Utilitarian Representation of Interim Efficient Mechanisms Given Continuum Types*

Charles Z. Zheng†

April 26, 2020

Abstract

In a general quasilinear model that allows for multidimensional types, countervailing incentives and continuum of types, any interim incentive efficient mechanism is a maximum of a one-dimensional welfare function subject to incentive compatibility, individual rationality and ex post feasibility. This welfare function aggregates individual preferences across players and across types through a profile of Radon measures, one for the types of each player.

Keywords: interim incentive efficient mechanism, multidimensional type, continuum types, separating hyperplane

JEL Classification Numbers: D44, D63, D82, H41

*I thank Mingshi Kang for discussions and comments. Much of this research was conducted during the author's visit at Carnegie Mellon University. The financial support from the Social Science and Humanities Research Council of Canada, Insight Grant R4809A04, is gratefully acknowledged.

†Department of Economics, University of Western Ontario, London, ON, Canada, N6A 5C2, charles.zheng@uwo.ca, https://sites.google.com/site/charleszhenggametheorist/.

1 Introduction

Initially formulated by Holmström and Myerson [4], interim incentive efficiency (IIE) is a profound concept for mechanism design that incorporates the interests from various parties across their various private information rather than merely optimize on behalf of only one of the parties. It is particularly relevant to political issues such as wealth distributions across individuals or across groups. There, the social surplus, or a simple sum of the expected payoffs across individuals, is insufficient to capture the various preference intensities across individuals and across groups.

IIE has been studied by Gresik [3] and Wilson [15] on bilateral trade, Laussel and Palfrey [6] and Ledyard and Palfrey [7] on public good mechanisms, and Ledyard and Palfrey [8] and Pérez-Nievas [10] on a broad class of quasilinear environments. Ledyard and Palfrey [8] have provided a forceful motivation for IIE from both normative and positive perspectives. Recently, in the IIE spirit of allowing for arbitrary welfare weights across player-types, Dworczak, Kominers and Akbarpour [2] consider redistributions in public finance.

Requiring immunity to Pareto improvement across types of any player, IIE is a multidimensional design objective. To characterize IIE through mechanism design techniques, the literature relies on a *utilitarian representation* assertion that IIE implies maximization of a one-dimensional objective that aggregates the preferences across the multiple player-types through an endogenous welfare weight distribution, which assigns measures to sets of player-types. While it is trivial given finitely many possible types, as in Holmström and Myerson [4], the utilitarian representation assertion, even if true, is nontrivial given a continuum of possible types, as occurs in many applications and in all the other works cited above. In those works except Pérez-Nievas [10], the assertion is either implicitly assumed or stated without a proof. Pérez-Nievas [10] gives a clever proof of the assertion, though the proof relies on the assumption that types are all one-dimensional.[1] This note proves the representation assertion in a quasilinear environment that is broader than those in the above-cited literature. I allow for multidimensional types (e.g., multiple-object auctions) and countervailing incentives when a player can act as a buyer or as a seller endogenously. Both cases are new to the above-cited works that assume continuum of types.[2]

The complication of the theorem caused by a continuum of types is analogous to the complication that macroeconomists face in decentralizing a planner's optimum into a price system given infinite-horizon models, with our profile of interim expected payoffs across

[1] It uses a clever maneuver of integration-by-parts along the dimension of each player's type. With multidimensional types, this step would not go through unless the endogenous welfare distribution is somehow stochastically independent across the dimensions of each player's type (cf. McAfee and McMillan [9]).

[2] Dworczak et al. [2] argue that their model can be extended to a case with two-dimensional types. They implicitly assume the utilitarian representation and do not state or prove it.

player-types corresponding to their consumption streams across periods. As is explained in Stokey, Lucas and Prescott [14, §15.4, §16.6], one would need an appropriate topological vector space for such infinite-dimensional objects to guarantee existence of a separating hyperplane and ensure that the separating hyperplane can be represented by an economically meaningful operator. My proof exploits a basic fact in mechanism design that a player's interim expected payoff is a continuous function of the type. That allows us to formulate the space of payoff profiles as the product of the spaces of continuous functions defined on compact cubes, thereby applying the Riesz-Markov theorem to represent the player-type preference aggregator by a profile of player-specific Radon measures of their types.

2 The Environment

There are n players ($n \geq 2$), each assumed risk neutral, and m kinds of transferable objects other than money ($m \geq 1$). Each player's type belongs to a set $T_i := \prod_{j=1}^{m}[a_{ij}, b_{ij}]$ such that either T_i is singleton or the type is independently drawn according to an absolutely continuous cdf F_i whose support is T_i.

Any *outcome* takes the form $((x_{ij})_{j=1}^{m}, y_i)_{i=1}^{n} \in (\mathbb{R}^m \times \mathbb{R})^n$ such that:

i. the non-monetary outcome $((x_{ij})_{j=1}^{m})_{i=1}^{n}$, with x_{ij} denoting the quantity of player i's net receipt of the jth kind of objects, belongs to a commonly known subset X of $\mathbb{R}^{mn}$;

ii. the monetary transfer configuration $(y_i)_{i=1}^{n}$, with y_i being player i's net monetary payment, satisfies an aggregate condition that $\sum_{i=1}^{n} y_i$ belongs to a subset $Y\left(((x_{ij})_{j=1}^{m})_{i=1}^{n}\right)$ of $\mathbb{R}_+$, where Y is a commonly known correspondence.

Denote $x_i := (x_{ij})_{j=1}^{m}$, $t_i := (t_{ij})_{j=1}^{m}$, and $x_i \cdot t_i := \sum_j x_{ij}t_{ij}$. Given any type $t_i \in T_i$, player i's preference relation on the outcomes is represented by the vNM utility function

$$(x_i, y_i) \longmapsto x_i \cdot t_i - y_i. \tag{1}$$

Assumption 1 *X is convex and contains the zero vector $\mathbf{0}$ in $\mathbb{R}^{mn}$; $0 \in Y(\mathbf{0})$.*

Assumption 2 *There exist $(x_i^0)_{i=1}^{n} \in X$ and $(y_i^0)_{i=1}^{n} \in \mathbb{R}^n$ for which $\sum_{i=1}^{n} y_i^0 \in Y((x_i^0)_{i=1}^{n})$ and $x_i^0 \cdot t_i - y_i^0 > 0$ for all $i \in \{1, \ldots, n\}$ and all $t_i \in T_i$.*

Assumption 3 *For any $x \in X$, $Y(x)$ is convex and, for any $x' \in X$ and any $\alpha \in [0, 1]$,*

$$Y(\alpha x + (1-\alpha)x') \supseteq \alpha Y(x) + (1-\alpha)Y(x'). \tag{2}$$

Assumption 1 says it is physically feasible that everyone gets zero payoff. Assumption 2 requires that there be an outcome that, if implemented, gives positive payoff to everyone. In most applications, the correspondence Y in Assumption 2 is simply the set $[0, \infty)$ constantly, requiring only ex post budget balancing. More generally, however, the assumption allows $Y(x)$ to vary with x so as to include the case of public good provision, where the minimal amount of aggregate payment is a function of the quantity of the public good. The convexity conditions in Assumptions 1 and 3 are natural because the representation theorem, essentially a separating hyperplane argument, inevitably involves convex analysis.[3]

This environment has two features. First, types can be multidimensional. Second, a player has countervailing incentives, playing the role of a buyer and that of a seller, depending on the realized type and the mechanism: x_{ij}, the counterpart to player i's winning probability of object j in the usual auction models, can be positive–so the player acts as a buyer—or negative—so the player acts as a seller. Following are the main special cases:

Provision of a public good This is the case where $m = 1$,

$$X = \left\{ (x_i)_{i=1}^n \in [0,1]^n \mid x_1 = \cdots = x_n \right\},$$

and, for some $c \geq 0$, $Y\left((x_i)_{i=1}^n\right) = [cx_1, \infty)$ for all $(x_i)_{i=1}^n \in X$. Here the parameter c is the per-unit cost of providing the public good.

Multiunit auctions with capacity constraints for any coalition Let $m = 1$. The capacity constraints, as is formulated by Che, Kim and Mierendorff [1], correspond to a pair (L, C) of functions $L, C : 2^{\{1,\ldots,n\}} \to \mathbb{R}_+$, with $L \leqq C$, such that any outcome $(x_i, y_i)_{i=1}^n \in \mathbb{R}^{2n}$ is required to satisfy $L(G) \leq \sum_{i \in G} x_i \leq C(G)$ for any subset G of $\{1, \ldots, n\}$. These inequalities obviously define a convex compact polytope X in $\mathbb{R}^n$.

Multiple-object auctions This is the case where $m > 1$. Various degrees of complementarity or substitutability across objects for player i can be captured by the various shapes of the projection of X onto the subspace for x_i. For instance, if the projection is a lattice with respect to the coordinate-wise $\geq$ partial ordering on $\mathbb{R}^m$, then acquiring a larger quantity x_{ij} of object j makes it feasible for player i to acquire a larger quantity x_{ik} of object k. By contrast, if the projection is a hyperplane in $\mathbb{R}^m$ then the various objects are perfect substitutes to player i.

[3] In particular, condition (2) plays only one role in the proof: if a payment configuration $(y_i)_{i=1}^n$ is physically feasible given non-monetary outcome $(x_i)_{i=1}^n$, and another payment configuration $(y_i')_{i=1}^n$ physically feasible given another outcome $(x_i')_{i=1}^n$, then any convex combination between $(y_i)_{i=1}^n$ and $(y_i')_{i=1}^n$ is physically feasible given the corresponding convex combination between $(x_i)_{i=1}^n$ and $(x_i')_{i=1}^n$.

Partnership dissolution This is the case where $m = 1$,

$$X = \left\{ (x_i)_{i=1}^n \in \prod_{i=1}^n [-e_i, 1 - e_i] \;\middle|\; \sum_{i=1}^n x_i = 0 \right\},$$

where $(e_i)_{i=1}^n \in [0,1]^n$ is a list of parameters for which $\sum_i e_i = 1$, and $Y(x) = \{0\}$ for all $x \in X$. Here, $e_i \in [0,1]$ is interpreted as player i's initial share of the partnership, $Y(x) = \{0\}$ the budget balance condition that allows for no undistributed monetary surplus, and x_i player i's net increase in the share of the partnership. This case is equivalent to an **exchange economy** where e_i is i's endowment of the good, and x_i player i's net trade thereof.

Allocation of a good and a NIMBY This is the case where $m = 1$,

$$X = \left\{ (x_i)_{i=1}^n \in [-c, 1]^n \;\middle|\; -c \leq \sum_{i=1}^n x_i \leq 1 \right\},$$

and $Y(x) = [0, \infty)$ for all $x \in X$. Here x_i corresponds to a lottery involving two items, a good and a bad (NIMBY), such that the value of the good is equal to one, and that of the bad equal to $-c$, to all players. Thus $x_i = \pi_{iA} - c\pi_{iB}$ for some $(\pi_{iA}, \pi_{iB}) \in [0,1]^2$ such that player i receives the good with probability π_{iA}, and the bad with probability π_{iB} (cf. Kang and Zheng [5]).

3 Incentive Efficiency and Utilitarian Representation

Denote $T := \prod_i T_i$. Without loss of generality, we shall restrict attention to *direct revelation mechanisms (DRM)*, each in the form of a profile $(q_i, p_i)_{i=1}^n$ of functions $q_i : T \to \mathbb{R}^m$ and $p_i : T \to \mathbb{R}$ that assigns, to every profile $t \in T$ of realized types across players, an outcome $(q_i(t), p_i(t))_{i=1}^n$. Recall the definition of outcomes in Section 2.

For any i, denote $T_{-i} := \prod_{j \neq i} T_j$ and F_{-i} for the cumulative distribution on T_{-i} generated by $(F_j)_{j \neq i}$. A *reduced form* means a profile $(Q_i, P_i)_{i=1}^n$ of functions $Q_i : T_i \to \mathbb{R}^m$ and $P_i : T_i \to \mathbb{R}$ for all i. A reduced form $(Q_i, P_i)_{i=1}^n$ is said to be *ex post feasible (XF)* iff there exists a DRM $(q_i, p_i)_{i=1}^n$ such that, for each i, Q_i is the marginal of q_i onto T_i, i.e., $Q_i(t_i) = \int_{T_{-i}} q_i(t_i, t_{-i}) dF_{-i}(t_{-i})$ for any $t_i \in T_i$, and P_i is the marginal of p_i onto T_i.

For any reduced form $(Q, P) := (Q_i, P_i)_{i=1}^n$, any player i and any $t_i \in T_i$, define

$$U_i(t_i | Q, P) := t_i \cdot Q_i(t_i) - P_i(t_i). \tag{3}$$

Incentive compatibility (IC) of (Q, P) means that, for any player i and any $t_i \in T_i$,

$$U_i(t_i | Q, P) = \max_{\hat{t}_i \in T_i} \left\{ t_i \cdot Q_i(\hat{t}_i) - P_i(\hat{t}_i) \right\}. \tag{4}$$

Individual rationality (IR) of a reduced form (Q, P) means that $U_i(t_i|Q, P) \geq 0$ for any player i and any type $t_i \in T_i$.

A reduced form (Q', P') *interim Pareto dominates* another reduced form (Q, P) if and only if (i) (Q', P') is XF, IC and IR; and (ii) $U_i(\cdot|Q', P') \geq U_i(\cdot|Q, P)$ a.e. $[F_i]$ on T_i for all player i, and there exists a player i and a subset $S \subseteq T_i$ such that S is of (strictly) positive F_i-measure and $U_i(\cdot|Q', P') > U_i(\cdot|Q, P)$ on S.

A reduced form (Q, P) is *interim incentive efficient (IIE)* if and only if (i) (Q, P) is XF, IC and IR; and (ii) it is not interim Pareto dominated by a reduced form.

Theorem *For any IIE (Q^*, P^*) there exists a profile $(\Lambda_i)_{i=1}^n$ such that Λ_i is a Radon measure on T_i for each i, with Λ_i not identically zero for some i, and (Q^*, P^*) maximizes*

$$\sum_{i=1}^{n} \int_{T_i} U_i(t_i|Q, P) d\Lambda_i(t_i) \tag{5}$$

among all (Q, P) that are IC, IR and XF.

Remark 1 In (5), individual player-type preferences are aggregated by the player-specific Radon measures $(\Lambda_i)_{i=1}^n$. Such form of aggregators affords as much tractability as the standard integration-by-parts routine in mechanism design requires. That is because Fubini's theorem applies to the product measure formed by Λ_i and F_i. See Kang and Zheng [5]) for an application in a countervailing incentive environment.

Remark 2 $\Lambda_i(T_i)$ corresponds to the ex ante expected welfare weight assigned to player i. In the extant literature on IIE such as Ledyard and Palfrey [7], mechanisms are not subject to the IR constraint and the counterpart of $\Lambda_i(T_i)$ is thus identical across players. In my model, mechanisms are required to be IR, it is possible that $\Lambda_i(T_i) > \Lambda_j(T_j)$. In that case, the ex ante expected value of monetary surplus raised from implementing an allocation is transferred, as lump sums, only to those players belonging to $\arg\max_i \Lambda_i(T_i)$.

Remark 3 Even if the welfare weight Λ_i on player i is identically zero on T_i, an optimal mechanism of (5) does not always give zero surplus to i. With player i assigned zero welfare weight, the optimal mechanism transfers the expected revenue extracted from i to those players who have maximum ex ante welfare weights ($\arg\max_j \Lambda_j(T_j)$). To extract such expected revenues from player i, the mechanism needs to concede some rent to i.

Remark 4 Since my model allows a player's type space to be singleton, if this player is the seller in an auction environment that has no private information, when Λ_i is identically zero for all bidders and nonzero only for the seller, a maximum of (5) becomes the revenue-maximizing auction in the traditional sense.

4 The Proof of the Theorem

Characterization of Incentive Compatibility For any integrable function $g : \mathbb{R}^m \to \mathbb{R}$ and any $v, w \in \mathbb{R}^m$, with $v := (v_j)_{j=1}^m$ and $w := (w_j)_{j=1}^m$, denote

$$\begin{aligned}\int_v^w g(s) \cdot ds \ &:= \ \int_{v_1}^{w_1} g(\tau, v_2, \ldots, v_m) d\tau + \int_{v_2}^{w_2} g(w_1, \tau, v_3, \ldots, v_m) d\tau + \cdots \\ &\quad + \int_{v_{m-1}}^{w_{m-1}} g(w_1, \ldots, w_{m-2}, \tau, v_m) d\tau + \int_{v_m}^{w_m} g(w_1, \ldots, w_{m-2}, w_{m-1}, \tau) d\tau.\end{aligned}$$

Lemma 1 *A reduced form (Q, P) is IC if and only if, for any player i, the function $U_i(\cdot|Q, P)$ defined by (3) is convex on T_i, and*

$$\forall t_i \in T_i : U_i(t_i|Q, P) - U_i(a_i|Q, P) = \int_{a_i}^{t_i} Q_i(s_i) \cdot ds_i. \tag{6}$$

Proof The proof is similar to Rochet [11, Prop. 2]. First, suppose that (Q, P) is IC. Then for any player i and any $t_i \in T_i$, (4) holds. Thus, $U_i(\cdot|Q, P)$, the supremum of linear functions, is convex. For any $t_i := (t_{ij})_{j=1}^m \in T_i$ and any $j \in \{1, \ldots, m\}$, denote

$$t_i^{\leq j} := (t_{ik})_{k=1}^j, \quad t_i^{\geq j} := (t_{ik})_{k=j}^m, \quad t_i^{\leq 0} := t_i^{\geq n+1} := \text{null}.$$

For any $j \in \{1, \ldots, m\}$, (4) implies that, for any $t_{ij} \in [a_{ij}, b_{ij}]$, setting $s_{ij} := t_{ij}$ solves

$$\max_{s_{ij} \in [a_{ij}, b_{ij}]} \left(t_i^{\leq j-1}, t_{ij}, a_i^{\geq j+1}\right) \cdot Q_i\left(t_i^{\leq j-1}, s_{ij}, a_i^{\geq j+1}\right) - P_i\left(t_i^{\leq j-1}, s_{ij}, a_i^{\geq j+1}\right).$$

Applying the Milgrom-Segal envelope theorem to the above problem, we have

$$U_i\left(t_i^{\leq j-1}, t_{ij}, a_i^{\geq j+1} \mid Q, P\right) - U_i\left(t_i^{\leq j-1}, a_{ij}, a_i^{\geq j+1} \mid Q, P\right) = \int_{a_{ij}}^{t_{ij}} Q_i\left(t_i^{\leq j-1}, s_{ij}, a_i^{\geq j+1}\right) ds_{ij}.$$

Summing this equation across all j, we obtain (6).

Conversely, suppose that $U_i(\cdot|Q, P)$ defined by (3) is convex and satisfies (6). For any $v \in T_i$, now that $U_i(\cdot|Q, P)$ is convex, any $x \in \mathbb{R}^m$ is a subgradient of $U_i(\cdot|Q, P)$ at v if

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha} \left(U_i(v + \alpha(w - v) \mid Q, P) - U_i(t_i \mid Q, P)\right) \geq x \cdot (w - v)$$

for any $w \in \mathbb{R}^m$ (by Rockafellar [12, Theorem 23.2, p216]). This inequality is satisfied by $x := Q_i(v)$ for all w, due to (6) applied to the case where $t_i = v + \alpha(w - v)$ and to the case where t_i is itself. Thus,

$$Q_i(v) \in \partial U_i(v|Q, P)$$

for all $v \in T_i$, with $\partial U_i(v|Q, P)$ denoting the subdifferential of $U_i(\cdot|Q, P)$ at v. For any $t_i, t_i' \in T_i$, by definition of $\partial U_i(t_i'|Q, P)$ we have

$$U_i(t_i|Q, P) \geq U_i(t_i'|Q, P) + (t_i - t_i') \cdot Q_i(t_i').$$

Plug (3) into this inequality to see that

$$t_i \cdot Q_i(t_i) - P_i(t_i) \geq t_i \cdot Q_i(t_i') - P_i(t_i').$$

This true for all i, t_i and t_i', we have shown that (Q, P) is IC. ■

Convexity of the Utility Possibility Set For each $i \in \{1, \ldots, n\}$ denote $C(T_i)$ for the space of continuous real functions defined on the compact cube T_i, with the maximum 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

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

By Lemma 1, $(U_i(\cdot \mid Q, P))_{i=1}^n \in \mathscr{C}$ for any IC reduced form (Q, P).

Lemma 2 $\mathbb{U}$ *is convex.*

Proof Pick any $(W_i^1)_{i=1}^n, (W_i^2)_{i=1}^n \in \mathbb{U}$. By definition of $\mathbb{U}$, there exist reduced forms $(Q_i^1, P_i^1)_{i=1}^n$ and $(Q_i^2, P_i^2)_{i=1}^n$ that are each IC, IR and XF and, for any $i \in \{1, \ldots, n\}$, $k \in \{1, 2\}$, and $t_i \in T_i$,

$$W_i^k(t_i) \leq t_i \cdot Q_i^k(t_i) - P_i^k(t_i). \tag{7}$$

The IR condition means that, for any i, k, and t_i,

$$0 \leq t_i \cdot Q_i^k(t_i) - P_i^k(t_i). \tag{8}$$

By Lemma 1 and definition (3), the IC condition means, with the shorthand

$$U_i^k := U_i(\cdot | Q^k, P^k),$$

that U_i^k is convex on T_i for each i and k, and

$$U_i^k(t_i) = U_i^k(a_i) + \int_{a_i}^{t_i} Q_i^k(s) \cdot ds \tag{9}$$

for any $t_i \in T_i$. The XF condition means that there exist DRMs $(q_i^1, p_i^1)_{i=1}^n$ and $(q_i^2, p_i^2)_{i=1}^n$ such that, for any $k \in \{1, 2\}$, any $t \in T$, any $i \in \{1, \ldots, n\}$ and any $t_i \in T_i$,

$$\begin{aligned}
(q_i^k(t))_{i=1}^n &\in X, && (10)\\
\sum_{i=1}^{n} p_i^k(t) &\in Y((q_i^k(t))_{i=1}^n), && (11)\\
Q_i^k(t_i) &= \int_{T_{-i}} q_i^k(t_i, t_{-i}) dF_{-i}(t_{-i}), && (12)\\
P_i^k(t_i) &= \int_{T_{-i}} p_i^k(t_i, t_{-i}) dF_{-i}(t_{-i}). && (13)
\end{aligned}$$

Pick any $\gamma \in [0,1]$. We shall show that the vector $(W_i)_{i=1}^n$ defined by

$$W_i := \gamma W_i^1 + (1-\gamma) W_i^2$$

for each i belongs to $\mathbb{U}$. To that end, define, for any i, any $t \in T$, and any $t_i \in T_i$,

$$\begin{aligned}
q_i(t) &:= \gamma q_i^1(t) + (1-\gamma) q_i^2(t), && (14)\\
p_i(t) &:= \gamma p_i^1(t) + (1-\gamma) p_i^2(t), && (15)\\
Q_i(t_i) &:= \int_{T_{-i}} q_i(t_i, t_{-i}) dF_{-i}(t_{-i}), &&\\
P_i(t_i) &:= \int_{T_{-i}} p_i(t_i, t_{-i}) dF_{-i}(t_{-i}), &&\\
U_i(t_i|Q,P) &:= t_i \cdot Q_i(t_i) - P_i(t_i). && (16)
\end{aligned}$$

By these definitions, for each i we have

$$\begin{aligned}
Q_i &= \gamma Q_i^1 + (1-\gamma) Q_i^2, && (17)\\
P_i &= \gamma P_i^1 + (1-\gamma) P_i^2, && (18)\\
U_i(\cdot|Q,P) &= \gamma U_i^1 + (1-\gamma) U_i^2. && (19)
\end{aligned}$$

It suffices to prove that the reduced form $(Q,P) := (Q_i, P_i)_{i=1}^n$ is IC, IR and XF, for then

$$W_i = \gamma W_i^1 + (1-\gamma) W_i^2 \leq \gamma U_i^1 + (1-\gamma) U_i^2 = U_i(\cdot|Q,P)$$

for all i. This, coupled with the fact that W_i is continuous on T_i (since W_i^1 and W_i^2 are each continuous by definition of $\mathbb{U}$), implies that $(W_i)_{i=1}^n \in \mathbb{U}$.

Thus we shall complete the proof by showing that $(Q_i, P_i)_{i=1}^n$ is IC, IR and XF. First, IR follows directly from (8) and (19). Second, we show XF. Since X is assumed convex, (10) and (14) together imply $(q_i(t))_{i=1}^n \in X$ for any $t \in T$. By the assumption (2) of Y, (10) implies that, for any $t \in T$,

$$\begin{aligned}
\gamma Y\left((q_i^1(t))_{i=1}^n\right) + (1-\gamma) Y\left((q_i^2(t))_{i=1}^n\right) &\subseteq Y\left(\gamma (q_i^1(t))_{i=1}^n + (1-\gamma)(q_i^2(t))_{i=1}^n\right)\\
&= Y\left((q_i(t))_{i=1}^n\right),
\end{aligned}$$

with the second line due to (14). It then follows from (11) and (15) that

$$\begin{aligned}
\sum_i p_i(t) = \gamma \sum_i p_i^1(t) + (1-\gamma) \sum_i p_i^2(t) &\in \gamma Y\left((q_i^1(t))_{i=1}^n\right) + (1-\gamma) Y\left((q_i^2(t))_{i=1}^n\right)\\
&\subseteq Y\left((q_i(t))_{i=1}^n\right)
\end{aligned}$$

for all $t \in T$. Thus, by the above definition of (Q_i, P_i), the reduced form is XF.

Third, we verify that (Q, P) is IC. By Lemma 1, it suffices to verify that, for any i, $U_i(\cdot|Q, P)$ is convex and (6) is satisfied. By (9), (17) and (19), (6) is satisfied. To show that $U_i(\cdot|Q, P)$ is convex, pick any $t_i', t_i'' \in T_i$ and any $\alpha \in [0, 1]$. Let $t_i := \alpha t_i' + (1-\alpha)t_i''$ and denote $U_i(\cdot) := U_i(\cdot|Q, P)$. We need only to show $U_i(t_i) \leq \alpha U_i(t_i') + (1-\alpha)U_i(t_i'')$:

$$\begin{aligned}
U_i(t_i) &= t_i \cdot \left(\gamma Q_i^1(t_i) + (1-\gamma)Q_i^2(t_i)\right) - \gamma P_i^1(t_i) - (1-\gamma)P_i^2(t_i) \\
&= \gamma \left(t_i \cdot Q_i^1(t_i) - P_i^1(t_i)\right) + (1-\gamma)\left(t_i \cdot Q_i^2(t_i) - P_i^2(t_i)\right) \\
&= \gamma U_i^1\left(\alpha t_i' + (1-\alpha)t_i''\right) + (1-\gamma)U_i^2\left(\alpha t_i' + (1-\alpha)t_i''\right) \\
&\leq \gamma \left(\alpha U_i^1(t_i') + (1-\alpha)U_i^1(t_i'')\right) + (1-\gamma)\left(\alpha U_i^2(t_i') + (1-\alpha)U_i^2(t_i'')\right) \\
&= \alpha \left(\gamma U_i^1(t_i') + (1-\gamma)U_i^2(t_i')\right) + (1-\alpha)\left(\gamma U_i^1(t_i'') + (1-\gamma)U_i^2(t_i'')\right) \\
&= \alpha U_i(t_i') + (1-\alpha)U_i(t_i''),
\end{aligned}$$

where the first line follows from (16), (17) and (18), the fourth line (inequality) from the fact that U_i^1 and U_i^2 are each convex, and the last line from (19). Thus, U_i is convex. ■

Separating Hyperplane Pick any IIE reduced form (Q^*, P^*). Denote $u_i^* := U_i(\cdot \mid Q^*, P^*)$ for each i. Then $(u_i^*)_{i=1}^n \in \mathbb{U}$. Let

$$\mathbb{V}((u_i^*)_{i=1}^n) := \left\{ (u_i)_{i=1}^n \in \mathscr{C} \;\middle|\; \begin{array}{l} \forall i \left[u_i \geq u_i^* \text{ a.e. } [F_i] \text{ on } T_i\right]; \\ \exists i \left[u_i > u_i^* \text{ on } S_i \subseteq T_i; F_i\text{-measure of } S_i \text{ is} > 0\right] \end{array} \right\}.$$

Lemma 3 *There exists a continuous linear functional ϕ on $\mathscr{C}$, not identically zero, such that for all $(u_i)_{i=1}^n \in \mathbb{U}$,*

$$\phi\left((u_i)_{i=1}^n\right) \leq \phi\left((u_i^*)_{i=1}^n\right). \tag{20}$$

Proof First, $\mathbb{U}$ is convex by Lemma 2, and $\mathbb{V}((u_i^*)_{i=1}^n)$ convex by its definition. Second, $\mathbb{U}$ contains an interior point: Consider a mechanism A whose outcome is always to have $x_i = \mathbf{0}$ and $y_i = 0$ for all i and all j. By Assumption 1, this mechanism is ex post feasible. It is obviously IC and IR and it gives every type of every player zero payoff. This payoff profile is an interior point of $\mathbb{U}$. To see that, consider another mechanism B whose outcome is always the $(x_i^0, y_i^0)_{i=1}^n$ specified in Assumption 2. By that assumption, this mechanism is XF and gives every type of every player strictly positive payoff. Clearly, it is also IC and IR. For each player i, with T_i compact,

$$u_i^0 := \min_{t_i \in T_i} x_i^0 \cdot t_i - y_i^0 > 0.$$

Now consider a third mechanism C that, for any profile of messages from the players, carries out mechanism A with probability $1-\epsilon$, and mechanism B with probability ϵ $(0 < \epsilon < 1)$. Clearly mechanism C is also IC, IR and XF, and the profile of expected payoff functions it generates is larger than that generated by mechanism A in every dimension by at least ϵu_i^0. Since this is true for all $\epsilon \in (0,1)$, the zero payoff profile generated by the mechanism A 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, there exists an IC, IR and XF reduced form that generates a payoff vector that is at least as large as this interior point in all dimensions. Then, by definition of $\mathbb{V}\left((u_i^*)_{i=1}^n\right)$, this reduced form interim Pareto dominates $(u_i^*)_{i=1}^n$, contradicting the premise that the latter is IIE.

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 $(\hat{u}_i)_{i=1}^n \in \mathbb{V}((u_i^*)_{i=1}^n)$,

$$\phi\left((u_i)_{i=1}^n\right) \leq w \leq \phi\left((\hat{u}_i)_{i=1}^n\right). \tag{21}$$

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 (21) to obtain (20). ■

Representation 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, ϕ_i 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,

$$\phi\left((u_i)_{i=1}^n\right) = \sum_{i=1}^n \phi_i(u_i) \tag{22}$$

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 $C(T_i)$.

Lemma 4 *For each $i \in \{1, \ldots, n\}$, ϕ_i is positive.*[4]

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 $\mathbb{U}$, and by Lemma 3 we derive a contradiction:

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

[4] 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 any i, since ϕ_i is a positive linear functional on $C(T_i)$, with $T_i = \prod_{j=1}^{m}[a_{ij}, b_{ij}]$ a compact Hausdorff space, the Riesz-Markov theorem (Royden and Fitzpatrick [13, p458]) implies that there exists a unique Radon measure Λ_i on the Borel σ-algebra $\mathscr{B}(T_i)$ associated with the Euclidean topology on T_i such that

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

for all $u_i \in C(T_i)$. By definition of Radon measures, $\Lambda_i \geq 0$. It follows that $\Lambda_i(S) > 0$ for some measurable subset S of T_i and some i: Otherwise $\Lambda_i = 0$ for all i, hence ϕ is identically zero on $\mathscr{C}$, contradicting Lemma 3. This, combined with (20) and (22), delivers the representation theorem.

5 Conclusion

Proving the representation theorem in a general model, this note provides a foundation for the application of the mechanism design techniques to IIE on a wide variety of environments, including multidimensional types and countervailing incentives. The theorem is still relevant if one cares only about the positive analysis on such environments, because, as explained by Ledyard and Palfrey [7], the theorem implies that there is no loss of generality from the positive perspective to consider the normative notion of a social planner designing an optimal mechanism given some welfare distribution across player-types.

An open question is whether the converse of the representation is true in general. Pérez-Nievas [10] provides an argument for the converse based on one-dimensional types. An argument for any multidimensional-type case has yet to be found.

References

[1] Yeon-Koo Che, Jinwood Kim, and Konrad Mierendorff. Generalized reduced-form auctions: A network-flow approach. *Econometrica*, 81(6):2487–2520, 2013. 2

[2] Piotr Dworczak, Scott Duke Kominers, and Mohammad Akbarpour. Redistribution through markets. Mimeo, March 28, 2019. 1, 2

[3] Thomas Gresik. Incentive-efficient equilibria of two-party sealed-bid bargaining games. *Journal of Economic Theory*, 68:26–48, 1996. 1

[4] Bengt Holmström and Roger Myerson. Efficient and durable decision rules with incomplete information. *Econometrica*, 51(6):1799–1820, 1983. 1

[5] Mingshi Kang and Charles Z. Zheng. Necessity of Auctions for Redistributive Optimality. Mimeo, April 2019. 2, 1

[6] Didier Laussel and Thomas Palfrey. Efficient equilibria in the voluntary contributions mechanism with private information. *Journal of Public Economic Theory*, 5:449–478, 2003. 1

[7] John Ledyard and Thomas Palfrey. A characterization of interim efficiency with public goods. *Econometrica*, 67:435–448, 1999. 1, 2, 5

[8] John Ledyard and Thomas Palfrey. A general characterization of interim efficient mechanisms for independent linear environments. *Journal of Economic Theory*, 133:441–466, 2007. 1

[9] R. Preston McAfee and John McMillan. Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46:335–354, 1988. 1

[10] Mikel Pérez-Nievas. Interim efficient allocation mechanisms. Working Paper 00-20, Departmento de Economia, Universidad Carlos III de Madrid, February 2000. 1, 5

[11] Jean-Charles Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16:191–200, 1987. 4

[12] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970. 4

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

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

[15] Robert Wilson. Incentive efficiency of double auctions. *Econometrica*, 53:1101–1105, 1985. 1