure } S_i \subseteq T_i]\}\,.$$

Obviously, $\mathbb{V}((u_i^*)_{i=1}^n)$ is convex.

Claim 1 *There exists a continuous linear functional ϕ on $\mathscr{C}$, not identically zero, such that for all $(u_i)_{i=1}^n \in \mathbb{U}$,*

$$\phi\left((u_i)_{i=1}^n\right) \leq \phi\left((u_i^*)_{i=1}^n\right). \tag{15}$$

Proof First, $\mathbb{U}$ is convex (Appendix A), and $\mathbb{V}((u_i^*)_{i=1}^n)$ convex as noted above. Second, $\mathbb{U}$ contains an interior point: Consider the mechanism that gives away the good A for free with probability $1/2$, else assigns neither item to anyone, and, in the former event, randomly assigns the good A (for free) to one of the n players with equal probability. This mechanism is IC, IR and BB, and it generates for everyone an interim expected payoff constantly equal to $1/(2n)$. Thus, this payoff profile belongs to $\mathbb{U}$. Now consider another mechanism that differs from the former only by that it assigns the good with probability $1/2+\epsilon$. The mechanism is also IC, IR and BB, and generates an expected payoff profile larger than the former in every dimension by ϵ/n. Since this is true for all $\epsilon \in (0,1/2]$, the payoff profile generated by the former mechanism is an interior point of $\mathbb{U}$ with respect to the max norm. Third, $\mathbb{V}\left((u_i^*)_{i=1}^n\right)$ contain no interior point of $\mathbb{U}$; otherwise, such an interior point, by definition of $\mathbb{V}\left((u_i^*)_{i=1}^n\right)$, interim Pareto dominates $(u_i^*)_{i=1}^n$, contradiction. Thus, by the Hahn-Banach theorem, there exists a continuous linear functional ϕ on $\mathscr{C}$, not identically zero, such that, for some constant w, for any $(u_i)_{i=1}^n \in \mathbb{U}$ and any $(\hat{u}_i)_{i=1}^n \in \mathbb{V}((u_i^*)_{i=1}^n)$,

$$\phi\left((u_i)_{i=1}^n\right) \leq w \leq \phi\left((\hat{u}_i)_{i=1}^n\right). \tag{16}$$

For any $\epsilon > 0$, the profile $(u_i^* + \epsilon)_{i=1}^n \in \mathbb{V}((u_i^*)_{i=1}^n)$. Thus

$$w \leq \phi\left((u_i^* + \epsilon)_{i=1}^n\right) = \phi\left((u_i^*)_{i=1}^n\right) + \epsilon\phi(\mathbf{1}),$$

with the equality due to linearity of ϕ, and $\mathbf{1}$ denoting the unit vector of $\mathscr{C}$. Since continuous linear functionals are bounded, $\epsilon\phi(\mathbf{1}) \to 0$ as $\epsilon \to 0$. Hence $w \leq \phi\left((u_i^*)_{i=1}^n\right)$. This coupled with the fact $(u_i^*)_{i=1}^n \in \mathbb{U}$ implies $\phi\left((u_i^*)_{i=1}^n\right) \leq w \leq \phi\left((u_i^*)_{i=1}^n\right)$, hence $\phi\left((u_i^*)_{i=1}^n\right) = w$. Plug this into (16) to obtain (15) and hence the claim. $\square$

For each $i \in \{1, \ldots, n\}$ and any $u_i \in C(T_i)$ let

$$\phi_i(u_i) := \phi\left(0, \ldots, 0, u_i, 0, \ldots, 0\right),$$

that is, the action of ϕ on the profile of payoff functions whose components are constantly zero except the one corresponding to player i's payoff function. By linearity of ϕ,

$$\phi\left((u_i)_{i=1}^n\right) = \sum_{i=1}^n \phi_i(u_i) \tag{17}$$

for all $(u_i)_{i=1}^n \in \mathscr{C}$. Obviously, for each i, ϕ_i is a continuous linear functional on $C(T_i)$. Thus ϕ_i is also a bounded functional on $C(T_i)$.

Claim 2 *For each $i \in \{1, \ldots, n\}$, ϕ_i is positive.*[5]

Proof Suppose, to the contrary, that $\phi_i(u_i) < 0$ for some $u_i \in C(T_i)$ such that $u_i \geq 0$ on T_i. Then $\left(u_i^* - u_i, (u_j^*)_{j \neq i}\right) \in \mathbb{U}$ by definition of $\mathbb{U}$, hence Claim 1 implies

$$\phi\left((u_j^*)_{j=1}^n\right) \geq \phi\left(\left(u_i^* - u_i, (u_j^*)_{j\neq i}\right)\right) = \sum_{j=1}^n \phi_j(u_j^*) - \phi_i(u_i) > \sum_{j=1}^n \phi_j(u_j^*) = \phi\left((u_j^*)_{j=1}^n\right),$$

contradiction. $\square$

For any i, since ϕ_i is a bounded linear functional on $C(T_i)$, with $T_i = [a_i, b_i]$ a closed, bounded interval, the Riesz representation theorem in its original version (Royden and Fitzpatrick [13, p468]) implies that there exists a unique function $\phi_i : T_i \to \mathbb{R}$, of bounded variation on $[a_i, b_i]$, continuous on the right on (a_i, b_i), and vanishing at a_i, such that

$$\phi_i(u_i) = \int_{T_i} u_i d\lambda_i$$

[5] A functional ϕ_i on $C(T_i)$ is *positive* iff $\phi_i(u_i) \geq 0$ for any $u_i \in C(T_i)$ such that $u_i \geq 0$ on T_i.

for all $u_i \in C(T_i)$. This, combined with (15) and (17), delivers Theorem 1 if (i) λ_i is also weakly increasing, (ii) its range belongs to $\mathbb{R}_+$, and (iii) $\lambda_i > 0$ on some positive-measure subset of T_i for some i.

Property (ii) follows from property (i) because $\lambda_i(a_i) = 0$. Then (iii) follows from (ii): Otherwise (ii) implies $\lambda_i = 0$ for all i, hence ϕ is identically zero on $\mathscr{C}$, contradiction to Claim 1. Thus it suffices to prove (i).

To that end, suppose, to the contrary, that $t_i < t_i'$ and $\lambda_i(t_i) > \lambda_i(t_i')$ for some $t_i, t_i' \in (a_i, b_i)$.[6] Then, since λ_i is right-continuous on (a_i, b_i), there exists a sufficiently small $\epsilon > 0$ such that for any $\delta \in (0, \epsilon)$, $\lambda_i(t_i+\delta) > \lambda_i(t_i'+\delta)$. It is easy to construct a continuous function $u_i : T_i \to [0, 1]$ whose support is contained by $[t_i, t_i' + \epsilon]$ such that $u_i = 1$ on $[t_i + \epsilon/2, t_i' + \epsilon/2]$. Then $u_i \geq 0$ on T_i, $u_i \in C(T_i)$, and yet

$$\phi_i(u_i) = \int_{T_i} u_i d\lambda_i = \int_{t_i}^{t_i'+\epsilon} d\lambda_i \leq \int_{t_i+\epsilon/2}^{t_i'+\epsilon/2} d\lambda_i = \lambda_i(t_i' + \epsilon/2) - \lambda_i(t_i + \epsilon) < 0,$$

contradicting Claim 2. That proves property (i) of λ_i. Thus Theorem 1 follows.

4.2 Calculating the Objective with Two-Part Operators

Theorem 1, coupled with (3), implies that any interim Pareto optimum is a maximum of

$$\sum_i \int_{T_i} Q_i(t_i) d\lambda_i(t_i) - \sum_i \int_{T_i} \frac{P_i(t_i)}{t_i} d\lambda_i(t_i) \tag{18}$$

among all mechanisms $(Q_i, P_i)_{i=1}^n$ subject to IC, IR and BB, given some profile $(\lambda_i)_{i=1}^n$ of distributions λ_i on T_i specified in Theorem 1. By (11), $t_i \mapsto 1/t_i$ is the Radon-Nikodym derivative $\frac{d\Lambda_i}{d\lambda_i}$ of Λ_i with respect to λ_i, hence the objective (18) is equal to

$$\sum_i \int_{T_i} Q_i d\lambda_i - \sum_i \int_{T_i} P_i d\Lambda_i. \tag{19}$$

One can calculate the integral $\int_{T_i} P_i d\Lambda_i$ in the second sum with the envelope theorem and integration-by-part routine in auction theory, thereby incorporating into (18) the first-order condition part of the IC constraint. The calculation amounts to modifying the ex ante expected revenue—measured by the endogenous Λ_i rather than the exogenous F_i—that one can extract from player i by the expected information rent necessary to incentivize i in

[6] There is no need to consider $t_i = a_i$ and $t_i' = b_i$ because we already have $\lambda_i(a_i) = 0$, and it is immaterial to change the value of λ_i at the singleton b_i.

implementing the allocation Q_i. Because i's incentive is that of a buyer when $Q_i(t_i) > 0$, and that of a seller (for the service of receiving the bad) when $Q_i(t_i) < 0$, the information rent calculation is an operation that bifurcates according to the sign of $Q_i(t_i)$ for each realized type t_i. Managing this operation tractably and understanding its properties are important for this and the next steps of the proof for Theorem 2. First we define such operations.

4.2.1 Two-Part Operators

For any player i and any three integrable functions $Q_i, \varphi_{i,+}, \varphi_{i,-} : T_i \to \mathbb{R}$, denote $\varphi_i := (\varphi_{i,+}, \varphi_{i,-})$, called *two-part function*, and define

$$\langle Q_i : \varphi_i| := \int_{T_i} Q_i^+(s)\varphi_{i,+}(s)ds - \int_{T_i} Q_i^-(s)\varphi_{i,-}(s)ds. \tag{20}$$

Thus the operation $Q_i \mapsto \langle Q_i : \varphi_i|$ acts on the function Q_i in two parts, one on the positive part Q_i^+ of Q_i, the other on the negative part, $-Q_i^-$. The asymmetric bracket of Q_i and φ_i is to highlight the asymmetry between the two arguments: Obviously, $\langle Q_i : \varphi_i|$ is not linear in Q_i unless $\varphi_{i,+} = \varphi_{i,-}$; by contrast, $\langle Q_i : \varphi_i|$ is a linear functional of φ_i. Any integrable function $g_i : T_i \to \mathbb{R}$ is the same as a two-part function $(g_{i,+}, g_{i,-})$ such that $g_{i,+} = g_{i,-} = g_i$.

A two-part function $\varphi_i := (\varphi_{i,+}, \varphi_{i,-})$ on T_i is said *well-ordered* iff $\varphi_{i,+} \leq \varphi_{i,-}$ a.e. on T_i. Next is an important property of well-ordered two-part functions (proved in Appendix B.1).

Lemma 1 *For any i and any two-part function φ_i that is well-ordered, $Q_i \mapsto \langle Q_i : \varphi_i|$ is a concave functional on the space of integrable functions defined on T_i.*

For any i and any distribution μ_i on T_i, define a two-part function $\rho(\mu_i) := (\rho_+(\mu_i), \rho_-(\mu_i))$ by letting, for any $t_i \in T_i$,

$$\rho_+(\mu_i)(t_i) \;:=\; -\int_{T_i} d\mu_i + \int_{a_i}^{t_i} d\mu_i, \tag{21}$$

$$\rho_-(\mu_i)(t_i) \;:=\; \int_{a_i}^{t_i} d\mu_i. \tag{22}$$

Obviously $\rho(\mu_i)$ is well-ordered. It will be clear that $\rho_+(\mu_i)$ reflects i's information rent density when i acts as a buyer, and $\rho_-(\mu_i)$, i's information rent density when i acts as a seller, had i's type been measured by μ_i. The bifurcated calculation of a player i's information rents mentioned previously, according to the next lemma (proved in Appendix B.2), involves a two-part operation on his allocation Q_i with $\rho(\mu_i)$:

Lemma 2 *For any IC mechanism (Q, P), any player i and any distribution μ_i on T_i, with the notation in (4) and (6),*

$$\int_{T_i} P_i d\mu_i = \int_{T_i} t_i Q_i(t_i) d\mu_i(t_i) + \langle Q_i : \rho(\mu_i)| - \tilde{U}_i(\tau(Q_i)) \int_{T_i} d\mu_i. \tag{23}$$

4.2.2 The Budget Balance Condition Combined with IC and IR

Denote $\mathbb{I}$ for the identity map $s \mapsto s$ on $\mathbb{R}$. For any functions $g, h : \mathbb{R} \to \mathbb{R}$, denote gh for the pointwise product between g and h, so that $(gh)(s) = g(s)h(s)$ for all $s \in \mathbb{R}$.

Lemma 3 *For any allocation $(Q_i)_{i=1}^n$ such that Q_i is weakly increasing on T_i for any i, if $(Q_i)_{i=1}^n$ constitutes an IC, IR and BB mechanism then*

$$\sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| \geq 0; \tag{24}$$

conversely, if (24) holds then there exists an ex post payment rule $(p_i)_{i=1}^n$,

$$\sum_i p_i(t) = \sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| \tag{25}$$

for any $t \in \prod_j T_j$, which coupled with Q constitutes an IC, IR and BB mechanism such that IR is binding at some $t_i \in T_i$ for every i.

Proof By Lemma 2, IC of (Q, P) implies (23). Plug $\mu_i = F_i$ into (23) and note $dF_i(s) = f_i(s)ds$ to obtain

$$\begin{aligned} \int_{T_i} P_i dF_i &= \int_{T_i} t_i Q_i(t_i) f_i(t_i) dt_i + \langle Q_i : \rho(F_i)| - \tilde{U}_i(\tau(Q_i)) \\ &= \langle Q_i : \mathbb{I} f_i| + \langle Q_i : \rho(F_i)| - \tilde{U}_i(\tau(Q_i)), \end{aligned}$$

where the second line comes from the notations of $\mathbb{I}$ and pointwise product $\mathbb{I} f_i$, and the fact that $\int_{T_i} Q_i(s)\psi(s)ds = \langle Q_i : \psi|$ for any integrable function $\psi : T_i \to \mathbb{R}$, as ψ is a special two-part function such that $\psi_+ = \psi_- = \psi$. Then, since $\varphi \mapsto \langle Q_i : \varphi|$ is linear,

$$\int_{T_i} P_i dF_i = \langle Q_i : \mathbb{I} f_i + \rho(F_i)| - \tilde{U}_i(\tau(Q_i)).$$

Sum this equation across $i = 1, \ldots, n$ to get

$$\sum_i \int_{T_i} P_i dF_i = \sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| - \sum_i \tilde{U}_i(\tau(Q_i)).$$

BB implies that the left-hand side is nonnegative and hence

$$\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq \sum_i \tilde{U}_i\left(\tau(Q_i)\right), \tag{26}$$

which coupled with IR (Section 2.3) implies (24). Thus (24) is a necessary condition for any profile Q of weakly increasing allocations to constitute an IC, IR and BB mechanism. The proof of the converse is routine and hence relegated to Appendix B.3. ■

4.2.3 The Objective with Optimal Payment Rules

Now we calculate the objective (18) by incorporating (23) (part of the IC constraint) and optimality of the payment rule. Denote $\mathscr{Q}$ for the set of all (reduced-form) allocations $(Q_i)_{i=1}^n$, each generated by some ex post allocation according to (2). Let $\mathscr{Q}_{\text{mon}}$ be the set of all $(Q_i)_{i=1}^n \in \mathscr{Q}$ such that Q_i is weakly increasing for every i.

Lemma 4 *For any profile $\lambda := (\lambda_i)_{i=1}^n$ of distributions specified in Theorem 1, denote Λ and α_λ by (11)–(12); then maximization of (10) subject to IC, IR and BB is equivalent to*

$$\begin{array}{rl} \max\limits_{Q\in\mathscr{Q}_{\text{mon}}} & \sum_i \langle Q_i : \alpha_\lambda\left(\mathbb{I}f_i + \rho(F_i)\right) - \rho(\Lambda_i)| \\ s.t. & \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0. \end{array} \tag{27}$$

Proof By Sections 2.3 and Lemma 3, the constraints $Q \in \mathscr{Q}_{\text{mon}}$ and $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$ together constitute the choice set for the problem. We still need to show that the objective in (27) is equal to (10), i.e., equal to (19). By Lemma 2, IC of (Q, P) implies (23). Plug into (23) the case $\mu_i = \Lambda_i$ and note $d\Lambda_i(s) = (1/s)d\lambda_i$ by (11) to obtain

$$\begin{aligned} \int_{T_i} P_i d\Lambda_i &= \int_{T_i} sQ_i(s)(1/s)d\lambda_i(s) + \langle Q_i : \rho(\Lambda_i)| - \tilde{U}_i\left(\tau(Q_i)\right)\int_{T_i} d\Lambda_i \\ &= \int_{T_i} Q_i d\lambda_i + \langle Q_i : \rho(\Lambda_i)| - \tilde{U}_i\left(\tau(Q_i)\right)\int_{T_i} d\Lambda_i. \end{aligned}$$

Sum this across i and plug the equation obtained thereof into (19) to see that the objective (18) is equal to

$$\sum_i \tilde{U}_i\left(\tau(Q_i)\right)\int_{T_i} d\Lambda_i - \sum_i \langle Q_i : \rho(\Lambda_i)|. \tag{28}$$

By (12) and (26),

$$\sum_i \tilde{U}_i\left(\tau(Q_i)\right)\int_{T_i} d\Lambda_i \leq \alpha_\lambda \sum_i \tilde{U}_i\left(\tau(Q_i)\right) \leq \alpha_\lambda \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)|.$$

Furthermore, the right end of this inequality can be attained: Pick a player i_* for whom $\int_{T_{i_*}} d\Lambda_{i_*} = \alpha_\lambda$; for any realized type profile $t \in \prod_i T_i$ and any $i \neq i_*$, set the money transfer $p_i^*(t)$ from i to others to be $p_i(t)$, with p_i being the ex post payment rule in (25); set the money transfer $p_{i_*}^*(t)$ from i_* to others as $p_{i_*}(t) - \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)|$. Given $(p_i^*)_{i=1}^n$, BB follows from (25), and $\tilde{U}_i(\tau(Q_i)) = 0$ for all $i \neq i_*$, while $\tilde{U}_i(\tau(Q_{i_*})) = \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)|$. Thus, given the allocation Q, when the payment is optimized, we have

$$\sum_i \tilde{U}_i(\tau(Q_i)) \int_{T_i} d\Lambda_i = \alpha_\lambda \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \, .$$

Hence there is no loss of generality to assume that (10), or (28), is equal to

$$\alpha_\lambda \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| - \sum_i \langle Q_i : \rho(\Lambda_i)| \, .$$

This, by linearity of $\varphi_i \mapsto \langle Q_i : \varphi_i|$, is equal to

$$\sum_i \left(\langle Q_i : \alpha_\lambda (\mathbb{I}f_i + \rho(F_i))| - \langle Q_i : \rho(\Lambda_i)|\right) = \sum_i \langle Q_i : \alpha_\lambda (\mathbb{I}f_i + \rho(F_i)) - \rho(\Lambda_i)| \, ,$$

the objective in (27). ■

Remark 2 By (26), any payment rule $\hat{p}$ that renders $\tilde{U}_i(\tau(Q_i)) > 0$ while $\int_{T_i} d\Lambda_i < \alpha_\lambda$ would make $\sum_i \tilde{U}_i(\tau(Q_i)) \int_{T_i} \Lambda_i < \alpha_\lambda \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)|$. Since the proof of Lemma 4 has shown that the right-hand side of this inequality is attainable, the payment rule $\hat{p}$ is suboptimal. This, coupled with the definitions of $\tau(Q_i)$ and $\tilde{U}_i$, implies Claim (b.ii) of Theorem 2. Claim (b.iii) is obvious. If it does not hold, there is a positive expected money surplus, which can be equally distributed to the players independently of their types thereby achieving Pareto improvement, contradiction.

4.3 Solving the Constrained Optimization Problem

To solve the optimization problem (27), first we reformulate it through the saddle point condition, which delivers the profile $(Z_{i,+}, Z_{i,-})_{i=1}^n$ of functions stated in Theorem 2. Then we maximize the associated Lagrangian thereby obtaining the formula for the optimal mechanism. The first step requires that the saddle point condition be a necessary condition for any solution of (27). The second step presents us a nonlinear programming problem, which we solve through bisecting it into two linear programmings. Properties of the two-part operators are instrumental to both steps.

4.3.1 Deriving the $(Z_i)_{i=1}^n$ Functions through the Saddle Point Condition

Recall that $\mathscr{Q}$ denotes the space of all allocations $(Q_i)_{i=1}^n$, each generated by some ex post allocation according to (2). It is easy to verify that $\mathscr{Q}$ belongs to a normed linear space. Endow $\mathscr{Q}$ with such a norm.[7] Also recall $\mathscr{Q}_{\text{mon}}$ as the set of $(Q_i)_{i=1}^n \in \mathscr{Q}$ such that Q_i is weakly increasing for any i. One can prove that $\mathscr{Q}_{\text{mon}}$ is convex (Appendix B.4).

Denote ν for the Lagrange multiplier of the constraint $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$ in (27). The Lagrangian associated with (27) is

$$\begin{aligned}
\mathscr{L}(Q,\nu) &:= \sum_i \langle Q_i : \alpha_\lambda \left(\mathbb{I}f_i + \rho(F_i)\right) - \rho(\Lambda_i)| + \nu \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \\
&= \sum_i \langle Q_i : \alpha_\lambda \left(\mathbb{I}f_i + \rho(F_i)\right) - \rho(\Lambda_i) + \nu \left(\mathbb{I}f_i + \rho(F_i)\right)| \\
&= \sum_i \left\langle Q_i : \left((\alpha_\lambda + \nu) \left(\mathbb{I} + \frac{\rho(F_i)}{f_i} \right) - \frac{\rho(\Lambda_i)}{f_i} \right) f_i \right|, \qquad (29)
\end{aligned}$$

with the second line due to linearity of $\varphi_i \mapsto \langle Q_i : \varphi_i|$, and the third line due to the fact that two-part functions constitute an algebra (allowing for multiplications and divisions).

Lemma 5 *Q^* is a solution for (27) if and only if there exists a $\nu^* \in \mathbb{R}_+$ such that (Q^*, ν^*) is a saddle point in the sense that, for all $Q \in \mathscr{Q}_{\text{mon}}$ and all $\nu \in \mathbb{R}_+$,*

$$\mathscr{L}(Q^*, \nu) \geq \mathscr{L}(Q^*, \nu^*) \geq \mathscr{L}(Q, \nu^*). \qquad (30)$$

Proof The "if" part is trivial. To prove the "only if" part, it suffices to verify the conditions corresponding to those in Luenberger [8, Corollary 1, p219]. To that end, we start with two claims about each player i, which are proved in Appendix B.5: First, the two-part functions $\mathbb{I}f_i + \rho(F_i)$ and $\alpha_\lambda \left(\mathbb{I}f_i + \rho(F_i)\right) - \rho(\Lambda_i)$ are each well-ordered. Second,

$$\langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq \tau(Q_i) \int_{T_i} Q_i dF_i \qquad (31)$$

for any weakly increasing $Q_i : T_i \to \mathbb{R}$.

Now that $\mathbb{I}f_i + \rho(F_i)$ and $\alpha_\lambda \left(\mathbb{I}f_i + \rho(F_i)\right) - \rho(\Lambda_i)$ are each well-ordered, Lemma 1 implies that, in (27), both the constraint expression $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)|$ and the objective

[7] For example, for each player i let $L^2(T_i)$ be the L^2-space of measurable real functions defined on T_i, endowed with the measure F_i. Clearly $\mathscr{Q} \in \prod_i L^2(T_i)$. Define the norm for $\prod_i L^2(T_i)$ by $\|Q\| := \sum_i \|Q_i\|_2$ for any $Q := (Q_i)_{i=1}^n \in \prod_i L^2(T_i)$.

$\sum_i \langle Q_i : \alpha_\lambda (\mathbb{I} f_i + \rho(F_i)) - \rho(\Lambda_i)|$ are concave functions of $(Q_i)_{i=1}^n$. Thus,

$$\left\{ (Q_i)_{i=1}^n \in \mathscr{Q} : \sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| \geq 0 \right\}$$

is a convex set, and the objective in (27) is concave in the choice variable. This, coupled with convexity of $\mathscr{Q}_{\text{mon}}$ (Appendix B.4), means that the proof is complete if there exists a $(Q_i)_{i=1}^n \in \mathscr{Q}_{\text{mon}}$ such that $\sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| > 0$. Such $(Q_i)_{i=1}^n$ exists: always assign the good to player 1 and never assign the bad at all. That is, $Q_1 = 1$, hence $\tau(Q_1) = a_1$, and $Q_i = 0$ for all $i \neq 1$. Note $(Q_i)_{i=1}^n \in \mathscr{Q}_{\text{mon}}$. By (31),

$$\sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| = \langle Q_1 : \mathbb{I} f_1 + \rho(F_1)| \geq a_1 \int_{T_1} Q_1 dF_1 = a_1 > 0.$$

Now that all conditions are verified, the saddle point characterization follows. ■

Coupled with Theorem 1, Lemma 5 implies that any Pareto optimal mechanism is necessarily a solution of $\max_{Q \in \mathscr{Q}_{\text{mon}}} \mathscr{L}(Q, \nu)$ with $\mathscr{L}(Q, \nu)$ defined by (29) for some profile $(\lambda_i)_{i=1}^n$ of distributions specified in Theorem 1, and some $\nu \in \mathbb{R}_+$. For each i, denote

$$Z_i := (\alpha_\lambda + \nu) \left(\mathbb{I} + \frac{\rho(F_i)}{f_i} \right) - \frac{\rho(\Lambda_i)}{f_i}. \tag{32}$$

Then (29) is the same as

$$\mathscr{L}(Q, \nu) = \sum_i \langle Q_i : Z_i f_i| \,. \tag{33}$$

The Z_i defined in (32) is exactly the two-part function that constitutes the profile $(Z_i)_{i=1}^n$ in Theorem 2. Plugging (21) and (22) into (32), and recalling $\mathbb{I}$ as the notation for the identity map, we obtain the explicit formula for the functions $Z_{i,+}$ and $Z_{i,-}$: for all i and all $t_i \in T_i$,

$$Z_{i,+}(t_i) = \alpha_\lambda t_i + \frac{\alpha_\lambda F_i(t_i) - \Lambda_i(t_i)}{f_i(t_i)} - \frac{\alpha_\lambda - \Lambda_i(b_i)}{f_i(t_i)} + \nu \left(t_i - \frac{1 - F_i(t_i)}{f_i(t_i)} \right), \tag{34}$$

$$Z_{i,-}(t_i) = \alpha_\lambda t_i + \frac{\alpha_\lambda F_i(t_i) - \Lambda_i(t_i)}{f_i(t_i)} + \nu \left(t_i + \frac{F_i(t_i)}{f_i(t_i)} \right). \tag{35}$$

By (34) and (35), $Z_{i,+}(t_i) - Z_{i,-}(t_i) = -\left(\alpha_\lambda - \Lambda_i(b_i) + \nu\right)/f_i(t_i)$, which is nonpositive because $\alpha_\lambda - \Lambda_i(b_i)$ and ν are each nonnegative by definition. Thus, $Z_{i,+} \leq Z_{i,-}$, as asserted in Theorem 2.

4.3.2 Maximizing the Lagrangian through Bisection

It follows that any interim Pareto optimal mechanism satisfies the saddle point condition (30), and hence maximizes the associated Lagrangian $\sum_i \langle Q_i : Z_i f_i|$, defined in (33).

To solve this Lagrange problem, recall from the end of the previous subsection that the two-part function Z_i is well-ordered. This property will be useful in the following, where we bisect the Lagrange problem into two independent linear programmings.

Let $\mathscr{Q}_+$ be the set of all $(Q_i)_{i=1}^n \in \mathscr{Q}$ such that $Q_i \geq 0$ for all i, and $\mathscr{Q}_-$ the set of all $(Q_i)_{i=1}^n \in \mathscr{Q}$ such that $Q_i \leq 0$ for all i. Obviously both $\mathscr{Q}_+$ and $\mathscr{Q}_-$ are convex. By (20) and (33), $\max_{Q\in\mathscr{Q}_{\mathrm{mon}}} \mathscr{L}(Q,\nu)$ is equivalent to

$$\max_{(Q_i)_{i=1}^n \in \mathscr{Q}_{\mathrm{mon}}} \left(\sum_i \int_{T_i} Q_i^+ Z_{i,+} dF_i - \sum_i \int_{T_i} Q_i^- Z_{i,-} dF_i \right) \tag{36}$$

$$\leq \max_{(Q_i)_{i=1}^n \in \mathscr{Q}_{\mathrm{mon}}} \sum_i \int_{T_i} Q_i^+ Z_{i,+} dF_i + \max_{(Q_i)_{i=1}^n \in \mathscr{Q}_{\mathrm{mon}}} \sum_i \int_{T_i} \left(-Q_i^-\right) Z_{i,-} dF_i \tag{37}$$

$$= \max_{(Q_i)_{i=1}^n \in \mathscr{Q}_{\mathrm{mon}} \cap \mathscr{Q}_+} \sum_i \int_{T_i} Q_i Z_{i,+} dF_i \tag{38}$$

$$+ \max_{(Q_i)_{i=1}^n \in \mathscr{Q}_{\mathrm{mon}} \cap \mathscr{Q}_-} \sum_i \int_{T_i} Q_i Z_{i,-} dF_i. \tag{39}$$

Thus, to solve (36), it suffices to first solve (38) and (39) individually and then construct from the two solutions a $Q^* \in \mathscr{Q}_{\mathrm{mon}}$ given which the objective in (36) attains the sum of the maximands in (38) and (39). The next two lemmas are proved in Appendix B.6.

Lemma 6 *$(\hat{Q}_i)_{i=1}^n \in \mathscr{Q}$ solves (38) if and only if, for some $(\hat{q}_i)_{i=1} \in \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$ and for each i, $\hat{Q}_i = \int_{T_{-i}} \hat{q}_i(\cdot, t_{-i}) dF_{-i}(t_{-i})$ on T_i.*

Lemma 7 *$(\check{Q}_i)_{i=1}^n \in \mathscr{Q}$ solves (39) if and only if, for some $(\check{q}_{iB})_{i=1} \in \mathscr{A}\left((-\overline{Z}_{i,-})_{i=1}^n\right)$ and for each i, $\check{Q}_i = c \int_{T_{-i}} \check{q}_i(\cdot, t_{-i}) dF_{-i}(t_{-i})$ on T_i.*[8]

As noted in Section 3.1, $\mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$ contains an element $(\hat{q}_{iA})_{i=1}$ such that $\hat{q}_{iA}(t_i,\cdot) = 0$ on T_{-i} whenever $\overline{Z}_{i,+}(t_i) \leq 0$. Likewise, $\mathscr{A}\left((-\overline{Z}_{i,-})_{i=1}^n\right)$ contains an element $(\check{q}_{iB})_{i=1}$ such that $\check{q}_{iB}(t_i,\cdot) = 0$ on T_{-i} whenever $\overline{Z}_{i,-}(t_i) \geq 0$. For each i, let $\hat{Q}_i$ be the marginal of $\hat{q}_{iA}$, and $\check{Q}_i$ the marginal of $c\check{q}_{iB}$. By Lemmas 6 and 7, $(\hat{Q}_i)_{i=1}^n$ solves (39), and $(\check{Q}_i)_{i=1}^n$ solves (39). Note that the support of $(\hat{Q}_i)_{i=1}^n$ and that of $(\check{Q}_i)_{i=1}^n$ have no overlapped interior: For each i,

$$\hat{Q}_i(t_i) \neq 0 \iff \hat{q}_i(t_i,\cdot) \not\equiv 0 \iff \check{q}_i(t_i,\cdot) \equiv 0 \iff \check{Q}_i(t_i) = 0. \tag{40}$$

That is because, by the choice of $(\hat{q}_i)_{i=1}^n$ and $(\check{q}_i)_{i=1}^n$,

$$\check{Q}_i(t_i) \neq 0 \;\Rightarrow\; \check{Q}_i(t_i) < 0 \;\Rightarrow\; t_i \leq \sup\left\{\tau_i \in T_i : \overline{Z}_{i,-}(\tau_i) < 0\right\},$$
$$\hat{Q}_i(t_i) \neq 0 \;\Rightarrow\; \hat{Q}_i(t_i) > 0 \;\Rightarrow\; t_i \geq \inf\left\{\tau_i \in T_i : \overline{Z}_{i,+}(\tau_i) > 0\right\},$$

[8]Recall (2) for the role of the coefficient c.

and, because Z_i is well-ordered, one can prove (Appendix B.7) that

$$\sup\left\{\tau_i \in T_i : \overline{Z}_{i,-}(\tau_i) < 0\right\} \leq \inf\left\{\tau_i \in T_i : \overline{Z}_{i,+}(\tau_i) \geq 0\right\}. \tag{41}$$

Thus, the following function Q_i^* is well-defined and weakly increasing on T_i:

$$Q_i^*(t_i) := \begin{cases} \check{Q}_i(t_i) & \text{if } \overline{Z}_{i,-}(t_i) < 0 \\ \hat{Q}_i(t_i) & \text{if } \overline{Z}_{i,+}(t_i) > 0 \\ 0 & \text{else.} \end{cases} \tag{42}$$

Because of (40), $(Q_i^*)^+ = \hat{Q}_i$ and $(Q_i^*)^- = -\check{Q}_i$ for any i. It follows that $(Q_i^*)_{i=1}^n$ is a solution for both problems in (37) simultaneously. By (41), (42) and monotonicity of $\hat{Q}_i$ and $\check{Q}_i$, each Q_i^* is weakly increasing; thus $(Q_i^*)_{i=1}^n$ is a feasible choice for (36), an upper bound of which is the maximand of (37), attained by $(Q_i^*)_{i=1}^n$. Thus, $(Q_i^*)_{i=1}^n$ is a solution of (36).

4.3.3 Proof of Theorem 2

Given any profile $(\lambda_i)_{i=1}^n$ of distributions, λ_i on T_i for each i, the objective (10) is defined. Let (Q, P) be a mechanism that maximizes (10) subject to IC, IR and BB. Then Q solves (27) and P obeys Claim (b) of the theorem with respect to Q (Lemma 4 and Remark 2). We still need to prove that Claim (a) of the theorem holds for Q. To do that, note from Q being a solution of (27) that (Q, ν) is a saddle point for some $\nu \geq 0$ with respect to the Lagrangian $\mathscr{L}$ defined by $(\lambda_i)_{i=1}^n$ via (32) and (33) (Lemma 5). Thus, Q solves (36). Section 4.3.2 has shown that the maximand of (36) is equal to the sum of the maximands in (37). Consequently, for Q to solve (36) it must solve the two problems in (37) simultaneously. That is, $(Q_i^+)_{i=1}^n$ solves (38) and $(-Q_i^-)_{i=1}^n$ solves (39). For $(Q_i^+)_{i=1}^n$ to solve (38), Lemma 6 requires that Q_i^+ be the marginal of some $\hat{q}_{iA}$, for each i, such that $(\hat{q}_{iA})_{i=1}^n \in \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$; for $(-Q_i^-)_{i=1}^n$ to solve (39), Lemma 7 requires that $-Q_i^-$ be the marginal of some $c\check{q}_{iB}$, for each i, such that $(\check{q}_{iB})_{i=1}^n \in \mathscr{A}\left((-\overline{Z}_{i,-})_{i=1}^n\right)$. That proves Claim (a) of the theorem. ■

5 Conclusion

Although the literature has long recognized auctions as the optimal means to allocate scarce resources to multiple individuals, the role of auctions has yet to be recognized, and sometimes deemed morally repugnant, when the issue is about redistribution among individuals.[9] Given wealth inequality, auctions are feared to benefit the rich and impoverish the poor. This paper, by contrast, argues that the role of auctions is essential to achieve redistributive optimality. Any interim Pareto optimum, no matter where it is located on the Pareto frontier, whether it weighs the poor more than it does the rich, or the rich more than the poor, is necessarily in the form of auctions, with the winner-selection rule adjusted to reflect the particular welfare weights associated with the particular Pareto optimum. Instead of mandating wealth transfers from one individual to another, whose idiosyncrasies are uncertain to regulators, a social planner could have used auctions to induce the right amount of wealth transfers among the right types of individuals.

This paper makes a methodology contribution to the mechanism design literature. Rather than assuming a utilitarian, one-dimension, design objective, we start with interim Pareto optimality, an objective with infinite dimensions, and show that any optimum in this infinite dimension space corresponds to a constrained optimization of a utilitarian objective obtained through the endogenous welfare ranking associated with the optimum. We introduce a new kind of operators to systematically keep track of each player's countervailing incentive of playing the role of a buyer sometimes and the role of a seller some other times. We devise a bisection technique to solve an optimal mechanism problem whose objective is nonlinear and binding constraints are multiple.

Our model is relevant to matching theory in the case where one side of the matching market has both desirable and undesirable items (e.g., toxic assets that need to be absorbed by other financial institutions; enrollment of schools in undesirable neighborhoods; thankless tasks to be carried out by some team members; donation of one's own kidney). While much of the matching theory literature assumes that money transfers are banned, our result suggests that it is suboptimal to ban money transfers from matching markets.

[9] Recall the indignant outcry expressed in the mass media when news broke that the locations of some international games were chosen through bidding, or the negative media coverage on less developed countries being paid to receive toxic, recycled materials.

A Details in Theorem 1: Convexity of $\mathbb{U}$

Pick any $(W_i^1)_{i=1}^n, (W_i^2)_{i=1}^n \in \mathbb{U}$. Thus, for some IC, IR and BB mechanisms $(Q_i^1, P_i^1)_{i=1}^n$ and $(Q_i^2, P_i^2)_{i=1}^n$ we have, for each $i = 1, \ldots, n$, each $k = 1, 2$, and any $t_i \in T_i$,

$$\begin{aligned} W_i^k(t_i) &\leq Q_i^k(t_i) - \frac{P_i^k(t_i)}{t_i}, & (43)\\ P_i^k(t_i') &= P_i^k(t_i) + \int_{t_i}^{t_i'} s dQ_i^k(s) \quad (\forall t_i' \in T_i), & (44)\\ 0 &\leq t_i Q_i^k(t_i) - P_i^k(t_i), & (45)\\ P_i^k(t_i) &= \int_{T_{-i}} p_i^k(t_i, t_{-i}) dF_{-i}(t_{-i}), & (46)\\ 0 &\leq \sum_i p_i^k(t) \quad (\forall t \in \prod_i T_i), & (47) \end{aligned}$$

and $Q^k \in \mathscr{Q}_{\text{mon}}$ and W_i^k a continuous function on T_i for each k. Here (44) coupled with $Q^k \in \mathscr{Q}_{\text{mon}}$ is equivalent to IC, (45) is equivalent to IR, (46) and (47) together mean BB. For any $\gamma \in [0, 1]$, define for each i

$$\begin{aligned} Q_i &:= \gamma Q_i^1 + (1-\gamma) Q_i^2,\\ p_i &:= \gamma p_i^1 + (1-\gamma) p_i^2. \end{aligned}$$

Then it follows from (46) that, for any i and any $t_i \in T_i$,

$$P_i = \gamma P_i^1 + (1-\gamma) P_i^2.$$

We shall show that $(Q_i, P_i)_{i=1}^n$ satisfies IC, IR and BB. Immediately from the definition of p_i and (47), BB follows. IR is proved by combining together the definition of Q_i, the fact $P_i = \gamma P_i^1 + (1-\gamma) P_i^2$, and (45) for both $k = 1, 2$. To verify IC, first note that $\gamma Q^1 + (1-\gamma) Q^2 \in \mathscr{Q}_{\text{mon}}$ by convexity of $\mathscr{Q}_{\text{mon}}$ (Appendix B.4). Second, by (44),

$$\gamma P_i^1(t_i') + (1-\gamma) P_i^2(t_i') = \gamma P_i^1(t_i) + (1-\gamma) P_i^2(t_i) + \int_{t_i}^{t_i'} s d\left(\gamma Q_i^1(s) + (1-\gamma) Q_i^2(s)\right)$$

for any $t_i, t_i' \in T_i$ and any i. Hence $(Q_i, P_i)_{i=1}^n$ is IC. Thus, $(Q_i, P_i)_{i=1}^n$ satisfies IC, IR and BB. Finally, plug $Q_i = \gamma Q_i^1 + (1-\gamma) Q_i^2$ and $P_i = \gamma P_i^1 + (1-\gamma) P_i^2$ into (43) to obtain

$$\gamma W_i^1(t_i) + (1-\gamma) W_i^2(t_i) \leq Q_i(t_i) - \frac{P_i(t_i)}{t_i}$$

for each i and any $t_i \in T_i$. This coupled with continuity of $\gamma W_i^1 + (1-\gamma) W_i^2$ implies $(\gamma W_i^1 + (1-\gamma) W_i^2)_{i=1}^n \in \mathbb{U}$, as desired.

B Details in Theorem 2

B.1 Proof of Lemma 1

For any integrable function $Q_i : T_i \to \mathbb{R}$ and any well-ordered two-part function $\varphi_i := (\varphi_{i,+}, \varphi_{i,-})$, use the definition of two-part operators and the fact $Q_i = Q_i^+ - Q_i^-$ to obtain

$$\begin{aligned}\langle Q_i : \varphi_i| &= \int_{T_i} Q_i^+(t_i)\varphi_{i,+}(t_i)dF_i(t_i) - \int_{T_i} Q_i^-(t_i)\varphi_{i,-}(t_i)dF_i(t_i) \\ &= \int_{T_i} Q_i(t_i)\varphi_{i,-}(t_i)dF_i(t_i) + \int_{T_i} Q_i^+(t_i)\left(\varphi_{i,+}(t_i) - \varphi_{i,-}(t_i)\right)dF_i(t_i).\end{aligned}$$

On the second line, the first sum on the second line is linear in Q_i; and the second sum concave in Q_i because $Q_i(t_i) \mapsto Q_i^+(t_i)$ is a convex mapping and, because $\varphi_{i,+} - \varphi_{i,-} \le 0$ a.e. on T_i (φ being well-ordered) and hence $Q_i^+(t_i)\left(\varphi_{i,+}(t_i) - \varphi_{i,-}(t_i)\right)$ is a concave function of $Q_i(t_i)$ for almost all t_i in T_i. Thus $\langle Q_i : \varphi_i|$ is concave in Q_i.

B.2 Proof of Lemma 2

Denote $t_i^0 := \tau(Q_i)$. Since (Q_i, P_i) is IC, (5) implies

$$\begin{aligned}\int_{T_i} P_i d\mu_i &= \int_{T_i}\left(t_iQ_i(t_i) - \int_{t_i^0}^{t_i} Q_i(s)ds - \tilde{U}_i(t_i^0)\right)d\mu_i(t_i) \\ &= \int_{T_i} t_iQ_i(t_i)d\mu_i(t_i) - \tilde{U}_i(t_i^0)\int_{T_i} d\mu_i - \int_{T_i}\int_{t_i^0}^{t_i} Q_i(s)dsd\mu_i(t_i).\end{aligned}$$

Decompose the last double integral to obtain

$$\begin{aligned}\int_{T_i}\int_{t_i^0}^{t_i} Q_i(s)dsd\mu_i(t_i) &= \int_{a_i}^{t_i^0}\int_{t_i^0}^{t_i} Q_i(s)dsd\mu_i(t_i) + \int_{t_i^0}^{b_i}\int_{t_i^0}^{t_i} Q_i(s)dsd\mu_i(t_i) \\ &= -\int_{a_i}^{t_i^0}\int_{t_i}^{t_i^0} Q_i(s)dsd\mu_i(t_i) + \int_{t_i^0}^{b_i}\int_{t_i^0}^{t_i} Q_i(s)dsd\mu_i(t_i) \\ &= -\int_{a_i}^{t_i^0}\int_{a_i}^{s} Q_i(s)d\mu_i(t_i)ds + \int_{t_i^0}^{b_i}\int_{s}^{b_i} Q_i(s)d\mu_i(t_i)ds \\ &= -\int_{a_i}^{t_i^0} Q_i(s)\int_{a_i}^{s} d\mu_i(t_i)ds + \int_{t_i^0}^{b_i} Q_i(s)\int_{s}^{b_i} d\mu_i(t_i)ds \\ &= \int_{a_i}^{t_i^0} Q_i^-(s)\int_{a_i}^{s} d\mu_i(t_i)ds + \int_{t_i^0}^{b_i} Q_i^+(s)\left(\int_{a_i}^{b_i} d\mu_i(t_i) - \int_{a_i}^{s} d\mu_i(t_i)\right)ds \\ &= \int_{a_i}^{t_i^0} Q_i^-(s)\rho_-(\mu_i)(s)ds - \int_{t_i^0}^{b_i} Q_i^+(s)\rho_+(\mu_i)(s)ds \\ &= -\left\langle Q_i : \rho(\mu_i)\right|,\end{aligned}$$

with the third equality due to Fubini's theorem, the second last equality due to (21) and (22), and the last equality due to (20). Plugging $\int_{T_i}\int_{t_i^0}^{t_i} Q_i(s)dsd\mu_i(t_i) = -\left\langle Q_i : \rho(\mu_i)\right|$ into the equation of $\int_{T_i} P_i d\mu_i$ displayed above, we get (23).

B.3 Proof of the Sufficiency of (24)

For each player i, denote $t_i^0 := \tau(Q_i)$ (τ defined in (6)). For each player i, define

$$c_i := t_i^0 Q_i(t_i^0) - \int_{a_i}^{t_i^0} s dQ_i(s) + \frac{1}{n-1}\sum_{j\neq i}\int_{a_j}^{b_j} s\left(1-F_j(s)\right)dQ_j(s) \tag{48}$$

and, for any $(t_i, t_{-i}) \in T_i \times T_{-i}$, let the money transfer from i to others be equal to

$$p_i(t_i, t_{-i}) := c_i + \int_{a_i}^{t_i} s dQ_i(s) - \frac{1}{n-1}\sum_{j\neq i}\int_{a_j}^{t_j} s dQ_j(s). \tag{49}$$

Integrating $p_i(t_i, t_{-i})$ across t_{-i} gives the envelope equation (5), which coupled with the monotonicity hypothesis of Q_i implies IC. The integration also implies $\tilde{U}_i(t_i^0) = 0$, hence IR follows. To complete the proof, we prove BB: It suffices to prove (25), $\sum_i p_i(t) = \sum_i \left\langle Q_i : \mathbb{I}f_i + \rho(F_i)\right|$ for all $t \in \prod_i T_i$, for then BB follows from (24). Hence pick any $t := (t_i)_{i=1}^n \in \prod_i T_i$. By (49),

$$\sum_i p_i(t) = \sum_i c_i + \sum_i \int_{a_i}^{t_i} s dQ_i(s) - \frac{1}{n-1}\sum_i\sum_{j\neq i}\int_{a_j}^{t_j} s dQ_j(s) = \sum_i c_i.$$

Thus, by (48),

$$\begin{aligned}
\sum_i p_i(t) &= \sum_i t_i^0 Q_i(t_i^0) - \sum_i \int_{a_i}^{t_i^0} s dQ_i(s) + \frac{1}{n-1}\sum_i\sum_{j\neq i}\int_{a_j}^{b_j} s\left(1-F_j(s)\right)dQ_j(s) \\
&= \sum_i t_i^0 Q_i(t_i^0) - \sum_i \int_{a_i}^{t_i^0} s dQ_i(s) + \sum_i \int_{a_i}^{b_i} s\left(1-F_i(s)\right)dQ_i(s) \\
&= \sum_i \left(t_i^0 Q_i(t_i^0) - \int_{a_i}^{t_i^0} s dQ_i(s) + \int_{a_i}^{b_i} s\left(1-F_i(s)\right)dQ_i(s)\right).
\end{aligned}$$

Calculate the two integrals in the last line through integration by parts and then combine terms to obtain

$$\begin{aligned}
\sum_i p_i(t) &= \sum_i \left(\int_{a_i}^{t_i^0} Q_i(s)ds - \int_{a_i}^{b_i} Q_i(s)\left(1-F_i(s)-sf_i(s)\right)ds\right) \\
&= \sum_i \left(\int_{a_i}^{t_i^0} Q_i(s)\left(1-\left(1-F_i(s)-sf_i(s)\right)\right)ds - \int_{t_i^0}^{b_i} Q_i(s)\left(1-F_i(s)-sf_i(s)\right)ds\right) \\
&= \sum_i \left\langle Q_i : \mathbb{I}f_i + \rho(F_i)\right|,
\end{aligned}$$

with the last line due to $t_i^0 = \tau(Q_i)$, (6), (20) and the definition of $\rho(F)$ (Eqs. (21) and (21)). That proves (25) and hence BB.

B.4 Convexity of $\mathscr{Q}_{\text{mon}}$

Let $\gamma \in [0,1]$ and $Q, \hat{Q} \in \mathscr{Q}_{\text{mon}}$. Since $Q \in \mathscr{Q}_{\text{mon}}$, it is generated by a $(q_{iA}, q_{iB})_{i=1}^n$ with $\sum_i q_{iA}(\cdot) \leq 1$ and $\sum_i q_{iB}(\cdot) \leq 1$ via (2), and Q_i is weakly increasing for all i. Likewise, $\hat{Q} = (\hat{Q}_i)_{i=1}^n$ is generated by a $(\hat{q}_{iA}, \hat{q}_{iB})_{i=1}^n$ with each $\tilde{Q}_i$ weakly increasing. Then $\sum_i (\gamma q_{iA} + (1-\gamma)\hat{q}_{iA}) \leq 1$ and $\sum_i (\gamma q_{iB} + (1-\gamma)\hat{q}_{iB}) \leq 1$; furthermore, for each i, $\gamma Q_i + (1-\gamma)\hat{Q}_i$ satisfies (2) with respect to $(\gamma q_{iA} + (1-\gamma)\hat{q}_{iA}, \gamma q_{iB} + (1-\gamma)\hat{q}_{iB})$, and is weakly increasing because both Q_i and $\hat{Q}_i$ are so. Thus $\left(\gamma Q_i + (1-\gamma)\hat{Q}_i\right)_{i=1}^n \in \mathscr{Q}_{\text{mon}}$.

B.5 Proving the Claims in the Proof of Lemma 5

$\mathbb{I}f_i + \rho(F_i)$ is well-ordered By (21) and (22)), $\rho_+(F_i) \leq \rho_-(F_i)$. This, coupled with the fact $(\mathbb{I}f_i)_+ = (\mathbb{I}f_i)_- = \mathbb{I}f_i$, implies $\mathbb{I}f_i + \rho_+(F_i) \leq \mathbb{I}f_i + \rho_-(F_i)$. ■

$\alpha_\lambda(\mathbb{I}f_i + \rho(F_i)) - \rho(\Lambda_i)$ is well-ordered Since $(\mathbb{I}f_i)_+ = (\mathbb{I}f_i)_- = \mathbb{I}f_i$, and α_λ is a coefficient, it suffices to show that $\alpha_\lambda \rho(F_i) - \rho(\Lambda_i)$ is well-ordered. To that end, let $t_i \in T_i$. By (11), (12), (21) and (22),

$$\begin{aligned}
\alpha_\lambda (\rho_+(F_i))(t_i) - (\rho_+(\Lambda_i))(t_i) &= \alpha_\lambda (-1 + F_i(t_i)) - (-\Lambda_i(b_i) + \Lambda_i(t_i)) \\
&= \alpha_\lambda F_i(t_i) - \Lambda_i(t_i) - (\alpha_\lambda - \Lambda_i(b_i)) \\
&\leq \alpha_\lambda F_i(t_i) - \Lambda_i(t_i) \\
&= \alpha_\lambda (\rho_-(F_i))(t_i) - (\rho_-(\Lambda_i))(t_i),
\end{aligned}$$

with the inequality due to (12). Thus, $\alpha_\lambda \rho(F_i) - \rho(\Lambda_i)$ is well-ordered, as desired. ■

Proof of (31) Let $Q : T_i \to \mathbb{R}$ be weakly increasing. Denote $t_i^0 := \tau(Q_i)$ for any i. Recall the notation $\mathbb{I}$ for the identity mapping $t_i \mapsto t_i$, plug (21) and (22) into $\rho(F)$, and use (20) the definition of two-part operators to obtain

$$\begin{aligned}
\langle Q_i : \mathbb{I}f_i + \rho(F_i)| &= \int_{a_i}^{b_i} Q_i^+(t_i)t_i dF_i(t_i) - \int_{a_i}^{b_i} Q_i^+(t_i)\left(1 - F_i(t_i)\right) dt_i \\
&\quad + \int_{a_i}^{b_i} Q_i^-(t_i)t_i dF_i(t_i) - \int_{a_i}^{b_i} Q_i^-(t_i)F_i(t_i)dt_i \\
&\overset{(6)}{=} -\int_{t_i^0}^{b_i} Q_i(t_i)d\left(t_i(1 - F_i(t_i))\right) + \int_{a_i}^{t_i^0} Q_i(t_i)d\left(t_iF_i(t_i)\right) \\
&= Q(t_i^0)t_i^0 + \int_{t_i^0}^{b_i} t_i(1 - F_i(t_i))dQ_i(t_i) - \int_{a_i}^{t_i^0} t_iF_i(t_i)dQ_i(t_i) \\
&\geq Q(t_i^0)t_i^0 + t_i^0\left[\int_{t_i^0}^{b_i}(1 - F_i(t_i))dQ_i(t_i) - \int_{a_i}^{t_i^0} F_i(t_i)dQ_i(t_i)\right] \\
&= Q(t_i^0)t_i^0 + t_i^0\left[\int_{a_i}^{b_i} Q_i(t_i)dF_i(t_i) - Q(t_i^0)\right] \\
&= t_i^0 \int_{T_i} Q_i(t_i)dF_i(t_i),
\end{aligned}$$

with the third and fourth equalities due to integration by parts, and the inequality due to Q_i being weakly increasing. ■

B.6 Proofs of Lemmas 6 and 7

Since (38) has a nonnegativity constraint $Q \in \mathscr{Q}_+$, and (39) a nonpositivity constraint $Q \in \mathscr{Q}_-$, we need to modify Myerson's [11, pp. 68–70] proof to prove these lemmas. For any continuous function $\varphi : [0, 1] \to \mathbb{R}$, denote the convex hull of φ by $\operatorname{conv}\varphi$, and define

$$\left(\operatorname{conv}_{\mathrm{L}}\varphi\right)(s) := \begin{cases} \min_{[0,1]}\varphi & \text{if } s \leq \inf\left(\arg\min_{r\in[0,1]}\varphi(r)\right) \\ \left(\operatorname{conv}\varphi\right)(s) & \text{else,} \end{cases} \tag{50}$$

$$\left(\operatorname{conv}_{\mathrm{R}}\varphi\right)(s) := \begin{cases} \min_{[0,1]}\varphi & \text{if } s \geq \sup\left(\arg\min_{r\in[0,1]}\varphi(r)\right) \\ \left(\operatorname{conv}\varphi\right)(s) & \text{else.} \end{cases} \tag{51}$$

For any integrable function $\psi_i : T_i \to \mathbb{R}$, $H_i\left(\psi_i\right) : [0, 1] \to \mathbb{R}$ is defined by (7) and continuous. Hence $\operatorname{conv}_{\mathrm{L}} H_i\left(Z_{i,+}\right)$ and $\operatorname{conv}_{\mathrm{R}} H_i\left(Z_{i,-}\right)$ are defined by (50) and (51). Each a convex function, their derivatives are defined for almost every $t_i \in T_i$, and weakly increasing on the set of these points:

$$\overline{\overline{Z}}_{i,+}(t_i) := \frac{d}{ds}\left(\left(\operatorname{conv}_{\mathrm{L}} H_i\left(Z_{i,+}\right)\right)(s)\right)\Big|_{s=F_i(t_i)}; \tag{52}$$

$$\overline{\overline{Z}}_{i,-}(t_i) := \frac{d}{ds}\left(\left(\operatorname{conv}_{\mathrm{R}} H_i\left(Z_{i,-}\right)\right)(s)\right)\Big|_{s=F_i(t_i)}. \tag{53}$$

Extend the definitions to all $t_i \in T_i$ to keep $\overline{\overline{Z}}_{i,+}$ and $\overline{\overline{Z}}_{i,-}$ monotone.

Proof of Lemma 6 Recall that any $(Q_i)_{i=1}^n \in \mathscr{Q}$ is generated by some ex post allocation $(q_{iA}, q_{iB})_{i=1}^n$ via (2). For any $(q_{iA}, q_{iB})_{i=1}^n$ and any i, denote

$$q_i := q_{iA} - cq_{iB}. \tag{54}$$

Then (38) is equivalent to

$$\begin{aligned} \max_{(q_i)_{i=1}^n \in \prod_i [-c,1]^{T_i}} \quad & \textstyle\sum_i \int_{T_i} \int_{T_{-i}} q_i(t_i, t_{-i}) Z_{i,+}(t_i) dF_{-i}(t_{-i}) dF_i(t_i) \quad (55) \\ \text{s.t.} \quad & \left(\textstyle\int_{T_{-i}} q_i(\cdot, t_{-i}) dF_{-i}(t_{-i})\right)_{i=1}^n \in \mathscr{Q}_{\text{mon}}, \\ & -c \le \textstyle\sum_i q_i \le 1 \quad \text{on} \quad \prod_i T_i, \\ & \textstyle\int_{T_{-i}} q_i(t_i, \cdot) dF_{-i} \ge 0 \quad \text{a.e. } t_i \in T_i \quad \forall i \in \{1, \ldots, n\}. \quad (56) \end{aligned}$$

Pick any $(\hat{q}_i)_{i=1}^n \in \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$. By the definition of $\mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$, $(\hat{q}_i)_{i=1}^n$ satisfies all the constraints in (55). Let $\hat{Q}_i$ denote the marginal of $\hat{q}_i$ for each i. We claim that $(\hat{q}_i)_{i=1}^n$ solves (55).

To that end, denote $G_i := H_i(Z_{i,+})$, $G_i^L := \operatorname{conv}_{\mathrm{L}} H_i(Z_{i,+})$, $T := \prod_i T_i$, $t := (t_i)_{i=1}^n$ and $F := \prod_i F_i$. Given any $(q_i)_{i=1}^n$, with $(Q_i)_{i=1}^n$ its marginal, the objective in (55) is equal to

$$\begin{aligned} & \int_T \sum_i q_i(t) \overline{\overline{Z}}_{i,+}(t_i) dF(t) - \sum_i \int_{T_i} \left(G_i\left(F_i(t_i)\right) - G_i^L\left(F_i(t_i)\right)\right) dQ_i(t_i) \\ + \quad & \sum_i Q_i(b_i)\left(G_i(1) - G_i^L(1)\right) - \sum_i Q_i(a_i)\left(G_i(0) - G_i^L(0)\right) \end{aligned}$$

by (52) the definition of $\overline{\overline{Z}}_{i,+}$ and integration-by-part. Here the third sum is zero because $G_i^L(1) = G_i(1)$ by definition of G_i^L, (50), for all i. Thus the objective in (55) is equal to

$$\underbrace{\int_T \sum_i q_i(t) \overline{\overline{Z}}_{i,+}(t_i) dF(t)}_{=:I((q_i)_{i=1}^n)} - \underbrace{\sum_i \int_{T_i} \left(G_i\left(F_i(t_i)\right) - G_i^L\left(F_i(t_i)\right)\right) dQ_i(t_i)}_{=:J((q_i)_{i=1}^n)} - \underbrace{\sum_i Q_i(a_i)\left(G_i(0) - G_i^L(0)\right)}_{=:K((q_i)_{i=1}^n)}.$$

Note from (52) the definition of $\overline{\overline{Z}}_{i,+}$ that, for any i and almost every $t_i \in T_i$,

$$\overline{\overline{Z}}_{i,+}(t_i) = \max\left\{0, \overline{Z}_{i,+}(t_i)\right\} \tag{57}$$

Thus, since $(\hat{q}_i)_{i=1}^n \in \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$, (9) the definition of $\mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$ implies

$$I\left((\hat{q}_i)_{i=1}^n\right) \geq I\left((q_i)_{i=1}^n\right),$$

with the inequality strict if $(q_i)_{i=1}^n \notin \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$. By definition of G_i^L, $G_i \geq G_i^L$ on $[0,1]$; by the monotonicity constraint in (55), Q_i and $\hat{Q}_i$ are weakly increasing, and $\hat{Q}_i$ by definition is constant on any interval where $G_i > G_i^L$. Thus,

$$J\left((\hat{q}_i)_{i=1}^n\right) = 0 \leq J\left((q_i)_{i=1}^n\right).$$

By (56), $Q_i(a_i) \geq 0$. If $\min \arg\min_{r\in[0,1]} \left(H_i(Z_{i,+})\right)(r) > 0$, then $\overline{Z}_{i,+} < 0$ on a neighborhood of a_i; since $(\hat{q}_i)_{i=1}^n \in \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$, that means $\hat{q}_i(t_i,\cdot) = 0$ on T_{-i} for all t_i sufficiently near a_i. Thus, with $(\hat{q}_i)_{i=1}^n$ generating $\hat{Q}$, we have $\hat{Q}_i(a_i) = 0$. If, on the other hand, $\min \arg\min_{r\in[0,1]} \left(H_i(Z_{i,+})\right)(r) = 0$, then by the definition of G_i^L and (50), we have $G_i(0) = G_i^L(0)$. Thus, in either case,

$$K\left((\hat{q}_i)_{i=1}^n\right) \leq K\left((q_i)_{i=1}^n\right).$$

It follows that $I\left((\hat{q}_i)_{i=1}^n\right) - J\left((\hat{q}_i)_{i=1}^n\right) - K\left((\hat{q}_i)_{i=1}^n\right) \geq I\left((q_i)_{i=1}^n\right) - J\left((q_i)_{i=1}^n\right) - K\left((q_i)_{i=1}^n\right)$, with the inequality strict if $(q_i)_{i=1}^n \notin \mathscr{A}\left((\overline{Z}_{i,+})_{i=1}^n\right)$. That proves the lemma. ∎

Proof of Lemma 7 This is analogous to the proof of Lemma 6, where constraint (56) is replaced here by

$$\int_{T_{-i}} q_i(t_i,\cdot)dF_{-i} \leq 0 \quad \text{a.e. } t_i \in T_i \quad \forall i \in \{1,\ldots,n\}.$$

Pick any $(\check{q}_i)_{i=1}^n \in \mathscr{A}\left((\overline{Z}_{i,-})_{i=1}^n\right)$. By the definition of $\mathscr{A}\left((\overline{Z}_{i,-})_{i=1}^n\right)$, $(c\check{q}_i)_{i=1}^n$ satisfies all the constraints in the counterpart of (55). Let $\check{Q}_i$ denote the marginal of $c\check{q}_i$ for each i. By the same token as the previous proof, $(c\check{q}_i)_{i=1}^n$ solves the counterpart of (55): Denote $G_i := H_i(Z_{i,-})$ and $G_i^R := \mathrm{conv}_{\mathrm{R}}\, H_i(Z_{i,-})$. We need only to make the following changes in the proof of Lemma 6: (i) $\sum_i Q_i(a_i)\left(G_i(0) - G_i^R(0)\right) = 0$ because $G_i(0) = G_i^R(0)$; (ii) $\sum_i Q_i(b_i)\left(G_i(1) - G_i^R(1)\right)$ is maximized by $\check{Q}$ because $\check{Q}_i(b_i) = 0$ when

$$\max\left(\arg\min_{r\in[0,1]} H_i(Z_{i,-})(r)\right) < 1,$$

and $G_i(1) = G_i^R(1)$ when the inequality does not hold. ∎

B.7 Proof of (41)

The following general observation on ironing (defined in (7) and (8)) implies (41).

Lemma 8 *For any two integrable functions φ and ϕ defined on T_i, if $\varphi \geq \phi$ on T_i then*

$$\sup \{t \in T_i : \overline{\varphi}(t) < 0\} \leq \inf \left\{t \in T_i : \overline{\phi}(t) \geq 0\right\}. \tag{58}$$

Proof Note from (7) and (8)) that the left-hand side of (58) is equal to

$$\inf \left(\arg\min_{t \in T_i} \left(H_i(\varphi)\right)\left(F_i(t)\right) \right),$$

and the right-hand side of (58) equal to

$$\inf \left(\arg\min_{t \in T_i} \left(H_i(\phi)\right)\left(F_i(t)\right) \right).$$

By (7), for any $t' > t$ the difference $\left(H(\varphi)\right)\left(F_i(t')\right) - \left(H(\varphi)\right)\left(F_i(t)\right) = \int_t^{t'} \varphi(s) dF_i(s)$ increases when φ increases pointwise. Thus, with $\varphi \geq \phi$ on T_i, $\arg\min_{t \in T_i} \left(H(\varphi)\right)\left(F_i(t)\right)$ is less than $\arg\min_{t \in T_i} \left(H(\phi)\right)\left(F_i(t)\right)$ in strong-set order (Milgrom and Shannon [9]), implying (58). ■

C Proof of Theorem 3

For each i, let $L^2(T_i)$ be the L^2-space of measurable real functions defined on T_i, endowed with the measure corresponding to F_i. Note that $L^2(T_i)$ is a normed linear space. Endow it with the weak topology.[10] Denote $\mathbb{B}_i$ for the closed ball in $L^2(T_i)$ such that the radius of $\mathbb{B}_i$ is equal to $\max\{c, 1\}$. Endow $\mathbb{L} := \prod_{i=1}^n L^2(T_i)$ with the product topology whose subspaces $L^2(T_i)$ are all in the weak topology.

Lemma 9 *$\mathscr{Q} \subset \prod_i \mathbb{B}_i$.*

Proof Let $Q \in \mathscr{Q}$. Then $Q = (Q_i)_{i=1}^n$ such that, by (2), Q_i is uniformly bounded within $[-c, 1]$ for all i. Thus, with the compact domain T_i, $Q_i \in L^2(T_i)$; furthermore, $\|Q_i\|_2 \leq \max\{1, c\}$ according to the probability measure F. Thus $Q_i \in \mathbb{B}_i$, hence $(Q_i)_{i=1}^n \in \prod_i \mathbb{B}_i$. ■

Lemma 10 *$\prod_i \mathbb{B}_i$ is compact in $\mathbb{L}^2$.*

[10] Although $\mathscr{Q}$ belongs to a normed linear space, topologizing $\mathscr{Q}$ according to the norm cannot guarantee compactness of any closed and bounded choice set (cf. Royden and Fitzpatrick [13, §13.3]).

Proof For each i, $L^2(T_i)$ is a Banach space; furthermore, $L^2(T_i)$ is the dual of itself by the Rieze representation theorem and the fact that the number 2 is its own conjugate. Thus, by Kakutani's theorem (Royden and Fitzpatrick [13, p301]), $\mathbb{B}_i$ is compact in the weak topology for each i. Then compactness of $\prod_i \mathbb{B}_i$ follows from the Tychonoff product theorem. ■

Lemma 11 *$\mathscr{Q}_{\text{mon}}$ is compact in $\mathbb{L}$.*

Proof Let $\left((Q_i^k)_{i=1}^n\right)_{k=1}^\infty$ be a sequence in $\mathscr{Q}_{\text{mon}}$. It suffices to extract an infinite subsequence that converges to some element of $\mathscr{Q}_{\text{mon}}$ in the topology of $\mathbb{L}$.

By Lemmas 9 and 10, $\left((Q_i^k)_{i=1}^n\right)_{k=1}^\infty$ has a subsequence $((Q_i^{k_l})_{i=1}^n)_{l=1}^\infty$ that converges in the topology of $\mathbb{L}$. Since the latter is the product topology of $L^2(T_i)$'s across i, each in weak topology, $(Q_i^{k_l})_{l=1}^\infty$ converges in weak topology for each $i \in \{1, \ldots, n\}$.

Pick any $i \in \{1, \ldots, n\}$. By the definition of $\mathscr{Q}_{\text{mon}}$, $Q_i^{k_l}$ is a weakly increasing function and bounded within $[-c, 1]$ for all i and all l. Thus, by Helley's selection principle,[11] $(Q_i^{k_l})_{l=1}^n$ has an infinite subsequence that converges pointwise to some $Q_i^* : T_i \to \mathbb{R}$. This being true for all $i \in \{1, \ldots, n\}$, we can extract an infinite subsequence, denoted by $((Q_i^k)_{i=1}^n)_{k=1}^\infty$ with superscripts relabeled, such that $Q_i^k \to_k Q_i^*$ pointwise on T_i for all i.

Claim: $(Q_i^*)_{i=1}^n \in \mathscr{Q}_{\text{mon}}$. For each i, with the convergence $Q_i^k \to_k Q_i^*$ pointwise, obviously Q_i^* is weakly increasing and bounded within $[-c, 1]$. Thus it suffices the claim to prove $(Q_i^*)_{i=1}^n \in \mathscr{Q}$, i.e., that $(Q_i^*)_{i=1}^n$ is generated by some ex post allocation via (2). To that end, note from $(Q_i^k)_{i=1}^n \in \mathscr{Q}$ that for each $k \in \{1, 2, \ldots\}$ there is a profile $(q_{iA}^k, q_{iB}^k)_{i=1}^n$, with $q_{iA}^k, q_{iB}^k : \prod_{j=1}^n T_j \to [0, 1]$, $\sum_i q_{iA}^k(\cdot) \leq 1$ and $\sum_i q_{iB}^k(\cdot) \leq 1$, that generates $(Q_i^k)_{i=1}^n$ according to (2). For each i let $((s_i^m)_{i=1}^n)_{m=1}^\infty$ be an enumeration of the points in $\prod_{k=1}^n T_i$ with rational coordinates. By the diagonal trick we can extract a subsequence $\left((q_{iA}^{k_m}, q_{iB}^{k_m})_{i=1}^n\right)_{m=1}^\infty$ of $\left((q_{iA}^k, q_{iB}^k)_{i=1}^n\right)_{k=1}^\infty$ such that $\left(q_{iA}^{k_m}\right)_{m=1}^\infty$ converges to some q_{iA}^*, and $\left(q_{iB}^{k_m}\right)_{m=1}^\infty$ to some q_{iB}^*, at $(s_i^m)_{i=1}^n$ for all $m = 1, 2, \ldots$. Extend the domain of $(q_{iA}^*, q_{iB}^*)_{i=1}^n$ to the rest of $\prod_i T_i$ by $q_{iA}^*(r) := \inf_{s^m \geqq r} q_{iA}^*(s^m)$ and $q_{iB}^*(r) := \inf_{s^m \geqq r} q_{iB}^*(s^m)$ for all $r \in \prod_i T_i$. Obviously, q_{iA}^* and q_{iB}^* are each a function $\prod_i T_i \to [0, 1]$, $\sum_i q_{iA}^*(\cdot) \leq 1$, and $\sum_i q_{iB}^*(\cdot) \leq 1$. For each $m = 1, 2, \ldots$, Eq. (2) implies

$$Q_i^{k_m}(\cdot) = \int_{T_{-i}} q_{iA}^{k_m}(\cdot, t_{-i}) dF_{-i}(t_{-i}) - c \int_{T_{-i}} q_{iB}^{k_m}(\cdot, t_{-i}) dF_{-i}(t_{-i})$$

[11] Kolmogorov and Fomin [6, p372].

on T_i. Now that $(q_{iA}^{k_m}, q_{iB}^{k_m}) \to_m (q_{iA}^*, q_{iB}^*)$ pointwise, the right-hand side of the above equation, by the bounded convergence theorem, converges to

$$\int_{T_{-i}} q_{iA}^*(\cdot, t_{-i}) dF_{-i}(t_{-i}) - c \int_{T_{-i}} q_{iB}^*(\cdot, t_{-i}) dF_{-i}(t_{-i})$$

pointwise on T_i. Meanwhile, $Q_i^{k_m} \to_m Q_i^*$ pointwise, as already established. Thus, for each i,

$$Q_i^*(\cdot) = \int_{T_{-i}} q_{iA}^*(\cdot, t_{-i}) dF_{-i}(t_{-i}) - c \int_{T_{-i}} q_{iB}^*(\cdot, t_{-i}) dF_{-i}(t_{-i})$$

on T_i. Consequently, $(Q_i^*)_{i=1}^n \in \mathscr{Q}$, which implies the claim $(Q_i^*)_{i=1}^n \in \mathscr{Q}_{\text{mon}}$.

Thus, to complete the proof, we show that the subsequence $((Q_i^k)_{i=1}^n)_{k=1}^\infty$ converges, in the topology of $\mathbb{L}$, to $(Q_i^*)_{i=1}^n$. It suffices to show that, for each bidder i, $(Q_i^k)_{k=1}^\infty$ converges to Q_i^* in the weak topology of $L^2(T_i)$. Thus consider any bounded linear functional Ψ on $L^2(T_i)$. By Riesz's representation theorem, the functional is the form of

$$\Psi(\varphi_i) = \int_{T_i} \varphi_i \psi dF_i$$

for some function $\psi \in L^2(T_i)$. Since $Q_i^k \to Q_i^*$ pointwise on T_i,

$$\lim_{k\to\infty} \Psi(Q_i^k) = \lim_{k\to\infty} \int_{T_i} Q_i^k \psi dF_i = \int_{T_i} \lim_{k\to\infty} Q_i^k \psi dF_i = \int_{T_i} Q_i^* \psi dF_i.$$

This being true for any continuous linear functional Ψ on $L^2(T_i)$, $(Q_i^k)_{k=1}^\infty$ converges to Q_i^* in the weak topology of $L^2(T_i)$. Hence $((Q_i^k)_{i=1}^n)_{k=1}^\infty$ converges, in the topology of $\mathbb{L}$, to $(Q_i^*)_{i=1}^n$, an element of $\mathscr{Q}_{\text{mon}}$. Thus $\mathscr{Q}_{\text{mon}}$ is compact in the topology of $\mathbb{L}$. ■

Lemma 12 *For any i and any distribution μ_i on T_i, $Q_i \mapsto \langle Q_i : \rho(\mu)|$ is continuous on $\mathscr{Q}$ in the topology of $\mathbb{L}$.*

Proof By (20), $\langle Q_i : \rho(\mu)| = \int_{T_i} Q_i^+(s) \left(\left(\rho_+(\mu_i) \right)(s) \right) ds - \int_{T_i} Q_i^-(s) \left(\left(\rho_-(\mu_i) \right)(s) \right) ds$, and Q^+ and Q^- are each continuous in Q. Thus it suffices to show continuity of

$$\begin{aligned} \varphi_i &\mapsto \int_{T_i} \varphi_i(s) \left(\left(\rho_+(\mu_i) \right)(s) \right) ds, \\ \varphi_i &\mapsto \int_{T_i} \varphi_i(s) \left(\left(\rho_-(\mu_i) \right)(s) \right) ds, \end{aligned}$$

in the weak topology of $L^2(T_i)$. By the definition of weak topology, it suffices to show that both mappings are bounded linear functionals on $L^2(T_i)$. With linearity obvious, we need only to show boundedness. To show that, note from (21) and (22) that

$$-\frac{\mu_i(b_i)}{\min_{T_i} f_i} \leq \frac{\rho_+(\mu_i)}{f_i} \leq 0,$$
$$0 \leq \frac{\rho_-(\mu_i)}{f_i} \leq \frac{\mu_i(b_i)}{\min_{T_i} f_i}.$$

Hence

$$\left|\int_{T_i} \varphi_i \cdot \frac{\rho_+(\mu_i)}{f_i} dF_i\right| \leq \frac{\mu_i(b_i)}{\min_{T_i} f_i} \int_{T_i} |\varphi_i|\, dF_i \leq \frac{\mu_i(b_i)}{\min_{T_i} f_i} \|\varphi_i\|,$$

and likewise for $\rho_-(\mu_i)$. Thus each is bounded on $L^p(T_i)$, as desired. ■

Proof of Theorem 3 By Lemma 4, it suffices to prove that problem (27) admits a solution. By Lemma 12, $\langle Q_i : \rho(F_i)|$ and $\langle Q_i : \rho(\Lambda_i)|$ are each a continuous function of Q_i in the topology of $\mathbb{L}$. Note that $\langle Q_i : \mathbb{I}f_i| = \int_{T_i} Q_i(s)sf_i(s)ds$, clearly a continuous function of Q_i in the topology of $\mathbb{L}$. Thus, $\sum_i \langle Q_i : \alpha_\lambda (\mathbb{I}f_i + \rho(F_i)) - \rho(\Lambda_i)|$, the objective in (27), is a continuous function of the choice variable Q in the topology of $\mathbb{L}$, and the constraint $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)\rangle \geq 0$ defines a closed subset of $\mathscr{Q}$ in the topology of $\mathbb{L}$. Since $\mathscr{Q}_{\text{mon}}$ is compact in the same topology (Lemma 11), the choice set of problem (27),

$$\mathscr{Q}^+_{\text{mon}} := \left\{(Q_i)_{i=1}^n \in \mathscr{Q}_{\text{mon}} : \sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0\right\},$$

is compact in the topology of $\mathbb{L}$. Since $t_i \mapsto 1/n$ (assigning the good for free randomly with equal probability among the n players and assigning the bad to none) is contained in $\mathscr{Q}^+_{\text{mon}}$, the set is also nonempty. It then follows from the generalized Weierstrass extreme value theorem that (27) admits a solution, as desired.

D Proofs of the Corollaries

D.1 Proof of Corollary 1

By hypothesis of this corollary, $F_i = F_j$ and $\lambda_i = \lambda_j$ for all players i and j. Thus, by (34) and (35), the pair $(Z_{i,+}, Z_{i,-})$ is identical across i. Hence Theorem 2 implies that Q_i is identical across i.

First, we claim that $Q_i > 0$ on $(b_i - \delta, b_i]$ for all i. To prove the claim, note from (34) that $Z_{i,+}(b_i) = (\alpha_\lambda + \nu)\, b_i > 0$ for all i. Thus, by continuity, $Z_{i,+} > 0$ on $(b_i - \delta, b_i]$ for some $\delta > 0$ for all i. Then for each i, $H_i(Z_{i,+})$ is strictly increasing on $(F_i(b_i - \delta), 1]$; thus, since 1 is the maximum of the domain for $H_i(Z_{i,+})$, its convex hull $\widehat{H}_i(Z_{i,+})$ is also strictly increasing on $(F_i(b_i - \delta), 1]$. It follows that $\overline{Z}_{i,+} > 0$ on $(b_i - \delta, b_i]$, hence Theorem 2.a implies that $Q_i > 0$ on $(b_i - \delta, b_i]$.

Second, by the hypothesis that the bad is assigned with strictly positive probability and the fact that Q_i is weakly increasing, $Q_i < 0$ on $[a_i, a_i + \epsilon]$ for some $\epsilon > 0$ and all i.

The first and second observations combined, there exists $x_i, y_i \in (a_i, b_i)$ $(\forall i)$ such that

$$Q_i(t_i) \begin{cases} < 0 & \text{if } t_i \in (a_i, x_i) \\ = 0 & \text{if } t_i \in (x_i, y_i) \\ > 0 & \text{if } t_i \in (y_i, b_i). \end{cases}$$

Recall the notation $\tilde{U}_i(\cdot \mid Q, P)$ from (4). By the envelope theorem, $\frac{d}{dt_i}\tilde{U}_i(t_i \mid Q, P) = Q_i(t_i)$ for almost every t_i, and $\tilde{U}_i(\cdot \mid Q, P)$ is absolutely continuous. Thus, $\tilde{U}_i(\cdot \mid Q, P)$ is strictly decreasing on $[a_i, x_i)$, constant on (x_i, y_i), and strictly increasing on $(y_i, b_i]$. Recall from (3) and (4) that

$$U_i(t_i \mid Q, P) = \frac{1}{t_i}\tilde{U}_i(t_i \mid Q, P)$$

for all $t_i \in T_i$. Thus, $U_i(\cdot \mid Q, P)$ is absolutely continuous on T_i[12] and, since $t_i \geq a_i > 0$, $U_i(\cdot \mid Q, P)$ is strictly decreasing on $[a_i, y_i)$. To complete the proof, we need only to show that $U_i(\cdot \mid Q, P)$ is strictly increasing on $(b_i - \delta, b_i]$ for some $\delta > 0$. To that end, pick any $t_i \in (a_i, b_i)$ at which $\tilde{U}_i(\cdot \mid Q, P)$ is differentiable and note

$$\frac{d}{dt_i}U_i(t_i \mid Q, P) = \frac{d}{dt_i}\left(\frac{\tilde{U}_i(t_i \mid Q, P)}{t_i}\right) = \frac{1}{(t_i)^2}\left(t_i Q_i(t_i) - \tilde{U}_i(t_i \mid Q, P)\right) = \frac{1}{(t_i)^2}P_i(t_i),$$

with the last equality due to (4). By the envelope equation (5), P_i is continuous and weakly increasing on T_i, hence $\lim_{t_i \uparrow b_i} P_i(t_i) = P_i(b_i) = \max_{T_i} P_i$. We claim that $P_i(b_i) > 0$, otherwise by Theorem 2.b.iii we have $P_i = 0$ on T_i, which contradicts (5), as Q_i has been proved to be

[12] It suffices to prove that $U_i(\cdot \mid Q, P)$ is Lipschitz on T_i: For any $t_i, t_i' \in T_i$, with $U_i := U_i(\cdot \mid Q, P)$,

$$\begin{aligned} |U_i(t_i') - U_i(t_i)| &= \left|\frac{1}{t_i'}\tilde{U}_i(t_i') - \frac{1}{t_i}\tilde{U}_i(t_i)\right| = \left|\frac{1}{t_i' t_i}\left(t_i\left(\tilde{U}_i(t_i') - \tilde{U}_i(t_i)\right) + \tilde{U}(t_i)\,(t_i - t_i')\right)\right| \\ &= \left|\frac{1}{t_i' t_i}\left(t_i \int_{t_i}^{t_i'} Q_i(s)ds + \tilde{U}(t_i)\,(t_i - t_i')\right)\right| \leq \left|\frac{1}{t_i' t_i}\left(t_i\,|t_i' - t_i| \max_{T_i}|Q_i| + \left|\tilde{U}(t_i)\right| |t_i - t_i'|\right)\right| \\ &\leq \frac{b_i}{a_i^2}\,(\max\{1, c\} + 1)\,|t_i' - t_i|\,, \end{aligned}$$

with the last inequality due to $-c \leq Q_i \leq 1$ and $0 \leq \tilde{U}_i \leq b_i$. Hence $U_i(\cdot \mid Q, P)$ is Lipschitz.

nonzero on positive-measure subsets of T_i. Now that $P_i(b_i) > 0$, $\lim_{t_i \uparrow b_i} P_i(t_i) = P_i(b_i)$ means that $P_i > 0$ on $(b_i - \delta, b_i]$ for some $\delta > 0$. Thus $\frac{d}{dt_i} U_i(t_i \mid Q, P) > 0$ at any differentiable point t_i in this interval. This, coupled with absolute continuity of $U_i(\cdot \mid Q, P)$, implies that $U_i(\cdot \mid Q, P)$ is strictly increasing on this interval, as desired.

D.2 Proof of Corollary 2

Theorems 1 and 2 combined, any interim Pareto optimal mechanism (Q, P) satisfies (a) in Theorem 2. That is, Q is generated by an ex post allocation $(q_{iA}, q_{iB})_{i=1}^n$ such that, for each i, the marginal of q_{iA} is equal to Q_i^+, and the marginal of cq_{iB}, Q_i^-. Since $\operatorname{supp} Q_i^+$ and $\operatorname{supp} Q_i^-$ have no overlapped interior, for any interior point t_i of $\operatorname{supp} Q_i^+$,

$$c \int_{T_{-i}} q_{iB}(t_i, \cdot) dF_{-i} = Q_i^-(t_i) = 0$$

and hence, since $c > 0$ by hypothesis, $q_{iB}(t_i, \cdot) = 0$ a.e. on T_{-i}. Since t_i is interior to $\operatorname{supp} Q_i^+$ if $0 < Q_i^+(t_i) = \int_{T_{-i}} q_{iA}(t_i, \cdot) dF_{-i}$, which holds if $q_{iA}(t_i, \cdot) > 0$ on a positive-measure subset of T_{-i}, we have:

$$q_{iA}(t_i, \cdot) > 0 \text{ on a positive-measure subset of } T_{-i} \Longrightarrow q_{iB}(t_i, \cdot) = 0 \text{ a.e. on } T_{-i}.$$

Analogously, the above holds when the roles of A and B are switched. Thus, $q_{iA} q_{iB} = 0$ a.e. on $\prod_k T_k$ for all i, i.e., the mechanism (Q, P) satisfies assignment exclusivity, as desired.

D.3 Proof of Corollary 3

Lemma 13 *For any solution Q of (27), if $Q_i \geq 0$ on $(a_i, b_i]$ for any i, then $\sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| > 0$.*

Proof For any i, $Q_i \geq 0$ on $(a_i, b_i]$ means that $\tau(Q_i) = a_i$ (with $\tau(Q_i)$ defined in (6)). Then (31) implies

$$\sum_i \langle Q_i : \mathbb{I} f_i + \rho(F_i)| \geq \left(\min_i a_i\right) \sum_i \int_{T_i} Q_i(t_i) dF_i(t_i).$$

We claim that $\sum_i \int_{T_i} Q_i(t_i) dF_i(t_i) > 0$. Otherwise, since $Q_i \geq 0$ for all i, $Q_i = 0$ for all i. Consequently, the objective in (27) is equal to

$$\sum_i \langle 0 : \alpha_\lambda (\mathbb{I} f_i + \rho(F_i)) - \rho(\Lambda_i)| = 0.$$

Thus, by Lemma 4, the social welfare $\sum_{i=1}^{n}\int_{T_i} U_i(\cdot \mid Q,P)d\lambda_i = 0$. But then Q is suboptimal because assigning the good to any i for free, for whom $\lambda_i > 0$ on a positive-measure subset of T_i (such i exists because, by Theorem 1, λ_k's are not identically zero), generates a positive social welfare. Thus $\sum_i \int_{T_i} Q_i(t_i)dF_i(t_i) > 0$, hence $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| > 0$. ■

Proof of Corollary 3 Let $(Q_i, P_i)_{i=1}^n$ be any Pareto optimum specified by the hypothesis. Then it is determined by $(Z_{i,+}, Z_{i,-})_{i=1}^n$ (Theorem 2). By (34) and (35) and continuous differentiability of f_i at a_i, and the definition of λ_i' $(= d\lambda_i/dF_i)$, one can show, for each i, that $\frac{d}{dt_i}Z_{i,+}$ and $\frac{d}{dt_i}Z_{i,-}$ are continuous at a_i and

$$\begin{aligned}
\frac{d}{dt_i}Z_{i,+}(a_i) &= 2\left(\alpha_\lambda + \nu\right) - \frac{\lambda_i'(a_i)}{a_i} + \frac{f_i'(a_i)}{f_i^2(a_i)}\left(\alpha_\lambda - \Lambda_i(b_i)\right) + \frac{\nu f_i'(a_i)}{f_i^2(a_i)}, \\
\frac{d}{dt_i}Z_{i,-}(a_i) &= 2\left(\alpha_\lambda + \nu\right) - \frac{\lambda_i'(a_i)}{a_i}.
\end{aligned}$$

Case (i): $\nu = 0$. Then, by (13), $\frac{d}{dt_i}Z_{i,-}(a_i) < 0$ for any i, and $\frac{d}{dt_{i_*}}Z_{i_*,+}(a_{i_*}) < 0$ for the i_* that maximizes $\Lambda_i(b_i)$ among all i (so $\alpha_\lambda - \Lambda_{i_*}(b_{i_*}) = 0$). Thus, since $\frac{d}{dt_i}Z_{i,+}$ and $\frac{d}{dt_i}Z_{i,-}$ are continuous at a_i, both $Z_{i_*,+}$ and $Z_{i_*,-}$ are strictly decreasing on $[a_{i_*}, a_{i_*} + \delta)$ for some $\delta > 0$. Then $H_{i_*}(Z_{i_*,+})$ and $H_{i_*}(Z_{i_*,-})$ by (7) are strictly concave, and hence their convex hulls affine, on $[F_{i_*}(a_{i_*}), F_{i_*}(a_{i_*} + \delta))$. Thus $\overline{Z}_{i_*,+}$ and $\overline{Z}_{i_*,-}$ are constant on $[a_{i_*}, a_{i_*} + \delta)$. Then Claims (a) of Theorem 2 implies that Q_{i_*} is constant on this neighborhood.

Case (ii): $\nu > 0$. We claim that there exists some i for whom $Q_i < 0$ on a neighborhood in T_i. Otherwise, the constraint $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$ is non-binding (Lemma 13), which coupled with the saddle point condition (30) implies that $\nu = 0$, contradiction. Now that $Q_i < 0$ on a neighborhood in T_i for some i, it follows from monotonicity (IC) of Q_i that $Q_i < 0$ on $[a_i, a_i + \eta)$ for some $\eta > 0$. Then Theorem 2.a implies $\overline{Z}_{i,-} < 0$ on $[a_i, a_i + \eta)$. By definition of ironing, if $\overline{Z}_{i,-}$ is not constant on a neighborhood of a_i, then $\overline{Z}_{i,-} = Z_{i,-}$ on that neighborhood and then $Z_{i,-} < 0$ on that neighborhood, contradicting the fact that $Z_{i,-}(a_i) = (\alpha_\lambda + \nu)\, a_i > 0$ and that $Z_{i,-}$ is continuous. Thus, on a neighborhood of a_i, $\overline{Z}_{i,-}$ is constant, and hence so is Q_i.

E Possibility of Binding IC, IR and BB Constraint

The bisection method in Section 4.3.2 is needed to solve the Lagrange problem because its objective $\sum_i \langle Q_i : Z_i f_i|$ may be nonlinear in the choice variable Q. Clearly $\sum_i \langle Q_i : Z_i f_i|$ is nonlinear if and only if $(Z_{i,+})_{i=1}^n \neq (Z_{i,-})_{i=1}^n$, which holds in general when Λ_i is not identical across all players i. Even if Λ_i is identical across all players i, $(Z_{i,+})_{i=1}^n \neq (Z_{i,-})_{i=1}^n$ still holds if $\nu > 0$, i.e., if the combined IC, IR and BB constraint, $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$, is binding.

Thus, consider the special case where for some constant β

$$\forall i = 1, \ldots, n : \Lambda_i(b_i) = \beta. \tag{59}$$

Then the objective in (27) becomes a linear functional on $\mathscr{Q}$:

$$\sum_i \int_{T_i} Q_i Y_i dF_i, \tag{60}$$

where, for any player i and any $t_i \in T_i$,

$$Y_i(t_i) := \beta t_i + \frac{\beta F_i(t_i) - \Lambda_i(t_i)}{f_i(t_i)}. \tag{61}$$

Thus, had $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$ been relaxed, Problem (27) would become a linear programming of maximizing (60) among all $Q \in \mathscr{Q}_{\text{mon}}$, thus for any solution Q^* for (27), $\left((Q_i^*)^+\right)_{i=1}^n$ would be the profile of marginals of an element of $\mathscr{A}\left((\overline{Y}_i)_{i=1}^n\right)$ and $\left((Q_i^*)^-\right)_{i=1}^n$ that of $\mathscr{A}\left((-\overline{Y}_i)_{i=1}^n\right)$. However, the next remark shows existence of parametric configurations that satisfies (59) and yet the constraint $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$ is binding in every solution for (27).

Remark 3 Let $0 < L < H$ such that $\ln(H/L) \geq 11/6$. Pick any $m = 1, 2, \ldots$ large enough for $L < H - 1/m < H$. For any $t \in [L, H]$, let

$$\phi(t) := t^3/3 - (H - 1/m)t^2 + (H - 1/m)^2 t + t/m^4.$$

Consider a symmetric-player case where the common distribution F is defined by

$$F(t) = \frac{\phi(t) - \phi(L)}{\phi(H) - \phi(L)}$$

for all t in its support $[L, H]$. Suppose further that $\lambda_i = 1$ for all players i. Then (59) is satisfied, and hence the objective in (27) becomes the linear form (60). By (12), (61), and the parametric condition $\ln(H/L) \geq 11/6$, one can show (Appendix E.1):

$$Y_i(H - 1/m) \to -\infty \quad \text{as} \quad m \to \infty. \tag{62}$$

Should the constraint $\sum_i \langle Q_i : \mathbb{I}f_i + \rho(F_i)| \geq 0$ be non-binding at a solution Q^* for (27), Q^* would be the concatenation of the allocation by the rank of $\left(\overline{Y}_i\right)_{i=1}^n$ and the allocation by the rank of $\left(-\overline{Y}_i\right)_{i=1}^n$. By (62), for all sufficiently large m, $\overline{Y}_i < 0$ on $[L, H-1/m)$ and hence $Q_i^* < 0$ on $[L, H-1/m)$ for all i. This, coupled with the fact that $\mathbb{I} + \rho_+(F_i)/f_i \le b_i = H$ and $\mathbb{I} + \rho_-(F_i)/f_i \ge a_i = L$ for all i (due to (21 and (22)), implies that

$$\begin{aligned}
\sum_i \langle Q_i^* : \mathbb{I} f_i + \rho(F_i)| &\le H \sum_i \int_{T_i} \max\{0, Q_i^*(t_i)\} dF(t_i) - L \sum_i \int_{T_i} \max\{0, -Q_i^*(t_i)\} dF(t_i) \\
&\le H\left(1 - F(H-1/m)^n\right) - Lc\left(1 - (1 - F(H-1/m))^n\right),
\end{aligned}$$

which is negative for all sufficiently large m, contradiction.

E.1 Proof of (62)

Denote $\Delta := \phi(H) - \phi(L)$. Note that the density function is, for all $t \in [L, H]$,

$$f(t) = \frac{1}{\Delta}\left((t - (H - 1/m))^2 + 1/m^4\right).$$

By the identical distribution and the definition of Λ_i in (12), for all i, $\Lambda_i(b_i) = \alpha_\lambda$ such that

$$\begin{aligned}
\alpha_\lambda &= \frac{1}{\Delta}\left(\frac{1}{2}(H^2 - L^2) - 2(H - 1/m)(H - L) + ((H-1/m)^2 + 1/m^4)\ln(H/L)\right) \\
&= \frac{1}{\Delta}\left(H^2 \ln(H/L) - (H - L)\left(\frac{3}{2}H - \frac{1}{2}L\right)\right) + O(1/m) \\
&> \frac{1}{\Delta} H^2\left(\ln(H/L) - 3/2\right) + O(1/m),
\end{aligned}$$

with the inequality due to $H > L$. Plug into this the definitions of Δ to obtain

$$\begin{aligned}
\alpha_\lambda - \frac{1}{H - 1/m} &> \frac{H^2\left(\ln(H/L) - 3/2\right)}{(H^3 - L^3)/3 - H(H^2 - L^2) + H^2(H - L)} - \frac{1}{H} + O(1/m) \\
&= \frac{H^2\left(\ln(H/L) - 3/2\right)}{(H - L)^3/3} - \frac{1}{H} + O(1/m) \\
&> \frac{1}{H}\left(3\left(\ln(H/L) - 3/2\right) - 1\right) + O(1/m) \\
&\ge O(1/m), \qquad (63)
\end{aligned}$$

with the last line due to $\ln(H/L) \ge 11/6$. By definitions of G and α_λ,

$$\begin{aligned}
G(H - 1/m) - \alpha_\lambda F(H - 1/m) &= \int_L^{H-1/m} \frac{1}{s} dF(s) - F(H - 1/m) \int_L^H \frac{1}{s} dF(s) \\
&= (F(H) - F(H - 1/m)) \int_L^H \frac{1}{s} dF(s) - \int_{H-1/m}^H \frac{1}{s} dF(s) \\
&> (F(H) - F(H - 1/m))\left(\alpha_\lambda - 1/(H - 1/m)\right) \\
&> O(1/m^2) O(1/m) \\
&= O(1/m^3),
\end{aligned}$$

with the second last line due to Taylor's formula and (63). Plug the above-derived inequality and the fact $f(H-1/m) = 1/(\Delta m^4) = o(1/m^3)$ into the definition of Y_i to obtain

$$Y_i(H-1/m) = \alpha_\lambda(H-1/m) - \left|\frac{O(1/m^3)}{o(1/m^3)}\right| = -\left|\frac{1}{O(1/m)}\right| \longrightarrow_m -\infty.$$

References

[1] Pierre-André Chiappori. *Matching with Transfers: The Economics of Love and Marriage*. Princeton University Press, 2017. 1

[2] Peter Cramton, Robert Gibbons, and Paul Klemperer. Dissolving a partnership efficiently. *Econometrica*, 55(3):615–632, May 1987. 1

[3] Piotr Dworczak, Scott Duke Kominers, and Mohammad Akbarpour. Redistribution through markets. Mimeo, March 28, 2019. 1, 1, 4

[4] Herodotus. *The Histories*. Penguin Books: Baltimore, 1954. Translated by Aubrey de Sélincourt; with an introduction and notes by A. R. Burn. 2

[5] M. Ali Khan and Edward E. Schlee. Money-metrics in applied welfare analysis: A saddlepoint rehabilitation. Mimeo, April 18, 2019. 1

[6] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis*. New York: Dover Publications, 1970. Translated by R. A. Silverman. 11

[7] Simon Loertscher and Cédric Wasser. Optimal structure and dissolution of parterships. *Theoretical Economics*. Forthcoming. 1

[8] David G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1969. 4.3.1

[9] Paul Milgrom and Chris Shannon. Monotone comparative statics. *Econometrica*, 62(1):157–180, 1994. B.7

[10] J. A. Mirrlees. An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38(2):175–208, 1971. 1

[11] Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, February 1981. 1, 3.2, B.6

[12] Roger B. Myerson and Mark A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29:265–281, 1983. 1

[13] H.L. Royden and P.M. Fitzpatrick. *Real Analysis*. Pearson Education, Boston, 4th edition, 2010. 4.1, 10, C

[14] Nancy L. Stokey, Robert E. Lucas, Jr., and Edward C. Prescott. *Recursive Methods in Economic Dynamics*. Harvard Univerity Press, 1989. 1