Incentive Compatible Transfers

Nenad Kos *
Dept. of Economics, IGIER
Bocconi University

Matthias Messner †
Dept. of Economics, IGIER
Bocconi University

April 16, 2009

PRELIMINARY AND INCOMPLETE

First Version: December 2008

Abstract

We characterize the boundaries of the set of transfers which implement a given allocation rule (extremal transfers) in a quasi-linear environment. On the one hand doing so allows us to generalize and unify several existing results of the mechanism design literature and on the other it allows us to obtain novel results. More specifically, our approach delivers a characterization result for the set of implementable allocation rules. Second, we obtain an exact characterization of the allocation rules which satisfy Revenue Equivalence. Third, we show that in problems of optimal mechanism design it is often without loss of generality for the designer to restrict his attention to the extremal transfer rules. That is, we show how mechanism design problems can be handled even when the Revenue Equivalence principle does not apply.

After proving our main results for general environments in the first part of the paper, we consider two specific and very commonly studied settings in the second part. Namely, a single unit auction framework and a bilateral trade setup. We use the first environment to show that our characterization of the extremal transfers takes a particularly convenient form when types are unidimensional. In the bilateral trade setting our tools allow us to give a simple characterization of the conditions under which agents will voluntarily engage in efficient budget-balanced trade which applies whether or not Revenue Equivalence holds.

JEL Code: C72,D44,D82.
Keywords: Incentive Compatibility, Revenue Equivalence, Mechanism Design

*Department of Economics and IGIER, Bocconi University, Via Roentgen 1, I-20136 Milan, e-mail: nenad.kos@unibocconi.it
†Department of Economics and IGIER, Bocconi University, Via Roentgen 1, I-20136 Milan, e-mail: matthias.messner@unibocconi.it

1 Introduction

We present a new approach to mechanism design problems in quasi-linear environments. While our approach applies also to multi agent settings it is most conveniently described for the single agent case. We build our analysis on a characterization of the boundaries of the transfer schemes which implement a given allocation rule. More precisely, as a starting point we describe for any given allocation rule and any given type of the agent (henceforth to be called reference type) an upper ($\bar{t}$) and a lower bound ($\underline{t}$) of the set of (expected) transfer schemes which i) are incentive compatible and which ii) vanish at the reference type. Most importantly, we show that when the allocation rule is implementable these bounds themselves constitute incentive compatible transfers. Thus they represent the transfer payments which, respectively, maximize and minimize the (expected) payments which can be extracted from the agent subject to the constraint that the reference type's transfer is zero.

Based on the characterization of the extremal transfers $\bar{t}$ and $\underline{t}$ we are able to derive three important results. First, it allows us to formulate necessary and sufficient conditions for the implementability of a given social choice function. In particular, an allocation rule is implementable if and only if $\underline{t} \leq \bar{t}$ for every type of the agent. Second, we obtain a characterization of the set of allocation rules which satisfy Revenue Equivalence. An allocation rule satisfies Revenue Equivalence if and only if the associated extremal transfers coincide, i.e. if and only if $\underline{t} = \bar{t}$for all types.

Finally, we show that for a mechanism designer it is often without loss of generality if he restricts his attention to the extremal transfers or translations thereof. In particular, we show that this holds whenever the following conditions are satisfied: i) the designer maximizes an objective which is weakly increasing in expected transfers; ii) there exists a type (low-payoff type) of the agent whose payoff from any decision is smaller than the corresponding payoff of any other type. Before we comment on the meaning of these conditions it is important to underline the relevance of this third result. Since the extremal transfers can be expressed in terms of the allocation rule, this result essentially tells us that a mechanism design can often be reduced - and thus be simplified - by the dimension of the transfers. In particular, in order for this simplification to be warranted it is not necessary that Revenue Equivalence obtains.

The role of the first condition is rather intuitive. If the planner's objective is increasing in (expected) transfers then he will want to choose them as large as possible. That is, there is a tendency to choosing an extremal transfer associated with the optimal allocation rule. The only reason for not doing so might be that the extremal transfers do not allow to satisfy the agent's individual rationality constraint at all types. This cannot happen if we know that it is always (i.e. independently of the allocation rule) the same type of the agent who is 'pivotal' for individual rationality (in the sense that a violation of individual rationality anywhere, together with incentive compatibility, implies a violation also for this type), for in that case the planner only needs to make sure that in calculating the extremal transfers he picks this type as reference type. The second condition imposes existence of such a crucial type. That is, it implies that whenever in an incentive compatible mechanism the individual rationality constraint of the low-payoff type is satisfied then so are all other types' individual rationality constraints.

Notice that both conditions we are imposing are satisfied for the most commonly studied mechanism design problems. As for the planner's objective it is standard to assume that he either maximizes his revenue or that he maximizes some welfare measure. Both of these objectives are linear in expected transfers. The existence of a low-payoff type is likewise satisfied in most applications for the agent's payoff is often assumed to be monotonic in his own type.

Previously described characterization results are derived for very general frameworks. The only specific assumptions which we need to impose are (measurability) assumptions which guarantee that all expectations which we are taking are well defined. While this generality is certainly a strength of our results, it comes at the cost that our characterizations remain somewhat abstract. In the second part of the paper we therefore explore the implications of our approach in more specific types of frameworks. In particular, we apply our tools to two of the most commonly considered settings, namely a single unit auction framework and a bilateral trade setup.

Our analysis of the auction framework unifies different approaches used in the literature to treat the case of discrete values and the case of values being distributed with a positive density over some interval. The characterization of the extremal payoffs takes a particularly simple form in the case of type spaces being subsets of $\mathbb{R}$. It is well known since the seminal paper by Myerson (1981), that in auction settings where bidders' type spaces are connected, (expected) transfers can be expressed as integrals over types with respect to the (expected) allocation rule. In the auction framework which we consider we allow type spaces to be disconnected. This notwithstanding we are able to show that transfers (including the extremal ones) can be expressed in a similar integral form. In particular, with the help of our earlier developed tools we show that one can define a class of extensions of the allocation rule to the smallest interval containing the type space such that integrating over types with respect to any such extension delivers an incentive compatible transfer. An interesting side product of our analysis is an alternative proof of Myerson's revenue equivalence result.

The study of the auction framework is also instructive for it allows us to assign a clear intuitive interpretation to the extremal transfers. In particular, if type spaces are finite and $\underline{t}$ and $\bar{t}$ are defined with respect to the lowest type, then $\underline{t}$ ($\bar{t}$) corresponds to the transfer rule for which all upward (downward) adjacent incentive constraints are binding.

The bilateral trade framework we analyze in Section 5 is standard, aside that we do not impose connectedness of the valuation spaces, as introduced by Myerson and Satterthwaite (1983). Myerson and Satterthwaite (1983) have shown that with connected type spaces there is no efficient trading mechanism in which all types of both agents would voluntary participate and whose transfers are ex post budget balanced. On the other hand they also show by means of an example with finite types, that in some disconnected settings this impossibility result does not hold.

Here we generalize their results by exactly characterizing the conditions under which voluntary trade is feasible. More precisely, we show that a trading rule can be implemented by ex post budget balanced transfers if and only if the buyer's highest transfer defined with respect to his lowest type ($\bar{t}_B$) and the seller's highest transfer defined with respect to his highest type ($\bar{t}_S$), run an ex ante budget surplus. The intuition underlying this result is straightforward. The seller's highest type and the buyer's lowest type are the crucial types in terms of individual rationality. If the highest transfers which can be extracted (or more appropriately in the case of the seller: the lowest payments which have to be made to him) from these types are not sufficient to guarantee a budget surplus in an ex ante sense then it cannot be possible to guarantee budget balancedness ex post. As an immediate corollary to this result we obtain an exact characterization of the conditions for the existence of an efficient ex post budget balanced, individually rational and incentive compatible mechanism.

Related Literature

The literature on mechanism design in quasi-linear environments is too vast to be detailedly discussed here. In what follows we will limit ourselves to a rather brief and schematic overview. We divide discussion of the literature in accordance with our results.

Implementability: Among the papers which provide a characterization of the set of implementable allocation rules, the most closely related to ours is Rochet (1987). He provides a 'no negative cycle'-condition which he shows to be necessary and sufficient for implementation in dominant strategies. Our characterization provides the missing economic intuition by recasting the problem in terms of minimal and maximal transfers.

Revenue Equivalence: The literature on Revenue Equivalence can be divided into two main lines. In a first and older strand of the literature Revenue Equivalence results are obtained in combination with formulas for the transfers, that is, formulas which express the transfers in terms of the allocation rule which they are supposed to implement. The most notable early proponents of such results were Vickrey (1961) and Myerson (1981). For a more recent paper in this line of the literature see Milgrom and Segal (2002).

In more recent years there have been a number of papers (see for instance Chung and Olszewski (2007) and Heydenreich et al. (2009)) which focus their attention exclusively on the characterization of the conditions under which Revenue Equivalence obtains, without providing an explicit description of transfers. From the perspective of the mechanism designer the results proposed in the earlier strand of the literature are more useful. Since this literature provides explicit formulas which describe the transfers in terms of the choice rule to be implemented, it enables the mechanism designer to eliminate the transfers altogether from his problem. Of course, doing so amounts to a major simplification of the mechanism design problem for it reduces it by the dimension of transfers.

Characterization of the extremal transfers: Chung and Olszewski (2007) show for a very general setup, that the incentive compatible transfer that extracts the most money from the agent exists. We go beyond this result by showing that i) also a (pointwise) minimal transfer scheme exists and by ii) providing a characterization of both of these extremal transfer. We also show that a mechanism designer whose objective function is monotonic in the interim transfers can restrict his attention to the extremal transfers only.

The auction framework: Our work is related to a large number of papers which consider settings with independent private values in $\mathbb{R}$ which do not satisfy the classical assumptions under which Revenue Equivalence holds (e.g. discrete types). In particular, our paper unifies the different approaches that are used to solve the problems with discrete types and the problems with a continuum of types. For a recent paper solving for an optimal auction without requiring connectedness of the underlying valuation spaces see Skreta (2006). Skreta (2006) solves a revenue maximization problem with such characteristics by showing that it is equivalent to a problem where all involved functions (utility functions, allocation rules, transfer schemes) and the definition of the incentive compatibility constraint are extended to the convex hull of the set of valuations. Our characterization does not impose incentive compatibility outside of the valuation space. Moreover, the class of problems to which our tools can be applied is far larger than the set of revenue maximization problems considered in Skreta (2006).

The bilateral trade setting: Our analysis of bilateral trade builds on the setup first developed in Myerson and Satterthwaite (1983). Moreover, our work is also closely related to Krishna and Perry (1998), Williams (1999) and Kos and Manea (2009). Krishna and Perry (1998) and Williams (1999) provide a provide an alternative proof and an interesting intuition for Myerson and Satterthwaite's (1983) central result which shows nonexistence of an ex post efficient, ex post budget balanced, incentive compatible and individually rational mechanism. In doing so they proceed in two steps. They first show that a mechanism of the before mentioned type exists if and only if a Vickrey-Clark-Groves (henceforth, VCG) mechanism with an appropriately chosen constant runs an expected budget surplus. In a second step they then prove that their VCG mechanism always runs an expected budget deficit. In terms of the framework both papers stick to Myerson and Satterthwaite's original assumption on the distribution of the valuation types. That is, in both it is assumed that i) type sets are intervals and ii) the distributions of the types admit density functions which are strictly positive everywhere on these intervals. Kos and Manea (2009) instead consider settings with finite valuation spaces. Their characterization of the type distributions for which an incentive compatible, individually rational, ex-post efficient and budget balanced trading mechanism exists, is similar to the one provided in this paper. Notice though that our characterization is not restricted to the case of finite valuation sets.

2 Framework

We start with the description of the quasi-linear, single agent framework which will be at the center of our attention throughout most of the following section. In terms of notation the extension to a multiple agent set up is straightforward. We will shortly comment on it at the end of Section 3 where we show how our results for the single agent case carry over to multi agent environments.

In what follows we consider a situation where a planner/mechanism designer faces a single agent. An allocation is a pair $(x, \tau) \in X \times \mathbb{R}$, where x is a (non-monetary) choice variable while τ represents a monetary transfer (to be paid by the agent). With a slight abuse of terminology we will often use the term 'allocation' to indicate the non-monetary component x of an allocation (x, τ). The agent's preferences over $X \times \mathbb{R}$ depend on his (preference-) type $v \in V$. We assume that they are quasi-linear for each $v \in V$. That is, type v's payoff at the allocation (x, τ) can be written in the form

$$u(x, v) - \tau,$$

where $u : X \times V \to \mathbb{R}$. It is important to point out that except for the quasi-linearity of preferences, our framework is very general. In particular, for the time being we do not impose any assumptions at all on the set of decisions X, the set of preference types V or the function u.

We refrain from giving a definition of general mechanisms for this environment since by the *Revelation Principle* we know that it is without loss of generality to consider only direct mechanisms. A *direct mechanism* is a pair of functions, (q, t), where

$$q : V \to X,$$ [1]

and

$$t : V \to \mathbb{R}.$$

The function q, to which we will refer as *allocation rule* or *choice function*, associates with each type of the agent $v \in V$ a decision $q(v) \in X$. Similarly, the so-called *transfer rule* t specifies for each type of the agent a transfer to be paid by this type.

[1]The restriction to deterministic mechanisms is made for notational convenience only. All our results extend to the case of non-deterministic allocation rules.

3 Incentive Compatible Transfers

The following definition recalls some standard terminology and concepts which we will frequently use throughout the paper.

Definition 1 (Incentive compatibility, Implementability, Revenue Equivalence)**.**

i) *A direct mechanism* (q, t) *is called* **incentive compatible** *if*

$$u(q(v), v) - t(v) \geq u(q(v'), v) - t(v')$$

holds for all type pairs $v, v' \in V$*.*

ii) *A choice function* q *is said to be* **implementable** *if there exists a transfer scheme* t *such that* (q, t) *is incentive compatible. Moreover, if* (q, t) *is incentive compatible then we say that the transfer rule* t **implements** *the choice function* q*.*

iii) *A social choice function* q *satisfies* **Revenue Equivalence** *if it is implementable and for any two incentive compatible mechanisms* (q, t) *and* (q, t') *there exists a constant* $c \in \mathbb{R}$ *such that*

$$t(v) = t'(v) + c,$$

for every $v \in V$*.*

It is straightforward to verify that the above definition of Revenue Equivalence is equivalent to the following one. A social choice function q satisfies Revenue Equivalence if it is implementable and for any $\hat{v} \in V$ and any two transfer schemes t and t' which implement q we have

$$t(v) - t(\hat{v}) = t'(v) - t'(\hat{v})$$

for all $v \in V$. We will refer to the function $t(\cdot) - t(\hat{v})$ as the *transfer differential* at the type $\hat{v}$. Whenever the reference type $\hat{v}$ is clear from the context we will not refer to it when using the term transfer differential.

We are now ready to introduce the central concepts for our analysis. For any pair of types $v, \hat{v} \in V$ let $S(v; \hat{v})$ denote the set of all finite sequences in the type space V which start at $\hat{v}$ and end at v. That is, a finite sequence $\{v^j\}_{j=0}^n$ belongs to $S(v; \hat{v})$ if and only if $v^j \in V$ for all $j = 0, \ldots, n$, $v^0 = \hat{v}$ and $v^n = v$. For any choice rule q and any $\hat{v} \in V$ we define the two functions $\bar{t}(\cdot; \hat{v}, q), \underline{t}(\cdot; \hat{v}, q) : V \to \overline{\mathbb{R}}$ as follows

$$\bar{t}(v; \hat{v}, q) \equiv \inf_{\{v^j\} \in S(v; \hat{v})} \sum_{j=1}^{n} \left[u(q(v^j), v^j) - u(q(v^{j-1}), v^j) \right], \qquad \text{and}$$

$$\underline{t}(v; \hat{v}, q) \equiv \sup_{\{v^j\} \in S(v; \hat{v})} \sum_{j=1}^{n} \left[u(q(v^j), v^{j-1}) - u(q(v^{j-1}), v^{j-1}) \right].$$

Finally, we write $I(v;\hat{v},q)$ for the interval $[\underline{t}(v;\hat{v},q),\bar{t}(v;\hat{v},q)]$. We say that $I(v;\hat{v},q)=\emptyset$ if $\bar{t}(v;\hat{v},q)<\underline{t}(v;\hat{v},q)$.

For the moment we postpone commenting on the above defined expressions since their meaning and interpretation will immediately become clear from the following results. At this point we limit ourself to making the following observation. First, notice that the functions $\underline{t}$ and $\bar{t}$ are not necessarily real valued. We have specified the extended real numbers $\overline{\mathbb{R}}$ as their range since we cannot rule out the possibility that for some $(v,\hat{v},q)$ they take on the value $-\infty$ or ∞. On the other hand, it is also easily seen that for each triple $(v,\hat{v},q)$, we have $\bar{t}(v;\hat{v},q)\leq u(q(v),v)-u(q(\hat{v}),v)<\infty$ and $\underline{t}(v;\hat{v},q)\geq u(q(v),\hat{v},q)-u(q(\hat{v}),\hat{v})>-\infty$. For later reference, it is important to point out that whenever $\bar{t}$ ($\underline{t}$) is finite then it vanishes at the type $\hat{v}$. [2] For the remainder of our analysis we will refer to this type as *reference* or *anchor* type.

Lemma 1. *Let (q,t) be an incentive compatible mechanism. Then, for any given $\hat{v}\in V$ we have*

$$\underline{t}(v;\hat{v},q)\leq t(v)-t(\hat{v})\leq\bar{t}(v;\hat{v},q)$$

for every $v\in V$.

Proof. Incentive compatibility implies

$$u(q(v'),v'')-u(q(v''),v'')\leq t(v')-t(v'')\leq u(q(v'),v')-u(q(v''),v') \tag{1}$$

for any pair $v',v''\in V$. Fix a $v\in V$ and let $\{v^j\}_{j=0}^n$ be a sequence in $S(v;\hat{v})$. Consider the first inequality in (1). Adding the inequality for all pairs of consecutive elements of the sequence, (v^{j-1},v^j), yields

$$\sum_{j=1}^{n}\left[u(q(v^{j-1}),v^j)-u(q(v^j),v^j)\right]\leq\sum_{j=1}^{n}\left[t(v^j)-t(v^{j-1})\right]=t(v)-t(\underline{v}).$$

Taking the supremum over all sequences in $S(v;\hat{v})$ yields $\underline{t}(v;\hat{v},q)\leq t(v)-t(\hat{v})$. The second inequality is obtained in an analogous fashion. □

Lemma 1 tells us that $\underline{t}(\cdot;\hat{v},q)$ and $\bar{t}(\cdot;\hat{v},q)$ represent respectively, a lower and an upper bound of the transfer differential with respect to $\hat{v}$. Put differently, $\underline{t}(\cdot;\hat{v},q)$ and $\bar{t}(\cdot;\hat{v},q)$ bound the set of all transfer schemes which implement q and require the anchor type $\hat{v}$ to pay a zero transfer. Our next result shows that these bounds are actually tight. That is, $\underline{t}$ and $\bar{t}$ themselves constitute incentive compatible schemes (which - by definition - vanish at the anchor type $\hat{v}$).

[2] Clearly $\bar{t}(\hat{v};\hat{v},q)\leq 0$, since the sequence $v_0=\hat{v},v_1=\hat{v}$ is available. If $\bar{t}(\hat{v};\hat{v},q)<0$, then there would exist a sequence (cycle) starting and ending in $\hat{v}$ whose sum would be negative. But then finite sequences could be constructed that incorporate this cycle arbitrarily many times making $\bar{t}(v;\hat{v},q)$ equal to $-\infty$ for all v.

Lemma 2. *Let q be such that for some $\hat{v} \in V$ we have $I(v; \hat{v}, q) \neq \emptyset$ for all $v \in V$. Then both $(q, \underline{t}(\cdot; \hat{v}, q))$ and $(q, \bar{t}(\cdot; \hat{v}, q))$ are incentive compatible mechanisms.*

Proof. See the Appendix. □

The preceding two lemmata constitute the basis for the following theorem.

Theorem 1. *For any $\hat{v} \in V$ the following statements are true.*

i) An allocation rule q is implementable if and only if $I(v; \hat{v}, q) \neq \emptyset$ for every $v \in V$.

ii) An allocation rule q satisfies Revenue Equivalence if and only if for all $v \in V$, $I(v; \hat{v}, q)$ is a singleton.

iii) Let q be an implementable allocation rule and let $\hat{v} \in V$. If t is a transfer which implements q, and $t(\hat{v}) = 0$, then

$$\underline{t}(v; \hat{v}, q) \leq t(v) \leq \bar{t}(v; \hat{v}, q),$$

for every $v \in V$.

Proof. See the Appendix. □

Theorem 1 provides a novel characterization of the set of incentive compatible transfers. Moreover, it sheds light on several existing results from the literature on mechanism design. The first statement of the theorem bears a close resemblance to the characterization of implementability in Rochet (1987). According to Rochet's characterization an allocation rule q is implementable if and only if there does not exist a finite sequence of valuations $\{v^j\}_{j=0}^n \subset V$, which satisfies the conditions i) $v^0 = v^n$ (i.e. the sequence forms a cycle) and ii) $\sum_{j=1}^n \big[u(q(v^j), v^j) - u(q(v^{j-1}), v^j)\big] < 0$.

The connection with our characterization is immediate. First, the existence of such a cycle implies that $\bar{t}(\cdot; \hat{v}, q) = -\infty$ and $\underline{t}(\cdot; \hat{v}, q) = \infty$, meaning that $I(\cdot; \hat{v}, q) = \emptyset$. In order to see this notice that if such a cycle exists then one can construct finite sequences $\{\tilde{v}^j\}_{j=0}^n$ which start at $\hat{v}$, end at v and which run multiple times through the cycle (forward or backward). The more often a sequences runs through the cycle in forward direction, the lower the value of the associated sum $\sum_{j=1}^n \big[u(q(v^j), v^j) - u(q(v^{j-1}), v^j)\big]$. Similarly, the more often a sequence contains the inverted version of the cycle, the higher will be the sum $\sum_{j=1}^n \big[u(q(v^j), v^{j-1}) - u(q(v^{j-1}), v^{j-1})\big]$. In fact, by including the cycle (or its inverse) sufficiently many times into a sequence one can reach an arbitrarily high (low) level.

Conversely, suppose we have $\underline{t}(v;\hat{v},q) \leq \bar{t}(v;\hat{v},q)$ for all $v \in V$. Then, in particular, $\underline{t}(\hat{v};\hat{v},q) \leq \bar{t}(\hat{v},\hat{v},q)$. Since for the constant sequences $\{v^j\}$, $v^j = \hat{v}$, $\forall j$, the sums in the definitions of the extremal transfers with anchor type $\hat{v}$ are zero, it follows that $\underline{t}(\hat{v};\hat{v},q) \geq 0$ and $\bar{t}(\hat{v};\hat{v},q) \leq 0$. Combining this with the previous observation we can conclude that whenever q is implementable $\underline{t}(\hat{v};\hat{v},q) = 0 = \bar{t}(\hat{v};\hat{v},q)$. Since $\bar{t}(\hat{v};\hat{v},q)$ is the lower bound of all values which can be written as a sum of the form $\sum_{j=1}^{n}\left[u(q(v^j),v^j) - u(q(v^{j-1}),v^j)\right]$ for some finite cycle which starts and ends at $\hat{v}$, the sum for no such cycle can be negative. But then there can be no negative-value cycle for any other anchor type either since we could always include (sufficiently many replica of) it into $\hat{v}$-based cycles in order to get a negative-value cycle at $\hat{v}$.

From the arguments which we have used in the previous two paragraphs it follows that our characterization of implementability in the first part of Theorem 1 can be formulated in the following slightly more compact way.

Corollary 1. *Fix $\hat{v} \in V$. An allocation rule q is implementable if and only if $\bar{t}(\hat{v};\hat{v},q) = 0$.*

Our characterization of Revenue Equivalence is formally equivalent to the one given in Heydenreich et al. (2009). The main advantage of our approach is that we formulate our result in terms of concepts (the extremal transfers) which allow for a very natural economic interpretation.

The most novel result of Theorem 1, is the one contained in statement iii). There we describe the boundaries of the set of transfer differentials for implementable allocation rules. Furthermore, by the preceding lemma we also know that these boundaries themselves constitute incentive compatible transfer rules. In fact, the two boundaries correspond to the transfers which extract the most money ($\bar{t}$) and the least money ($\underline{t}$), respectively, given that the transfer of the reference type $\hat{v}$ is zero. That is, any transfer rule which i) implements the given allocation rule and ii) requires the reference type $\hat{v}$ to pay a zero transfer is pointwise dominated by (dominates) the transfer scheme $\bar{t}$ ($\underline{t}$).

This result is of utmost relevance for mechanism design. Most of the mechanism design literature assumes that the objective of the designer is monotonic in transfers (welfare maximization, revenue maximization). But then imposing assumptions which deliver Revenue Equivalence is actually an overkill. In fact, in what follows we will argue that for a designer with preferences which are monotonic in (expected) transfers it is often without loss of generality to restrict his attention to the transfers $\underline{t}$ or $\bar{t}$ (or translated versions of them).

In order to make this point more precise it is convenient to consider a standard principal-agent set up. Suppose there is an uniformed principal who interacts with an agent who privately observes his type $v \in V$. v enters both players' payoffs and is distributed according to the distribution function F (which is commonly known). Notationally, we accommodate the second player by indexing all the relevant variables and functions by P (principal) and A (agent), respectively. The planner designs the mechanism with the objective to maximize a weighted sum of the agent's and the principal's expected utilities. That is, the payoff which he perceives at the mechanism (q, t_A, t_P) is

$$W(q, t_A, t_P) = \lambda_P \int [u_P(q(v), v) - t_P(v)] dF(v) + \lambda_A \int [u_A(q(v), v) - t_A(v)] dF(v),$$

where $\lambda_P \geq 0$ and $\lambda_A \geq 0$ are the welfare weights which the planner attaches to the principal and the agent, respectively.

In maximizing W the planner has to respect three types of constraints. First, he must choose such a mechanism that it is incentive compatible for the agent to reveal his type. Second, he must guarantee both individuals a minimal payoff level at the interim stage in order to make sure that both of them are willing to participate in the mechanism (individual rationality constraint). In particular, we assume that both individuals are willing to participate as long as their payoff is non-negative. Finally, the planner must also respect the budget constraint $t_A(v) + t_P(v) \geq 0$. Formally, we can write his problem as follows.

$$\begin{aligned} \sup_{(q,t) \in \mathcal{Q} \times \mathcal{T}} \quad & W(q, t_A, t_P) \\ \text{s.t.} \quad & u_A(q(v), v) - t_A(v) \geq u_A(q(v'), v) - t_A(v'), \quad \forall\, v, v' \in V \\ & u_A(q(v), v) - t_A(v) \geq 0, \quad \forall\, v \in V, \\ & \int [u_P(q(v), v) - t_P(v)] dF(v) \geq 0 \\ & t_A(v) + t_P(v) \geq 0, \end{aligned} \tag{P1}$$

where $\mathcal{Q}$ and $\mathcal{T}$ are the sets of all allocation and transfer rules for which W is well defined (i.e. those rules for which the integral in the definition of W is well defined). The following proposition establishes that for a planner who puts a larger weight on the principal's welfare (i.e. $\lambda_P \geq \lambda_A$) solving problem (P1) is essentially equivalent (in a sense to be made more precise) to solving the following simpler problem (P2).

$$\begin{aligned} \sup_{(q,c) \in \mathcal{Q} \times \mathbb{R}} \quad & W(q, \bar{t}_A(v; \hat{v}, q) + c, -\bar{t}_A(v; \hat{v}, q) - c) \\ \text{s.t.} \quad & \underline{t}_A(v; \hat{v}, q) \leq \bar{t}_A(v; \hat{v}, q), \forall v \in V \\ & u_A(q(\hat{v}), \hat{v}) - \bar{t}_A(\hat{v}; \hat{v}, q) - c \geq 0, \forall v \in V \end{aligned} \tag{P2}$$

Proposition 1. *Suppose that there exists a $\hat{v} \in V$ such that $u_A(x, v) \geq u_A(x, \hat{v})$ for all $x \in X$ and for all $v \in V$. Also, assume that $\lambda_P \geq \lambda_A$ and that the problems (P1) and (P2) admit a solution. Then the values of the maximization problems (P1) and (P2) coincide. Moreover, if (q^*, c^*) solves (P2), then $(q^*, \bar{t}_A(\cdot; \hat{v}, q^*) + c^*, -\bar{t}_A(\cdot; \hat{v}, q^*) - c^*)$ solves (P1). Conversely, if (q^*, t_A^*, t_P^*) solves (P1) then there exists a c^* such that (q^*, c^*) solves (P2). Finally, if $\lambda_P > \lambda_A$ and (q^*, c^*) is a solution of (P2), then $(t_A^*, t_P^*) = (\bar{t}_A(\cdot; \hat{v}, q^*) + c^*, -\bar{t}_A(\cdot; \hat{v}, q^*) - c^*)$ is the unique transfer pair which together with q^* solves problem (P1).*

Proof. See the Appendix. □

In the above proposition the assumption on the agent's payoff function serves the purpose of simplifying the agent's individual rationality constraint. In combination with incentive compatibility it guarantees that individual rationality must hold everywhere if it holds at the type $\hat{v}$. Using type $\hat{v}$ as the anchor type one can thus make sure that $\bar{t}$ satisfies individual rationality.

In problems (P1) and (P2) it is assumed that the agent has an outside option which he values zero independently of his type. There is a rather straightforward way in which the above proposition can be generalized to environments with a type-dependent outside option. In such cases, the condition that there exists a type who assigns to any allocation a smaller value than any other type can be replaced with the condition that there exists a type, $\hat{v}$, whose 'net' valuation at any allocation is smaller than the corresponding 'net' valuations of all other types. By net valuation here we simply mean the difference between the agent's valuation, $u_A(x, v)$, and the value of the outside option at the type v.

The proposition only deals with the case where the planner puts a larger weight on the principal's payoff, i.e. $\lambda_P > \lambda_A$. The opposite strict inequality is easily dealt with, though not as interesting for us since the optimal mechanism solving the problem does not necessarily involve the extremal transfers. Indeed, when $\lambda_P < \lambda_A$ the problem (P1) can be shown to be equivalent to the problem in which the expectation of $\lambda_A(u_A(q(v), v) + u_P(q(v), v))$ is maximized subject to ex post budget balance, binding principal's IR constraint, agent's incentive compatibility and agent's individual rationality. It is not difficult to see that the set of the allocation rules that solve the problem is the same for all pairs of welfare weights, (λ_A, λ_P), such that $\lambda_A > \lambda_P$.

We lay no claims as to how easily $\underline{t}$ and $\bar{t}$ are computed. We will demonstrate, however, the use of the tools we have developed by applying them to a well known example from Holmstrom (1979). Before we proceed with this example, a brief comment on how to extend the model to multiple agents is in order.

Multiple agent settings: Extending our analysis to a framework with multiple agents whose valuations are independently distributed does not generate any major conceptual difficulties. The only difference between the single agent case which we have considered so far and a multi agent environment with independently distributed types lies in the fact that in the latter the agents' incentive constraints are defined in terms of agents' expected utility (where the expectation is taken with respect to the other agents' types). Consequently, the bounds on transfer differentials in the multi agent case are to be interpreted as bounds on (interim) expected transfers.

In order to see this, suppose there are I agents, indexed by $i = 1, \ldots, I$. Let V_i be the set of types of player i and write $V = \times_i V_i$ for the set of all possible type profiles. It is commonly known that player i's type is drawn from some distribution over V_i and that this distribution is stochastically independent of any other player's type distribution. Player i's payoff function is allowed to depend not only on his own type, v_i, but also on the type profile of all other players, v_{-i}. Thus we write $u_i(x, v) - \tau_i$ for the payoff of agent i when the type profile is v, the choice is x and agent i's transfer is τ_i.

In this multi agent environment a direct mechanism is a pair of functions (Q, T) which specify for each profile of valuations specify an allocation $(Q : V \to X)$ and a profile of transfers $(T : V \to \mathbb{R}^I)$, respectively. A mechanism (Q, T) is said to be (Bayesian) incentive compatible if for each player $i = 1, \ldots, I$ we have

$$E_{v_{-i}}[u_i(Q(v_i, v_{-i}), v_i, v_{-i}) - T_i(v_i, v_{-i})] \geq E_{v_{-i}}[u(Q(v'_i, v_{-i}), v_i, v_{-i}) - T_i(v'_i, v_{-i})] \qquad \forall\, v_i, v'_i \in V_i.$$

Letting $t_i(v_i) = E_{v_{-i}}[T_i(v_i, v_{v_{-i}}]$ and $\tilde{u}_i(v'_i, v_i; Q) = E_{v_{-i}}[u_i(Q(v'_i, v_{-i}), v_i, v_{-i})]$ this condition can be rewritten as

$$\tilde{u}_i(v_i, v_i; Q) - t_i(v_i) \geq \tilde{u}_i(v'_i, v_i; Q) - t_i(v'_i) \qquad \forall\, v_i, v'_i \in V_i.$$

This condition has the same form as the incentive constraint in the single agent case. Thus, using the functions $\tilde{u}_i$ we can define for each allocation rule Q and any $\hat{v}_i \in V_i$ upper and lower bounds on the (interim) *expected* transfer differential $t_i(\cdot) - t_i(\hat{v}_i)$. All results then follow by adding the appropriate quantifier for i.

As anticipated above, we now reconsider a much celebrated example first studied by Holmstrom (1979).[3]

Example 1 (Holmstrom (1979))**.** There are two agents with valuation sets $V_1 = V_2 = [0, 1]$. The choice set is $X = [0, 1]$. Players have quasi-linear preferences. Their utilities from the non-monetary component x of the social decision are described by the following two functions.

$$u_1(x, v_1) = \begin{cases} 0 & \text{if} \quad x < v_1 \\ -(x - v_1) & \text{if} \quad x \geq v_1 \end{cases} \qquad \text{and} \qquad u_2(x, v_2) = v_2 + \frac{1}{2}x.$$

For later reference notice that in this simple environment the unique efficient allocation at the type profile (v_1, v_2) is to set x equal to v_1. Hence, the unique efficient allocation rule, Q^*, is defined by $Q^*(v_1, v_2) = v_1$.

The objective of Holmstrom (1979) was to identify conditions under which Revenue Equivalence holds at the efficient allocation rule. The above example is proposed as an instance for an environment which fails to satisfy this condition. The reason for this failure lies of course in

[3]The purpose of this example is to put our machinery to work. We do not claim to be the first to point out how the extremal transfers are computed in this case. Indeed Ely (2001) shows how to do that. His treatment, however, applies only to the case where valuations are distributed over an interval with a positive density.

the fact that the utility function of player 1 is not smooth (differentiable) in the social decision at the efficient allocation.

Before we proceed to show how the boundaries of the set of transfers which implement Q^* can be described using our tools a remark on the role of Revenue Equivalence is in order. Holmstrom (1979) restricts his attention to the efficient allocation rule Q^*. This rule can be characterized without any reference to transfers whatsoever. Thus, the purpose of Revenue Equivalence in this case is not to simplify the problem of calculating the best allocation rule. Instead it is useful only in that it allows us to obtain in a simple way a full characterization of the set of transfers which implement Q^*: If Revenue Equivalence holds we immediately know that only Vickrey-Clark-Groves transfer schemes can do the job. But if the question is the one of characterizing the set of all incentive compatible mechanisms then that can be done also for his counterexample.

We fix as anchor types the two agents' lowest types, i.e. $\hat{v}_1 = \hat{v}_2 = 0$. Now consider first agent 1 and observe that since u_1 does not depend on v_2 we have $\tilde{u}_1(v_1', v_1; Q^*) = u_1(Q^*(v_1', v_2), v_1)$. Next notice that the difference $u_1(Q^*(v_1, v_2), v_1) - u_1(Q^*(v_1', v_2), v_1)$ vanishes if $v_1 \geq v_1'$. But since $\hat{v}_1 = 0$ this implies that $\bar{t}_1(v_1; \hat{v}_1, Q^*) = 0$ for every $v_1 \in V_1$. As for the lower bound $\bar{t}$ it is easily seen that for any sequence, $\{v_1^j\}$, starting at $\hat{v}_1 = 0$ and ending at v_1, the corresponding sum of the payoff differences $u_1(Q^*(v_1^j, v_2), v_1^{j-1}) - u_1(Q^*(v_1^{j-1}, v_2), v_1^{j-1}))$ is equal to $-v_1$. It thus immediately follows that $\underline{t}_1(v_1; \hat{v}_1, Q^*) = -v_1$ for every $v_1 \in V_1$. Summarizing, we conclude that any transfer rule t_1 for which player 1's incentive constraint holds, satisfies

$$-v_1 \leq t_1(v_1) - t_1(0) \leq 0,$$

for every $v_1 \in V_1$.

Next, consider agent 2 and observe that $\tilde{u}_2(Q^*(v_1, v_2), v_2) = v_2 + E(v_1)/2$. Since this is an expression which is linear in v_2 it is straightforward to see that we must have $\underline{t}_2(v_2; \hat{v}, Q^*) = \bar{t}_2(v_2; \hat{v}, Q) = v_2$. Therefore, any transfer rule t_2 for which player 2's incentive constraint holds satisfies

$$t_2(v_2) - t_2(0) = v_2,$$

for every $v_2 \in V_2$.

4 Auctions

In the previous section we have considered very general environments. As a consequence our characterization of the bounds for the transfer differentials has remained rather abstract. Here we consider a more specific class of mechanism design problems on which we impose stronger assumptions regarding the form of the type sets, the choice set X and the agents' payoff functions. In particular, we study a single unit auction framework with one-dimensional types and risk neutral bidders. It is well-known since Myerson (1981) that if one assumes that type sets are intervals and that the type distributions allow for density functions which are strictly positive everywhere, then Revenue Equivalence holds and the transfers which implement a given allocation rule can be written in integral form. More generally, Revenue Equivalence was established to hold for distribution functions with connected support.

On the other hand, it has often been observed that when V_i is disconnected the transfers can be quite unruly. For instance, when V_i is finite then incentive compatibility leaves plenty of room for the specification of the transfers. These facts notwithstanding we will show that our approach allows to obtain a simple characterization of incentive compatible transfers also when connectedness of the type space is not satisfied. In fact, we will show that incentive compatible transfers can still be written in integral form, but not so with respect to the actual allocation rule, but with respect to an appropriately defined extension of the latter.

We start our analysis with a more careful description of the setting which we are considering. There is a seller who owns a single unit of a good. He faces I bidders, indexed by $i = 1, 2, ..., I$. Each bidder has valuations distributed according to a distribution function F_i on $\mathbb{R}_+$. The support of the distribution function F_i, i.e. the set of elements v such that for every $\epsilon > 0$, $F_i(v + \epsilon) - F_i(v - \epsilon) > 0$, is denoted by V_i. Observe that V_i is a closed set. We assume that V_i is bounded and denote its smallest and largest elements by $\underline{v}_i$ and $\bar{v}_i$, respectively. Finally, we label the set of all valuation profiles by $V = \times_{i=1}^{I} V_i$.

Valuations are independently distributed, and the distribution functions, F_i, are commonly known among all players. Bidders are risk neutral, and have utility functions separable in money and the good; i.e., if a bidder with a valuation v_i wins the good with probability x_i and pays the transfer τ_i his utility is

$$x_i v_i - \tau_i.$$

It is important to point out that the assumption of linearity of payoffs in types and choices is made for convenience only. Most of the results in this section easily extend to the framework with quasi-linear utilities and strictly increasing differences in (x_i, τ_i).[4]

A direct mechanism is a pair of functions, (Q, T), which for each profile of types $v \in V$ specify with which probability each bidder obtains the object ($Q : V \to [0, 1]^I$, $\sum_i Q_i(v) \leq 1$ for all v) and how much each of them has to pay ($T : V \to \mathbb{R}^I$).

We use q_i to denote bidder i's expected probability of winning (i.e. $q_i(v_i) = E_{v_{-i}}[Q_i(v)]$),

[4]See Milgrom and Shannon (1994) for an in-depth treatment of comparative statics analysis under increasing differences.

and t_i for the expected transfer he pays ($t_i(v_i) = E_{v_{-1}}[T_i(v)]$). We call a mechanism (Q, T) (Bayesian) incentive compatible if for every bidder i

$$q_i(v_i)v_i - t_i(v_i) \geq q_i(v_i')v_i - t_i(v_i'),$$

for every $v_i, v_i' \in V_i$. It is well known that for a function q_i there exists a t_i such that the pair (q_i, t_i) satisfies this condition if and only if q_i is non-decreasing. We say that (Q, T) is (interim) individually rational if for every bidder i we have

$$q_i(v_i)v_i - t_i(v_i) \geq 0,$$

for every $v_i \in V_i$.

Since in this framework incentive compatibility and individual rationality can be expressed in terms of the functions q_i and t_i only, it is convenient to conduct our analysis in terms of these functions rather than in terms of the actual mechanism (Q, T).[5] With a slight abuse of terminology we will therefore refer to a pair (q, t) as a mechanism and to its components as allocation rules (q_i) and transfer rules (t_i). We will also say that a pair (q_i, t_i) is incentive compatible if it satisfies bidder i's incentive compatibility constraint. In this case we say that t_i implements q_i. Finally, we say that q_i satisfies Revenue Equivalence if all the expected transfers which implement it differ by a constant.

We start our analysis by showing that in this framework the bounds on the transfer differentials with respect to the lowest types $\underline{v}_i$ can be defined in a simpler and more convenient fashion compared to the definition which we have seen in Section 3. For every bidder i let $S_i^*(v_i; \underline{v}_i)$ be the set of all finite nondecreasing sequences in V_i which start at $\underline{v}_i$ and which end at v_i. Furthermore, let $\underline{t}_i^*(\cdot; \underline{v}_i, q_i)$ and $\bar{t}_i^*(\cdot; \underline{v}_i, q_i)$ be the functions which one obtains by replacing in the definitions of $\underline{t}_i(\cdot; \underline{v}_i, q_i)$ and $\bar{t}_i(\cdot; \underline{v}_i, q_i)$, the set $S(\cdot; \underline{v}_i)$ with the set $S^*(\cdot; \underline{v}_i)$.

Lemma 3. *Suppose q_i is nondecreasing. Then*

$$\bar{t}_i^*(v_i; \underline{v}_i, q_i) = \bar{t}_i(v_i; \underline{v}_i, q_i), \quad \textit{and} \quad \underline{t}_i^*(v_i; \underline{v}_i, q_i) = \underline{t}_i(v_i; \underline{v}_i, q_i),$$

for every $v_i \in V_i$.

Proof. See the Appendix. □

For the above considered transfer differentials $\bar{t}_i(\cdot; \underline{v}_i, q_i)$ and $\underline{t}_i(\cdot; \underline{v}_i, q_i)$ we will now provide a tighter and simpler characterization. In order to do so we introduce the following notation.

[5] For an in depth treatment of the question which pairs (q, t) can arise from actual mechanism see Border (1991).

For every implementable allocation rule q_i we define two functions, $\underline{q}_i$ and $\bar{q}_i$, which extend q_i from V_i to $[\underline{v}_i, \overline{v}_i]$, as follows

$$\begin{aligned} \underline{q}_i(v_i) &= \inf\{q_i(\tilde{v}_i) : \tilde{v}_i \in V_i, \tilde{v}_i \geq v_i\} \qquad \text{and} \\ \bar{q}_i(v_i) &= \sup\{q_i(\tilde{v}_i) : \tilde{v}_i \in V_i, \tilde{v}_i \leq v_i\}. \end{aligned}$$

Notice that q_i, $\underline{q}_i$ and $\overline{q}_i$ coincide on V_i, given that q_i is nondecreasing.

The following theorem shows that the upper and lower bound of the transfer differentials with respect to the lowest types, can be expressed as integrals with respect to the two extensions $\overline{q}_i$ and $\underline{q}_i$, respectively. It is important to point out that $\overline{q}_i$ and $\underline{q}_i$ are merely auxiliary tools which allow us to express the extremal transfers in a simple and convenient way. In particular, their use does not require to impose incentive compatibility outside V_i.

Theorem 2. *Let q_i be an implementable allocation rule. Then*

$$\bar{t}_i(v_i; \underline{v}_i, q_i) = \int_{\underline{v}_i}^{v_i} x d\overline{q}_i, \tag{4}$$

$$\underline{t}_i(v_i; \underline{v}_i, q_i) = \int_{\underline{v}_i}^{v_i} x d\underline{q}_i, \tag{5}$$

for every $v_i \in V_i$.

Proof. See the Appendix. □

Note that the transfer $\bar{t}_i(\cdot; \underline{v}_i, q_i)$ implies 'tight downward incentive constraints' at all types. By 'tight downward incentive constraints' for type v_i we simply mean that

$$v_i q_i(v_i) - \bar{t}_i(v_i; \underline{v}_i, q_i) = \sup_{v_i' \leq v_i} q_i(v_i')v_i - \bar{t}_i(v_i'; \hat{v}_i, q_i).$$

That is, type v_i's incentive constraints are downward tight if there are lower types with respect to which his incentive compatibility constraint is binding or comes arbitrarily close to being so. In the case of discrete types this leads to a well known concept of binding downward-adjacent IC constraints. $\underline{t}_i(\cdot; \underline{v}_i, q_i)$ corresponds to 'tight upward incentive constraints'.

Our next result is an adaption of Proposition 1 to the current framework. In particular, in the following proposition we use the characterizations of the extremal transfers given above. We assume that the planner maximizes some weighted sum of the expected utilities of the I bidders and the expected utility of the seller. The seller's valuation of the good is assumed to be 0. Thus, if we denote the planner's payoff at the mechanism (q, t) by $W(q, t)$, we have

$$W(q, t) = \lambda_0 \sum_{i=1}^{I} \int_{\underline{v}_i}^{\bar{v}_i} t_i(v_i) dF_i(v_i) + \sum_{i=1}^{I} \lambda_i \int_{\underline{v}_i}^{\bar{v}_i} [v_i q_i(v_i) - t_i(v_i)] dF_i(v_i),$$

where $\lambda_i \geq 0$, $i = 1, \dots, I$ is the welfare weight associated with bidder i while $\lambda_0 \geq 0$ represents the seller's weight. We also remind the reader that $\mathcal{Q}$ stands for the set of expected allocation rules, $q = (q_1, \dots, q_I)$, which are derived from some well behaved (in terms of measurability) allocation rule Q; similarly, we denote the set of all expected transfer rules, $t = (t_1, \dots, t_I)$ which can be obtained from some well behaved transfer rule T by $\mathcal{T}$.

Proposition 2 shows that following two problems are equivalent provided that the principal's weight exceeds the weights of the agents.

$$\begin{aligned}
\max_{(q,t)\in\mathcal{Q}\times\mathcal{T}} \quad & W(q,t) && \text{(P3)}\\
\text{s.t.} \quad & q_i(v_i)v_i - t_i(v_i) \geq q_i(v_i')v_i - t_i(v_i'), \quad \forall\, v_i, v_i' \in V_i, \quad \forall\, i = 1, \dots, I \\
& q_i(v_i)v_i - t_i(v_i) \geq 0, \quad \forall\, v_i \in V_i, \quad \forall i = 1, \dots, I \\
& \sum_{i=1}^{I} \int_{\underline{v}_i}^{\bar{v}_i} t_i(v_i) dF_i(v_i) \geq 0,
\end{aligned}$$

and

$$\begin{aligned}
\max_{(q,c)\in\mathcal{Q}\times\mathbb{R}^I} \quad & W(q,t) && \text{(P4)}\\
\text{s.t.} \quad & q_i \text{ is nondecreasing} \quad \forall i = 1, \dots, I, \\
& t_i(v_i) = c_i + \int_{\underline{v}_i}^{v_i} x d\overline{q}_i, \quad \forall\, i = 1, \dots, I \\
& q_i(\underline{v}_i)\underline{v}_i - t_i(\underline{v}_i) \geq 0, \quad \forall\, i = 1, \dots, I
\end{aligned}$$

Proposition 2. *Suppose that $\lambda_0 \geq \lambda_i$ for all $i = 1, \dots, I$ and assume that problem (P3) admits a solution. Then the values of problems (P3) and (P4) coincide. Moreover, if (q^*, c^*) solves problem (P4) and t^* is defined by $t_i^*(v_i) = c_i^* + \int_{\underline{v}_i}^{v_i} x d\overline{q}_i^*$ for all $i = 1, \dots, I$, then (q^*, t^*) solves problem (P3). Conversely, if (q^*, t^*) is a solution of problem (P3), then there exists a c^* such that (q^*, c^*) solves program (P4).*

Proof. See the Appendix. □

The above proposition considers the case where the principal is assigned the largest welfare weight. Other cases can be handled similarly. For example, if $I \geq 2$ and λ_1 is the largest of the weights, all the transfers will go to bidder 1 and as much as possible will be extracted from all the other agents.

Thus far we have considered two extensions of q_i, namely $\overline{q}_i$ and $\underline{q}_i$. We have seen that they define the boundaries on transfers which can be extracted by an incentive compatible mechanism which specifies a zero payment for the lowest type $\underline{v}_i$. Our next objective is it to show that integrating with respect to other appropriately defined extensions of q_i to $[\underline{v}_i, \overline{v}_i]$ yields incentive compatible transfers as well. Indeed, we will prove that any incentive compatible transfer scheme for an allocation rule q_i can be obtained in such a way. In order to show this we first formally define the set of relevant extensions of q_i.

Definition 2. *A function $\tilde{q}_i : [\underline{v}_i, \bar{v}_i] \to [0, 1]$ is called a* **monotonic range preserving extension** *(henceforth, MRP extension) of an implementable allocation rule $q_i : V_i \to [0, 1]$ if it coincides with q_i on V_i, is non-decreasing, and satisfies has the same range as q.*

To illustrate the concept of MRP extensions we provide a simple example.

Example 2. Let $V_1 = \{0, 1\}$ (other bidders are irrelevant for the example) and $q_1(0) = 0, q_1(1) = 1/2$. Observe first that q_1 is increasing and thus implementable. Remember that any range preserving extension $\tilde{q}_1$ needs to be weakly increasing and map the interval $[0, 1]$ into the set $\{0, 1/2\}$. This simply means that any such function can be characterized by a pair $(\alpha, x) \in [0, 1] \times \{0, 1/2\}$, where α indicates the point at which the function jumps from 0 to 1/2 and x is the value which the function takes at α. If $x = 0$ then the extension is left continuous, while $x = 1/2$ means that the extension is right continuous.

The following figure shows the extensions characterized by the pairs $(1/4, 0)$ (green) and $(3/4, 1/2)$ (blue).

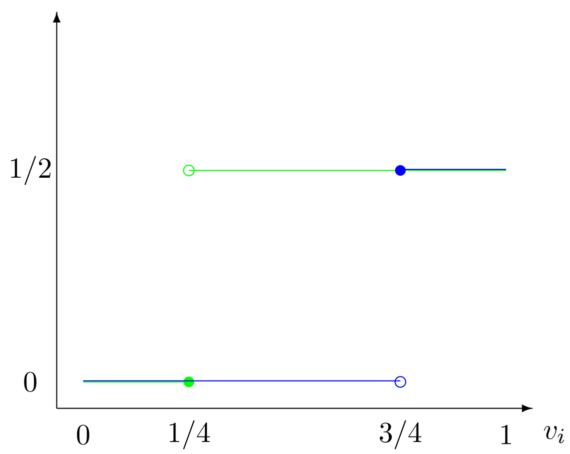

Figure 1: A left-continuous (green) and a right-continuous (blue) MRP extension of q

Before we proceed to the general result we provide a characterization for the case of connected valuation spaces V_i. This was, of course, originally done by Myerson (1981). Nevertheless, it is instructive to restate the result for our approach yields a simple alternative proof.

Proposition 3. *Let V_i be compact and connected, and let (q_i, t_i) be incentive compatible. Then there exists a $c_i \in \mathbb{R}$ such that*

$$t_i(v_i) = c_i + \int_{\underline{v}_i}^{v_i} x dq_i(x),$$

for all $v_i \in V_i$.

Proof. Let q_i be an implementable allocation rule. By the definition of the Riemann-Stieltjes integral we have

$$\int_{\underline{v}_i}^{v_i} x dq_i(x) = \bar{t}_i^*(v_i; \underline{v}_i, q_i) = \underline{t}_i^*(v_i; \underline{v}_i, q_i),$$

for every $v_i \in V_i$. Using

$$\underline{t}_i^*(v_i; \underline{v}_i, q_i) = \underline{t}_i(v_i; \underline{v}_i, q_i) \leq t_i(v_i) - t_i(\underline{v}_i) \leq \bar{t}_i(v_i; \underline{v}_i, q_i) = \bar{t}_i^*(v_i; \underline{v}_i, q_i),$$

and setting $c_i = t_i(\underline{v}_i)$, yields the result. □

The idea of the above proposition is as follows. From incentive compatibility we obtain the following bounds on the transfer differential with respect to $\underline{v}_i$:

$$\underline{v}_i[q_i(v_i) - q_i(\underline{v}_i)] \leq t_i(v_i) - t_i(\underline{v}_i) \leq v_i[q_i(v_i) - q_i(\underline{v}_i)].$$

Since q_i is non-decreasing it follows that for every $v_i > \underline{v}_i$ one can think of $v_i[q_i(v_i) - q_i(\underline{v}_i)]$ and $\underline{v}_i[q_i(v_i) - q_i(\underline{v}_i)]$ as as the upper and lower Riemann-Stieltjes sums, respectively, of the identity function on $[\underline{v}_i, v_i]$ with respect to q_i for the trivial partition of the interval $[\underline{v}_i, v_i]$. Now take any sequence $\{v_i^j\}$ in $S^*(v_i; \underline{v}_i)$. This sequence defines a finer partition of $[\underline{v}_i, v_i]$. Again one can use the incentive constraints to bound the differentials $t_i(v_i^j) - t_i(v_i^{j-1})$ by $v_i^{j-1}[q_i(v_i^j) - q_i(v_i^{j-1})]$ and $v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})]$, respectively. Summing these approximations we get the upper and lower Riemann-Stieltjes sums for the new partition. In particular, we have

$$\begin{aligned} \sum_j v_i^{j-1}[q_i(v_i^j) - q_i(v_i^{j-1})] \quad &\leq \quad \sum_j [t(v_i^j) - t(v_i^{j-1})] = t_i(v_i) - t(\underline{v}_i) \qquad \text{and} \\ \sum_j v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})] \quad &\geq \quad \sum_j [t(v_i^j) - t(v_i^{j-1})] = t_i(v_i) - t(\underline{v}_i). \end{aligned}$$

It is easily seen that if the partitions become finer (in the sense that the length of all cells goes to zero uniformly), both the upper and the lower Riemann-Stieltjes sums converge to the same number. In other words, the Riemann-Stieltjes integral is well defined and it exactly describes the transfer differential with respect to $\underline{v}_i$. Thus, we can indeed write

$$t_i(v_i) - t_i(\underline{v}_i) = \int_{\underline{v}_i}^{v_i} x dq(x).$$

The proof of the above result requires connectedness of the type set V_i. We now proceed to show how an equally simple and convenient full characterization of the transfers can be obtained when V_i fails to be connected. It is instructive to start with the particularly simple case in which V_i consists of only two valuations.

Lemma 4. *Let $V_i = \{\underline{v}_i, \overline{v}_i\}$, and let q_i be nondecreasing. Transfer t_i implements q_i if and only if there exists a constant c_i and a MRP extension $\tilde{q}_i$ of q_i such that*

$$t_i(v_i) = c_i + \int_{\underline{v}_i}^{v_i} x d\tilde{q}_i(x), \tag{8}$$

for every $v_i \in V_i$.

Proof of Lemma 4. ($\Leftarrow$) Let $\tilde{q}_i$ be some MRP extension of q_i with a jump at the point $\alpha \in [\underline{v}_i, \overline{v}_i]$ (see the discussion in the preceding example). Moreover, take any $c_i \in \mathbb{R}$ and let t_i be defined by $t_i(v_i) = c_i + \int_{\underline{v}_i}^{v_i} x d\tilde{q}_i(x)$. Then

$$\underline{v}_i[q_i(\bar{v}_i) - q_i(\underline{v}_i)] \leq \alpha[q_i(\bar{v}_i) - q_i(\underline{v}_i)] = t_i(\bar{v}_i) - t_i(\underline{v}_i) = \alpha[q_i(\bar{v}_i) - q_i(\underline{v}_i)] \leq \bar{v}_i[q_i(\bar{v}_i) - q_i(\underline{v}_i)],$$

thus rendering (q_i, t_i) incentive compatible.

($\Rightarrow$) As for the other direction assume that (q_i, t_i) is incentive compatible so that

$$\underline{v}_i[q_i(\overline{v}_i) - q_i(\underline{v}_i)] \leq t_i(\overline{v}_i) - t(\underline{v}_i) \leq \overline{v}_i[q_i(\overline{v}_i) - q_i(\underline{v}_i)].$$

But then there exists an $\alpha \in [\underline{v}_i, \overline{v}_i]$ such that $t_i(\overline{v}_i) - t_i(\underline{v}_i) = \alpha[q_i(\overline{v}_i) - q_i(\underline{v}_i)]$, and a MRP extension $\tilde{q}_i$ of q_i with a jump at α. c_i is set to $t_i(\underline{v}_i)$. □

The ideas on which the proof of the preceding result is built, carry over to the more general case where V_i may contain more than two elements.

Proposition 4. *Let q_i be an implementable allocation rule. Transfer t_i implements q_i if and only if there exists an MRP extension $\tilde{q}_i$ of q_i and a constant c_i such that*

$$t_i(v_i) = c_i + \int_{\underline{v}_i}^{v_i} x d\tilde{q}_i(x), \tag{9}$$

for every $v_i \in V_i$. Moreover, the extension $\tilde{q}_i$ is unique up to a set of (Lebesgue) measure zero.

Proof. See the Appendix. □

5 Bilateral Trade

In this section we show how our extremal transfers can be fruitfully applied to bilateral. In particular, we will use them to formulate the exact condition for the existence of budget balanced transfers which implement a given allocation rule (in particular, the efficient one).

Before we do so we shortly describe the framework. Essentially, our set up can be seen as a generalized version of the one considered in Myerson and Satterthwaite (1983), the only difference being that we do not restrict individuals' sets of valuations to be intervals.

A single indivisible object is to be allocated between two individuals, the initial owner of the good or seller (S) and a (potential) buyer (B). Player i's payoff when he obtains (retains) the object with probability x_i and pays a transfer of τ_i is

$$x_i v_i - \tau_i,$$

where v_i indicates the player's privately observed valuation of the object. Notice that we stick to our convention that transfers are interpreted as payments to be made by the individuals. In the environment which we are considering in this section, the seller must be induced to give up the object which he owns already. Of course, this requires that he *receives* a monetary compensation. That is, the transfer T_S will typically be a negative quantity.

We assume that v_i is drawn from a distribution function $F_i : \mathbb{R} \to [0, 1]$ whose support, V_i, is not necessarily an interval. We only impose that V_i is bounded and write $\underline{v}_i$ and $\bar{v}_i$ for its lower and upper bound, respectively. Notice, that since V_i is closed (being the support of F_i) it follows that $\underline{v}_i, \bar{v}_i \in V_i$. The two individuals' valuations are independent and their respective distributions commonly known.

A direct mechanism for this environment is a pair (Q, T), where

$$Q : V_B \times V_S \to [0, 1]^2 \qquad \text{and} \qquad T : V_B \times V_S \to \mathbb{R}^2.$$

The function Q describes the probabilities with which each individual gets the object. Since the object either remains with the seller or is transferred to the buyer, Q is required to satisfy the condition

$$Q_B(v_B, v_S) = 1 - Q_S(v_B, v_S)$$

for all (v_B, v_S). Notice, that the probability with which the good is assigned to the buyer, Q_B, can also be interpreted as the probability of trade.

Following our notational conventions from earlier sections we denote the expectations of Q_i and T_i, $i = B, S$, with respect to F_{-i} by q_i and t_i, respectively. Thus, under mechanism (Q, T) the interim utility of type v_i of individual $i = B, S$ from reporting his true valuation is

$$q_i(v_i)v_i - t_i(v_i).$$

We say that a mechanism (Q, T) is *interim individually rational* (or simply, individually rational) if each player's interim payoff from truth-telling is no smaller than his payoff from not participating in the mechanism. For the buyer this means that his payoff should be at least zero, while for the seller the interim utility must be no smaller than the utility which he gets from keeping the object. That is, for the seller participating in the mechanism is individually rational if

$$q_S(v_S)v_S - t_S(v_S) \geq v_S$$

, or equivalently $[q_S(v_S) - 1]v_S - t_S(v_S) \geq 0$.

The question of central interest in the literature on bilateral trade is whether it is possible that the two individuals voluntarily engage in an exchange of the good (against the payment of some transfer) whenever it is efficient to do so. Or put in more technical terms, can the efficient allocation be implemented in a way so that both individuals are at least as well off as at the initial allocation and the transfers extracted from the buyer suffice in order to compensate the seller for giving up the good. This latter condition is typically referred to as budget balancedness of a mechanism and is defined more formally as follows.

Definition 3. *A mechanism* (Q, T) *is (ex post)* **budget balanced** *if*

$$T_B(v_B, v_S) + T_S(v_B, v_S) = 0$$

for every $(v_B, v_S) \in V_B \times V_S$. *It is* **ex ante budget balanced** *if*

$$\int t_B(v_B)dF_B + \int t_S(v_S)dF_S = 0.$$

Finally, we say that it **runs an ex ante budget surplus** *if*

$$\int t_B(v_B)dF_B + \int t_S(v_S)dF_S \geq 0.$$

We are now ready to state the main result of this section which shows that there exists a close link between the possibility of implementing an allocation rule in an individually rational and budget balanced way, and the sign of the ex ante budget of the extremal transfers $\bar{t}_B(\cdot; \underline{v}_B, q_B)$ and $\bar{t}_S(\cdot; \bar{v}_S, q_S)$.

Proposition 5. *Let Q be an implementable allocation rule. There exists a transfer rule T such that* (Q, T) *is budget balanced, incentive compatible and interim individually rational if and only if the transfer scheme* $\bar{T}$ *defined by* $\bar{T}(v_B, v_S) = (\underline{v}_B q_B(\underline{v}_B) + \bar{t}_B(v_B; \underline{v}_B, q_B), \bar{v}_S[q_S(\bar{v}_S) - 1] + \bar{t}_S(v_S; \bar{v}_S, q_S))$, *runs an ex ante budget surplus.*

Proof. See the Appendix. □

By Theorem 1 we know that the (interim) transfers, $\bar{t}_B$ and $\bar{t}_S$, as defined in the statement of the above proposition, extract the most money from the buyer and give the least to the seller (sellers transfer is negative) among all incentive compatible mechanisms which satisfy the individual rationality constraints of the lowest buyer type, $\underline{v}_B$, and the highest seller type, $\bar{v}_S$, with equality. Hence, these (interim) transfers create the highest budget surplus among all incentive compatible and individually rational transfers. Therefore, if $(Q, \bar{T})$ is an incentive compatible mechanism which runs a strict ex ante budget deficit, then there cannot exist an ex post budget balanced incentive compatible and individually rational mechanism (Q, T). We show that the converse is also true: if $(Q, \bar{T})$ runs an ex ante budget surplus, then an incentive compatible, individually rational and ex post budget balanced mechanism can be constructed. The next Corollary puts the machinery we developed above to work.

Corollary 2. *Let Q^* be an ex post efficient allocation rule and let $\bar{T}$ be defined by $\bar{T}_B(v_B) = \underline{v}_B q_B^*(\underline{v}_B) + \bar{t}_B(v_B; \underline{v}_B, q_B^*)$ and $\bar{T}_S(v_S) = \bar{v}_S[q_S^*(\bar{v}_S) - 1] + \bar{t}_S(v_S; \bar{v}_S, q_S^*)$. There exists a transfer T such that (Q^*, T) is budget balanced, individually rational and incentive compatible if and only if $(Q^*, \bar{T})$ runs a budget surplus.*

Since we allow for general distributions there may exist several ex post efficient allocation rules, therefore the 'an' in the statement of the theorem. Now notice that the transfers $\bar{t}_i$, $i = B, S$ depend both on the allocation rule $Q*$ and the distributions of the individuals' valuations, F_i, $i = B, S$. Thus, once the allocation rule $Q*$ is fixed, the condition that $(Q*, \bar{T})$ runs an ex ante budget surplus is really a condition on the distributions.

By the celebrated result of Myerson and Satterthwaite (1983) we know that if the sets V_i are intervals which overlap, then no transfer which implements an efficient allocation rule $Q*$ so that individual rationality is satisfied can be budget balanced. On the other hand, Myerson and Satterthwaite (1983) also provide an example with finite valuation sets in which the efficient allocation rule can be implemented respecting both individual rationality and budget balancedness. A general treatment of the case of finite type spaces is given in Kos and Manea (2009). They provide necessary and sufficient conditions on discrete distributions under which budget balanced and individually rational implementation of a given allocation rule is feasible.

6 Conclusion

We provide a characterization of the highest and the lowest transfers - given the fixed payment of some type - which implement a given allocation rule. We obtain this characterization under very general assumptions on the environment. In particular, we do not impose any restrictions on the set of types or the distributions from which they are drawn, the set of allocations or the payoff functions which go beyond the minimal mathematical conditions guaranteeing that all expectations are well defined.

Based on our characterization of the extremal incentive compatible transfers we obtain a characterization of the set of implementable allocation rules and a characterization of the set of allocation rules which satisfy Revenue Equivalence. But we do not stop at the Revenue Equivalence. In fact, we show that often the extremal transfers are all a mechanism designer needs to care about. In such cases the designer's problem can be simplified by the transfer dimension whether or not the Revenue Equivalence holds.

In the last two sections we consider more specific environments with one-dimensional type sets and payoff functions which are linear in types. For such environments the existing literature proposes different approaches for solving the corresponding mechanism design problem depending on the 'size' (cardinality) of the valuation sets. Since our tools are applicable independently of the cardinality of the type spaces we are able to unify those methods. In fact, we show that our approach leads to a characterization of the transfers which takes a very simple and convenient form independently of the size of the type spaces.

Finally, in a bilateral trade setting we formulate necessary and sufficient conditions for the existence of an incentive compatible, individually rational, ex post efficient and budget balanced trading mechanism. We thus generalize the results obtained in the much celebrated paper by Myerson and Satterthwaite (1983) where it is shown that such mechanisms never exist in environments where the supports of the type distributions are intervals and that they may exist in environments with finite type sets.

The results thus far obtained apply to static environments. We explore possible ways of extending our approach and findings to dynamic setups in the companion paper.

7 Appendix

Proof of Lemma 2. $I(v;\hat{v},q) \neq \emptyset$ and $-\infty < \underline{t}(v;\hat{v},q)$ imply $-\infty < \bar{t}(v;\hat{v},q)$, for every $v \in V$. We will show that $\bar{t}$ implements q. That is, we have to verify that

$$u(q(v),v) - \bar{t}(v;\hat{v},q) \geq u(q(v'),v) - \bar{t}(v';\hat{v},q),$$

for every $v, v' \in V$. By the definition of $\bar{t}(v';\hat{v},q)$ for every $\epsilon > 0$ there exists a sequence $\{v^j\}_{j=0}^n \in S(v';\hat{v})$ such that

$$\sum_{j=1}^{n} \left[u(q(v^j),v^j) - u(q(v^{j-1}),v^j)\right] - \epsilon \leq \bar{t}(v';\hat{v},q). \tag{10}$$

Now,

$$\begin{aligned} \bar{t}(v;\hat{v},q) - \bar{t}(v';\hat{v},q) &\leq \bar{t}(v;\hat{v},q) - \sum_{j=1}^{n} \left[u(q(v^j),v^j) - u(q(v^{j-1}),v^j)\right] + \epsilon \\ &\leq u(q(v),v) - u(q(v'),v) + \epsilon, \end{aligned} \tag{11}$$

where the first inequality follows from (10) and the second from the fact that $\bar{t}(v;\hat{v},q) \leq u(q(v),v) - u(q(v'),v) + \sum_{j=1}^{n} \left[u(q(v^j),v^j) - u(q(v^{j-1}),v^j)\right]$ (that is, being the infimum over all sums, $\bar{t}(v;\hat{v},q)$ can be no larger than the sum corresponding to the sequence which is obtained by adding v as the final element to the sequence $\{v^j\}_{j=0}^n$). Since the inequality (11) holds for any ϵ we are done. □

Proof of Theorem 1. Statement i). The fact that implementability of q implies that the correspondence $I(\cdot;\hat{v},q)$ is nonempty is a direct consequence of Lemma 1. The other direction instead follows from Lemma 2.

Statement ii). (⇒) Suppose q satisfies Revenue Equivalence. Since $\bar{t}(\cdot;\hat{v},q)$ and $\underline{t}(\cdot;\hat{v},q)$ are two transfer schemes which implement q, Revenue equivalence implies that they can differ only by a constant. But given that the two transfers coincide at the anchor type $\hat{v}$ this constant must be zero, meaning that they coincide everywhere.

(⇐) Let q be such that $|I(v;\hat{v},q)| = 1$ for every $v \in V$. By the first statement of the theorem q is implementable. So let t, t' be two transfer rules which implement q. Since $|I(v;\hat{v},q)| = 1$ is equivalent to $\bar{t}(v;\hat{v},q) = \underline{t}(v;\hat{v},q)$ for every $v \in V$, we have

$$\begin{aligned} \underline{t}(v;\hat{v},q) &= t(v) - t(\hat{v}) = \bar{t}(v;\hat{v},q), \quad \text{and} \\ \underline{t}(v;\hat{v},q) &= t'(v) - t'(\hat{v}) = \bar{t}(v;\hat{v},q), \end{aligned}$$

for every $v \in V$. But then

$$t(v) - t(\hat{v}) = t'(v) - t'(\hat{v})$$

for every $v \in V$, which is a restatement of revenue equivalence.

Statement iii) is simply a restatement of Lemma 1. □

Proof of Proposition 1. In what follows we will formally prove only the equivalence of the two programs in terms of value. It is straightforward to see that the argument also extends to the problems' solutions. We first argue that for any mechanism (q, t_A, t_P) which is feasible in (P1) there exists a pair (q, c) which satisfies the constraints of (P2)) and which gives the planner a weakly higher payoff than (q, t_A, t_P). This implies that the value of the second problem can be no smaller than the value of the first program. The converse statement simply follows from the observation that for any pair (q, c) which satisfies the constraints of (P2) the corresponding mechanism $(q, \bar{t}_A(\cdot; \hat{v}, q) + c, -\bar{t}_A(\cdot; \hat{v}, q) - c)$ is feasible in the first problem.

Suppose (q, t_A, t_P) satisfies the constraints of program (P1). We will argue now that the pair $(q, t_A(\hat{v}))$ is feasible in the second problem. If (q, t_A, t_P) is feasible in the first program then it must satisfy the agent's incentive compatibility constraint. By Theorem 1 we know that implementability of q is equivalent to $\bar{t}_A(\cdot; \hat{v}, q) \geq \underline{t}_A(\cdot; \hat{v}, q)$. Hence the first constraint of (P2) is certainly satisfied. Next, notice that the transfer $\tilde{t}_A(v) = t_A(\hat{v}) + \bar{t}_A(v; \hat{v}, q)$ not only implements q but dominates all other transfers of the agent which do so and require the type $\hat{v}$ to pay $t_A(\hat{v})$. This implies that the corresponding transfer to the principal, $-\tilde{t}_A$ must be smaller than t_P ($\tilde{t}_A \geq t_A$ iff $-\tilde{t}_A \leq -t_A$; combining this with the budget constraint $t_P + t_A \geq 0$ delivers $t_P \geq -t_A \geq -\tilde{t}_A$). Thus, given that the principal's individual rationality constraint was satisfied in (P1) it must also be satisfied in (P2). Finally, it remains to be verified that the pair $(q, \tilde{t}_A)$ is individually rational for the agent. But this is implied by the following observation:

$$\begin{aligned} u_A(q(v), v) - \tilde{t}_A(v) &\geq u_A(q(\hat{v}), v) - \tilde{t}_A(\hat{v}) \\ &\geq u_A(q(\hat{v}), \hat{v}) - \tilde{t}_A(\hat{v}) \geq 0. \end{aligned}$$

The first inequality simply follows from incentive compatibility. The second one is a consequence of our assumption that $u_A(\cdot, \hat{v}) \leq u_A(\cdot, v)$. The third one instead follows from the construction of $\tilde{t}_A$. At $\hat{v}$, $\tilde{t}_A$ coincides with t_A and thus the inequality follows from the fact that t_A satisfies individual rationality in (P1).

Finally, we have to show that $W(q, \tilde{t}_A, -\tilde{t}_A) - W(q, t_A, t_P) = \int [\lambda_P(\tilde{t}_A + t_P) - \lambda_A(\tilde{t}_A - t_A)] dF \geq 0$. But in order to see this just observe that $[\tilde{t}_A + t_P] - [\tilde{t}_A - t_A] = t_P + t_A \geq 0$. Combining this with $\lambda_P \geq \lambda_A \geq 0$ the result follows immediately. □

Proof of Lemma 3. We will prove the first equality, the second one is obtained similarly. Notice first that $\bar{t}_i(v_i; \underline{v}_i, q_i) \leq \bar{t}_i^*(v_i; \underline{v}_i, q_i)$ follows directly from $S_i^*(v_i; \underline{v}_i) \subset S_i(v_i; \underline{v}_i)$.

In order to prove the opposite inequality we will show that for any non-monotonic sequence $\{v_i^j\}_{j=0}^n$ in $S_i(v_i; \underline{v}_i)$ there exists a (monotonic) sequence $\{\tilde{v}_i^j\}_{j=0}^r$ in $S_i^*(v_i; \underline{v}_i)$ such that

$$\sum_{j=1}^{r} \tilde{v}_i^j [q_i(\tilde{v}_i^j) - q_i(\tilde{v}_i^{j-1})] \leq \sum_{j=1}^{n} v_i^j [q_i(v_i^j) - q_i(v_i^{j-1})]. \tag{12}$$

Our proof is constructive. That is, we will show how $\{\tilde{v}_i^j\}_{j=0}^r$ can be obtained from $\{v_i^j\}_{j=0}^n$ in a finite iterative procedure which successively eliminates the downward jumps of the latter sequence.

Take a sequence $\{v_i^j\}_{j=0}^n$ from $S_i(v_i; \underline{v}_i)$ which is not (weakly) increasing and let v_k be the first element of this sequence which is larger than its successor, i.e. which satisfies $v_k > v_{k+1}$. Since $v_i^0 = \underline{v}_i$, it follows that $k > 0$. Next, let v_m be the last element of the monotonic subsequence $\{v_i^j\}_{j=0}^k$ which is no larger than v_i^{k+1}. The existence of v_m is guaranteed by the fact that $v_i^0 = \underline{v}_i \leq v_i^j$ for all $j = 0, \ldots, n$. Now eliminate from $\{v_i^j\}_{j=0}^n$ all elements between v_i^m and v^{k+1} (the spike around v_k) and denote the resulting new sequence by $\{\hat{v}_i^j\}_{j=0}^l$ (that is, $\{\hat{v}_i^j\}_{j=0}^l$ is identical to $(v_i^0, \ldots, v_i^m, v_i^{k+1}, v_i^{k+2}, \ldots, v_i^n)$). If $\{\hat{v}_i^j\}_{j=0}^l$ is monotonically increasing we are done. Otherwise, we iterate the procedure. Since we are considering only finite sequences we must obtain a non-decreasing sequence $\{\tilde{v}_i^j\}_{j=0}^r$ in a finite number of steps .

It remains to be shown that $\{\tilde{v}_i^j\}_{j=0}^r$ satisfies condition (12). In order to do so it is sufficient to show that

$$\sum_{j=1}^{l} \hat{v}_i^j [q_i(\hat{v}_i^j) - q_i(\hat{v}_i^{j-1})] \leq \sum_{j=1}^{n} v_i^j [q_i(v_i^j) - q_i(v_i^{j-1})],$$

since that implies that also any further step of the iterative procedure will not lead to a sequence for which the corresponding sum is larger than $\sum_{j=1}^n v_i^j [q_i(v_i^j) - q_i(v_i^{j-1})]$.

For this last step observe that

$$\begin{aligned}
&\sum_{j=m+1}^{k+1} v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})] \\
&\quad = \sum_{j=m+2}^{k+1} \left[v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})\right] + v_i^{m+1}[q_i(v_i^{m+1}) - q_i(v_i^m)] \\
&\quad = \sum_{j=m+2}^{k+1} \left[v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})]\right] + v_i^{m+1}[q_i(v_i^{m+1}) - q_i(v_i^{k+1}) + q_i(v_i^{k+1}) - q_i(v_i^m)] \\
&\quad \geq \sum_{j=m+2}^{k+1} \left[v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})]\right] + v_i^{m+1}[q_i(v_i^{m+1}) - q_i(v_i^{k+1})] + v_i^{k+1}[q_i(v_i^{k+1}) - q_i(v_i^m)] \\
&\quad \geq v_i^{k+1}[q_i(v_i^{k+1}) - q_i(v_i^m)],
\end{aligned}$$

where the first inequality follows from $v_i^{m+1} \geq v_i^{k+1} \geq v_i^m$ and q_i being nondecreasing, and the second from the fact that q_i is implementable. That is, in the discussion following Theorem 1 we have argued that implementability of q_i is equivalent to the non-existence of a cycle of valuations over which the above sum, $\sum_{j=m+2}^{k+1}\left[v_i^j[q_i(v_i^j) - q_i(v_i^{j-1})]\right] + v_i^{m+1}[q_i(v_i^{m+1}) - q_i(v_i^{k+1})]$, is negative. □

Proof of Theorem 2. Before we prove the theorem, it is worth pointing out that the integrals in the statement are to be seen as Riemann-Stieltjes integrals. Notice, that they are well defined since the integrators $\underline{q}_i$ and $\overline{q}_i$ are monotonic and the integrand is continuous.

We will prove the first equality; the second one is handled similarly. From Lemma 3 we know

$$\bar{t}_i(v_i; \underline{v}_i, q_i) = \bar{t}_i^*(v_i; \underline{v}_i, q_i), \tag{13}$$

therefore it is enough to show

$$\int_{\underline{v}_i}^{v_i} x d\overline{q}_i = \bar{t}_i^*(v_i; \underline{v}_i, q_i). \tag{14}$$

By definition the Riemann-Stieltjes integral of the identity function with respect to the function $\bar{q}_i$ is the infimum of the upper Riemann-Stieltjes sums which take the form $\sum_{j=1}^m v_i^j[\overline{q}_i(v_i^j) - \overline{q}_i(v_i^{j-1})]$ where the sequence $\{v_i^j\}_{j=0}^m$ is increasing with $v_i^0 = \underline{v}_i$ and $v_i^m = v_i$; the infimum is taken over all such sequences in $[\underline{v}_i, v_i]$. But then $\bar{t}_i^*(v_i; \underline{v}_i, q_i)$, which is the infimum of the same type of sums over sequences in $V_i \cap [\underline{v}_i, v_i]$, is no smaller. Hence

$$\int_{\underline{v}_i}^{v_i} x d\overline{q}_i \leq \bar{t}_i^*(v_i; \underline{v}_i, q_i).$$

We are left to show the opposite inequality. Suppose to the contrary that

$$\Delta \equiv \bar{t}_i^*(v_i; \underline{v}_i, q_i) - \int_{\underline{v}_i}^{v_i} x d\overline{q}_i > 0. \tag{15}$$

By the definition of the integral an increasing sequence $\{v_i^j\}_{j=0}^n$ in $[\underline{v}_i, v_i]$ with $v_i^0 = \underline{v}_i$ and $v_i^n = v_i$ exists such that

$$0 \leq \sum_{j=1}^{n} v_i^j \left[\overline{q}_i(v_i^j) - \overline{q}_i(v_i^{j-1})\right] - \int_{\underline{v}_i}^{v_i} x d\overline{q}_i < \Delta/2. \tag{16}$$

Next, consider the sequence $\{\tilde{v}_i^j\}_{j=0}^n$ defined as follows: $\tilde{v}_i^j = v_i^j$ for all j such that $v_i^j \in V_i$; otherwise set $\tilde{v}_i^j = \max\{v \in V_i : v < v_i^j\}$. Notice that $\{\tilde{v}_i^j\}_{j=0}^n$ is well defined (since V_i is closed) and entirely contained in V_i. Furthermore, by the definition of $\bar{q}_i$, we have $\bar{q}_i(v_i^j) = \bar{q}_i(\tilde{v}_i^j)$ for all j. By construction we have $\tilde{v}_i^j \leq v_i^j$ and $\overline{q}_i(v_i^j) - \overline{q}_i(v_i^{j-1}) = q_i(\tilde{v}_i^j) - q_i(\tilde{v}_i^{j-1})$ for every j, therefore

$$\sum_{j=1}^{n} \tilde{v}_i^j \left[q_i(\tilde{v}_i^j) - q_i(\tilde{v}_i^{j-1})\right] \leq \sum_{j=1}^{n} v_i^j \left[\overline{q}_i(v_i^j) - \overline{q}_i(v_i^{j-1})\right]. \tag{17}$$

Inequalities (16) and (17) imply

$$\sum_{j=1}^{n} v_i^j \left[q_i(\tilde{v}_i^j) - q_i(\tilde{v}_i^{j-1})\right] - \int_{\underline{v}_i}^{v_i} x d\overline{q}_i < \frac{\Delta}{2}, \tag{18}$$

which, using the definition of $\bar{t}_i^*(\cdot; \underline{v}_i, \bar{q}_i)$ contradicts (15). □

Proof of Proposition 2. In what follows we only show the equivalence of the two optimization programs in terms of the values which they deliver. The argument is similar to the one which we have used in the proof of Proposition 1.

We first argue that for each pair $(\tilde{q}, \tilde{t})$ which is feasible in program (P3) there is a c such that $(\tilde{q}, c)$ satisfies the constraints of the second program and the corresponding payoff of the planner is no smaller than $W(q, t)$. First, since $\tilde{q}$ is implementable, each of its components must be non-decreasing. Second, set $c_i = \tilde{t}_i(\underline{v}_i)$ and denote the transfers which this implies in program (P4) by t. Clearly, we have $t_i(\underline{v}_i) = \tilde{t}_i(\underline{v}_i)$ for all i. Thus the lowest type's payoff is the same in both cases and so the last constraint of (P4) is satisfied. Finally, observe that by construction we have $t_i(v_i) \leq \tilde{t}_i(v_i)$ for all $i = 1, \ldots, I$. Since $\lambda_0 \geq \lambda_i$ for all $i \geq 1$ we can therefore conclude that $W(\tilde{q}, \tilde{t}) \leq W(\tilde{q}, t)$.

For the other direction let $(\tilde{q}, \tilde{c})$ be a feasible pair of program (P4) and denote the corresponding transfers by $\tilde{t}$. We will argue that $(\tilde{q}, \tilde{t})$ is feasible in program (P3). Monotonicity of $\tilde{q}$ is equivalent to implementability. Furthermore, $\tilde{t}$ is such that the lowest types' individual rationality constraint is satisfied. But since the equilibrium utility in our linear environment is increasing in the bidders' valuations it follows that at the pair $(\tilde{q}, \tilde{t})$ the individual rationality constraint of all types of all bidders is satisfied. □

Proof of Proposition 4. ($\Leftarrow$) Let $\tilde{q}_i$ be some integrable extension of q_i and let t_i be defined by 9. We need to show that (q_i, t_i) is incentive compatible.

Let $v_i, v_i' \in V_i$, and $v_i > v_i'$ (the other case is handled similarly), then

$$\begin{aligned} v_i q_i(v_i) - t_i(v_i) - [v_i q_i(v_i') - t_i(v_i')] &= v_i[q_i(v_i) - q_i(v_i')] - \int_{v_i'}^{v_i} x d\tilde{q}_i(x) \\ &\geq v_i[q_i(v_i) - q_i(v_i')] - \int_{v_i'}^{v_i} v_i d\tilde{q}_i(x) \\ &= v_i[q_i(v_i) - q_i(v_i')] - v_i[\tilde{q}_i(v_i) - \tilde{q}_i(v_i')] \\ &= 0, \end{aligned}$$

where the first line follows from the definition of t_i, the second from the fact that q_i is nondecreasing, and the fourth by the fact that $\tilde{q}_i$ and q_i coincide on V_i.

($\Rightarrow$) Let (q_i, t_i) be incentive compatible. Interval $[\underline{v}_i, \overline{v}_i]$ can be written as a disjoint union of V_i and a countable set of disjoint open intervals, which we call holes. Indeed, denote the complement of V_i in $\mathbb{R}$ by V_i^C. The latter is open, thus can be written as a countable union of disjoint intervals (we are operating on the real line). But then so can be $[\underline{v}_i, \overline{v}_i] \backslash V_i$.

Likewise there can only be countably many disjoint nontrivial connected closed intervals contained in V_i. By Proposition 3 on each of these subintervals, say $[\underline{w}_i, \overline{w}_i]$, the transfer difference $t_i(\overline{w}_i) - t(\underline{w}_i)$, can be written as an integral. But now, for every $v_i \in V_i$ the transfer differential, $t_i(v_i) - t(\underline{v}_i)$ can be split into a countable sum of transfer differences of which each either corresponds to a closed interval or to boundaries of a hole. In the first case the integral is taken with respect to the actual allocation rule. In the second case we get the extension by Lemma 4.

The fact that all MRP extensions which yield the desired transfer are equivalent almost everywhere follows from inspection of Lemma 4. This lemma implies that the extensions which deliver the same transfer differential can only differ at the point of discontinuity, i.e. the point α. Since the number of holes is countable, as argued in Step 1, the two extensions can only differ on a countable set of points. □

Proof of Proposition 5. ($\Rightarrow$) We will prove this part by proving the contrapositive statement. So suppose that $\bar{T}$ runs a strict ex ante budget deficit. By Theorem 1 $\bar{t}_B$, (pointwise) dominates all other transfer schemes which implement q_B and leave the lowest type of the buyer with a zero utility. Hence, it extracts the most money from the buyer among all incentive compatible and individually rational transfers.

An analogous argument applies in the case of the seller and the transfer $\bar{t}_S$. The only difference is that in the case of the seller the crucial type for individual rationality is the one which values the object the most, $\bar{v}_S$. Of course $\bar{t}_S$, being negative, is now interpreted as the transfer that gives the least money to the seller. Now if $\bar{T}$ is the transfer which extracts the most money from the individuals (in fact gives the least to the seller) among all transfers which satisfy incentive compatibility and individual rationality it immediately follows that whenever $\bar{T}$ runs a strict ex ante budget deficit then so must any other incentive compatible and individually rational mechanism. This in turn implies that every incentive compatible and individually rational mechanism must violate (ex post) budget balancedness.

($\Leftarrow$) Suppose $(Q, \bar{T})$ runs a budget surplus of size $\Pi \leq 0$ and consider the transfer rule T defined by

$$T_B(v_B, v_S) = -T_S(v_B, v_S) = \bar{t}_B(v_B) - \bar{t}_S(v_S) - \int \bar{t}_B dF_B + \alpha\Pi,$$

where α is some number from the unit interval. By construction (Q, T) satisfies budget balance. We will show now that (Q, T) is also incentive compatible and individually rational. Observe first that for all $v_B \in V_B$ we have $t_B(v_B) = \bar{t}_B(v_B) - (1 - \alpha)\Pi$. Thus t_B differs from $\bar{t}_B$ only by a constant. But then, given that $\bar{t}_B$ implements q_B, the same also holds for t_B. Moreover, since $\bar{t}_B$ is individually rational, $t_B \leq \bar{t}_B$ implies that t_B is individually rational as well.

By an analogous argument it can be shown that $t_S = \bar{t}_S + \alpha\Pi$ implements q_S in an individually rational way and so we are done. □

References

Border, K. C. (1991). Implementation of reduced form auctions: A geometric approach. *Econometrica 59*(4), 1175–1187.

Chung, K. and W. Olszewski (2007). A non-differentiable approach to revenue equivalence. *Theoretical Economics 2*(4), 469–487.

Ely, J. (2001). Revenue equivalence without differentiability assumptions. Working Paper.

Heydenreich, B., R. Müller, M. Uetz, and R. Vohra (2009). Characterization of revenue equivalence. *Econometrica 77*(1), 307–316.

Holmstrom, B. (1979). Groves' scheme on restricted domains. *Econometrica 47*(5), 1137–1144.

Kos, N. and M. Manea (2009). Efficient trade mechanisms with discrete values. Working Paper.

Krishna, V. and M. Perry (1998). Efficient mechanism design. Pennsylvania State University.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica 62*(1), 157–180.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Myerson, R. and M. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*, 265–281.

Rochet, J. (1987). A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics 16*, 191–200.

Skreta, V. (2006). Mechanism design for arbitrary type spaces. *Economic Letters 91*, 293–299.

Vickrey, W. (1961). Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance 16*, 8–37.

Williams, S. (1999). A characterization of efficient, Bayesian incentive compatible mechanisms. *Economic Theory 14*, 155–180.