Detecting Profitable Deviations

David Rahman*
University of Minnesota

Preliminary and Incomplete

April 16, 2009

Abstract

This paper characterizes implementability of an allocation in a quasi-linear context for arbitrary type spaces. Intuitively, an allocation is implementable if and only if every profitable deviation is detectable. The result is compared to Rochet's Theorem. Several natural extensions are provided.

JEL Classification: D21, D23, D82.
Keywords: mechanism design, implementation, duality.

---

*I thank Sushil Bikhchandani for numerous conversations and encouragement.

1 Model

Consider an agent with private information parametrized by a *type* t in some set T, interpreted as the collection of all types that the principal deems possible. Let X be a nonempty set of outcomes, and denote by $v(t,x) \in \mathbb{R}$ his utility from outcome $x \in X$ when his type is $t \in T$. An *allocation* is any map $\mathbf{x} : T \to X$. An *incentive scheme* (or simply *scheme*) is any map $\xi : T \to \mathbb{R}$, interpreted as report-contingent linear transfers intended to induce the agent to report his type truthfully.

An allocation $\mathbf{x}$ is called *implementable* if a scheme ξ exists such that

$$v(t,\mathbf{x}(s)) - v(t,\mathbf{x}(t)) \;\leq\; \xi(s) - \xi(t) \qquad \forall (t,s) \in T \times T. \tag{1}$$

A *finite reporting strategy* (or simply *reporting strategy*) is any map $\pi : S \to \Delta(S)$ defined on some finite subset S of T, where $\pi(s|t)$ is interpreted as the conditional probability that the agent's report equals s when his true type equals t. For instance, the *truthful* reporting strategy θ is defined for every pair (s,t) by $\theta(s|t) = 1$ if $s = t$ and 0 otherwise. A *finite deviation* (or simply *deviation*) is any reporting strategy that isn't truthful, i.e., it lies with positive probability conditional on some type.

A reporting strategy π is called *undetectable* if

$$\sum_{t \in S} \pi(s|t) \;=\; 1 \qquad \forall s \in S. \tag{2}$$

In other words, π is doubly stochastic. Otherwise, π is called *detectable*. Suppose that types are drawn from S according to the uniform distribution (so the probability of each type equals $1/|S|$).[1] Intuitively, π is undetectable if the probability distribution over reports coincides with that of actual types, $1/|S|$.

A reporting strategy π is $\mathbf{x}$-*profitable* if it yields a higher ex ante payoff than truthful reporting assuming that the incentive scheme is identically zero, i.e.,

$$\sum_{(t,s)} \pi(s|t)[v(t,\mathbf{x}(s)) - v(t,\mathbf{x}(t))] \;>\; 0. \tag{3}$$

Otherwise, π is called $\mathbf{x}$-*unprofitable*. Intuitively, (3) says that the expected utility from reporting according to π is greater than that from reporting truthfully.

Theorem 1. *The following statements are equivalent for a given allocation* $\mathbf{x}$*:*
(i) $\mathbf{x}$ *is implementable.*
(ii) *Every* $\mathbf{x}$*-profitable deviation is detectable.*

---

[1] Although the uniform assumption is useful, it is by no means necessary for any of the results.

2 Proof of Theorem 1

First of all, let us prove the theorem under the restriction that T is a finite set.

Lemma 1. *If T is a finite set then an allocation $\mathbf{x}$ is implementable if and only if every $\mathbf{x}$-profitable deviation is detectable.*

Proof. By the Theorem of the Alternative (see, e.g., Rockafellar, 1970, page 198), a scheme $\xi \in \mathbb{R}^T$ exists such that $v(t, \mathbf{x}(s)) - v(t, \mathbf{x}(t)) \leq \xi(s) - \xi(t)$ for every $t, s \in T$ if and only if there does not exist a vector $\lambda \geq 0$ satisfying (i) $\sum_s \lambda(s,t) = \sum_s \lambda(t,s)$ for all $t \in T$, and (ii) $\sum_{(t,s)} \lambda(s,t)[v(t,\mathbf{x}(s)) - v(t,\mathbf{x}(t))] > 0$. Each of these two conditions on λ is independent of $\lambda(t,t)$ for all $t \in T$, so assume without loss of generality that $\lambda(t,t) = \max\{\sum_{s \neq r} \lambda(s,r) : r \in T\} - \sum_{s \neq t} \lambda(s,t)$ for all $t \in T$. Now λ is proportional to a doubly stochastic matrix—in other words, a reporting strategy, call it π—which satisfies (i) and (ii) if and only if λ satisfies (i) and (ii). But (i) is just the requirement that π be undetectable, and (ii) states that π is $\mathbf{x}$-profitable. $\square$

Now suppose that T is not necessarily finite. We begin with some preliminaries.

For any set Z, let $\mathbb{R}^Z$ be the space of all real-valued functions on Z endowed with the product topology, and let $\mathbb{R}^Z_+ = \{f \in \mathbb{R}^Z : f(z) \geq 0 \ \forall z \in Z\}$ denote its positive cone. Let $\mathbb{R}^{(Z)}$ be the subspace of all real-valued functions g on Z with finite support, i.e., such that the set $\{z \in Z : g(z) \neq 0\}$ is finite. Any $g \in \mathbb{R}^{(Z)}$ is described by a finite set $\operatorname{supp} g = \{z_1, \ldots, z_m\}$ of elements in Z (the support of g) together with a finite-dimensional vector $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$. Such a g acts on $\mathbb{R}^Z$ as follows:

$$g(f) \ = \ \sum_{k=1}^{m} \lambda_k f(z_k) \qquad\qquad \forall f \in \mathbb{R}^Z.$$

An important example is the *evaluation* functional $\mathbf{e}_z \in \mathbb{R}^{(Z)}$, defined by $\mathbf{e}_z(f) = f(z)$. Clearly, any $g \in \mathbb{R}^{(Z)}$ can be written in terms of these evaluations as $g = \sum_k \lambda_k \mathbf{e}_{z_k}$. It is well known (see, e.g., Holmes, 1975) that $\mathbb{R}^{(Z)}$ is the topological dual of $\mathbb{R}^Z$, i.e., the space of continuous linear functionals on $\mathbb{R}^Z$. Denote its positive cone by $\mathbb{R}^{(Z)}_+$.

Any $g \in \mathbb{R}^{(Z \times Z)}$ is given by a finite support $\{(z_{11}, z_{21}), \ldots, (z_{1m}, z_{2m})\}$ and a vector $(\lambda_1, \ldots, \lambda_m)$. We will describe it instead by the subset $\{z : z = z_{ik} \text{ for some } i, k\}$ of Z with, say, n elements, denoted by $\operatorname{supp}_Z g = \{z_1, \ldots, z_n\}$ together with the $n \times n$ matrix $(\lambda_{11}, \ldots, \lambda_{1n}, \ldots, \lambda_{n1}, \ldots, \lambda_{nn})$ defined by $\lambda_{k\ell} = \lambda_i$ if $(z_k, z_\ell) = (z_{1i}, z_{2i})$ and 0 if no such i exists. Clearly, both descriptions are equivalent.

Let $w \in \mathbb{R}^{T\times T}$ be the function defined pointwise by $w(t,s) = v(t,\mathbf{x}(s)) - v(t,\mathbf{x}(t))$. Define pointwise the following operator $D : \mathbb{R}^{(T\times T)} \to \mathbb{R}^{(T)}$. Given $g \in \mathbb{R}^{(T\times T)}$, let $Dg = \sum_{(k,\ell)} \lambda_{k\ell}(\mathbf{e}_{t_k} - \mathbf{e}_{t_\ell})$. Hence, $Dg(f) = \sum_{(k,\ell)} \lambda_{k\ell}[f(t_k) - f(t_\ell)]$ for all $f \in \mathbb{R}^T$.

Lemma 2. *The following are equivalent:*
(i) *For every $g \in \mathbb{R}_+^{(T\times T)}$, $Dg = \mathbf{0}$ implies that $g(w) \leq 0$.*[2]
(ii) *There exists a net $\{\xi_\delta\}$ such that $w(t,s) \leq \liminf_\delta \xi_\delta(s) - \xi_\delta(t)$ for all (t,s).*

Proof. Let $X = \mathbb{R}^{(T\times T)}$ and $Y = \mathbb{R}^{(T)} \times \mathbb{R}$. Let $A : X \to Y$ be the operator defined pointwise by $A(g) = (Dg, g(w))$. Since $\mathbb{R}_+^{(T\times T)}$ is a cone, (i) fails if and only if there exists $g \in \mathbb{R}_+^{(T\times T)}$ such that $Dg = \mathbf{0}$ and $g(w) = 1$, i.e., $A(g) = (\mathbf{0}, 1)$. The operator A is clearly linear and continuous, so by Lemma A.1, $A(g) = (\mathbf{0}, 1)$ if and only if given any number ε, incentive scheme ξ and net $\{(w_\delta, \xi_\delta) \in \mathbb{R}_+^{T\times T} \times \mathbb{R}^T\}$,

$$\xi(s) - \xi(t) + \varepsilon w(t,s) \;=\; \lim w_\delta(s,t) - [\xi_\delta(s) - \xi_\delta(t)] \quad \forall (t,s) \quad \Rightarrow \quad \varepsilon \geq 0.$$

Since $w_\delta \geq 0$, this condition is equivalent to

$$\xi(s) - \xi(t) + \varepsilon w(t,s) \;\geq\; \limsup -[\xi_\delta(s) - \xi_\delta(t)] \quad \forall (t,s) \quad \Rightarrow \quad \varepsilon \geq 0.$$

Rearranging, multiplying by -1 and replacing without any loss of generality ξ_δ with $\xi_\delta + \xi$ yields the equivalent condition

$$-\varepsilon w(t,s) \;\leq\; \liminf \xi_\delta(s) - \xi_\delta(t) \quad \forall (t,s) \quad \Rightarrow \quad \varepsilon \geq 0.$$

Therefore, (i) holds if and only if there exists a number $\varepsilon > 0$ and a net $\{\xi_\delta\}$ such that $\varepsilon w(t,s) \leq \liminf \xi_\delta(s) - \xi_\delta(t)$ for all (t,s). This last requirement is clearly equivalent to (ii) by dividing both sides by ε and replacing ξ_δ with $\varepsilon\xi_\delta$, as claimed. □

It is easy to see that (i) is necessary and sufficient for every undetectable deviation to be $\mathbf{x}$-unprofitable. Indeed, given $t \in T$ let $f_t \in \mathbb{R}^T$ be the indicator function of t, i.e., $f_t(s) = 1$ if $s = t$ and 0 otherwise. For sufficiency, if $g \in \mathbb{R}_+^{(T\times T)}$ satisfies $\sum_\ell \lambda_{k\ell} = 1$ for all k then g is a reporting strategy. If $Dg = \mathbf{0}$ then $\sum_k \lambda_{k\ell} = 1$, too, since $Dg(f_{t_\ell}) = \sum_k \lambda_{k\ell} - \lambda_{\ell k}$ for every $t_\ell \in \operatorname{supp}_T g$, so g is undetectable. Finally, it is clear that $g(w) \leq 0$ is equivalent to g being $\mathbf{x}$-unprofitable. For necessity, every $g \in \mathbb{R}_+^{(T\times T)}$ is proportional to a reporting strategy, and the value of Dg is determined by $Dg(f_{t_\ell})$ for every $t_\ell \in \operatorname{supp}_T g$, so if it is doubly stochastic then $Dg = \mathbf{0}$.

The last step in our proof of Theorem 1 is to show that (ii) implies implementability.

---

[2] $\mathbf{0} \in \mathbb{R}^{(T)}$ denotes the zero functional such that $\mathbf{0}(f) = 0$ for all $f \in \mathbb{R}^T$.

Lemma 3. *The following statements are equivalent:*
(i) *There exists a net* $\{\xi_\delta\}$ *such that* $w(t,s) \leq \liminf_\delta \xi_\delta(s) - \xi_\delta(t)$ *for all* (t,s).
(ii) *There exists an incentive scheme* ξ *such that* $w(t,s) \leq \xi(s) - \xi(t)$ *for all* (t,s).

Proof. That (ii) implies (i) is immediate. For the converse, without loss of generality we may fix any $t_0 \in T$ and assume that $\xi_\delta(t_0) = 0$ for all δ in the net, since it will not affect the any of the differences $\xi_\delta(s) - \xi_\delta(t)$. By hypothesis,

$$w(t,t_0) \;\leq\; \liminf \xi_\delta(t) \;\leq\; \limsup \xi_\delta(t) \;=\; -\liminf -\xi_\delta(t) \;\leq\; -w(t_0,t) \qquad \forall t \in T.$$

Hence, $\xi(t) = \liminf \xi_\delta(t)$ is bounded. Since the lim inf function is superadditive, it follows that $\liminf \xi_\delta(s) - \xi_\delta(t) + \liminf \xi_\delta(t) \leq \liminf \xi_\delta(s)$ for every (t,s). Hence, $\liminf \xi_\delta(s) - \xi_\delta(t) \leq \xi(s) - \xi(t)$. By (i), $w(t,s) \leq \liminf \xi_\delta(s) - \xi_\delta(t)$. Collecting these last two inequalities finally yields $w(t,s) \leq \xi(s) - \xi(t)$, as required. □

3 Rochet's Theorem

Let us compare Theorem 1 to Rochet's (1987) Theorem, which states that a given allocation $\mathbf{x}$ is implementable if and only if it is *cyclically monotone*, i.e., for every finite cycle $(t_1, \ldots, t_{m+1})$ such that $t_1 = t_{m+1}$,

$$\sum_{k=1}^{m} v(t_{k+1}, \mathbf{x}(t_k)) - v(t_k, \mathbf{x}(t_k)) \;\leq\; 0. \tag{4}$$

Rochet's proof of this result (adapted from Rockafellar, 1970) is remarkable not only for its simplicity, but also because it is constructive: if an allocation is implementable, the proof produces an incentive scheme that implements it. We include it below.

Proof of Rochet's Theorem. For sufficiency, suppose that $\mathbf{x}$ is implementable and let $(t_1, \ldots, t_{m+1})$ be a finite cycle, so $t_1 = t_{m+1}$. By hypothesis, there exists a scheme ξ such that $v(t_{k+1}, \mathbf{x}(t_k)) - v(t_{k+1}, \mathbf{x}(t_{k+1})) \leq \xi(t_k) - \xi(t_{k+1})$ for every $k \in \{1, \ldots, m\}$. Adding up all these inequalities yields $\sum_{k=1}^m v(t_{k+1}, \mathbf{x}(t_k)) - v(t_{k+1}, \mathbf{x}(t_{k+1})) \leq 0$, or equivalently $\sum_{k=1}^m v(t_{k+1}, \mathbf{x}(t_k)) - v(t_k, \mathbf{x}(t_k)) \leq 0$. Conversely, fix $t_0 \in T$ and define $U(t_0, t) = \sup \sum_{k=1}^m v(t_{k+1}, \mathbf{x}(t_k)) - v(t_k, \mathbf{x}(t_k))$, where the sup is with respect to all finite sequences $(t_1, \ldots, t_m)$ such that $t_1 = t_0$ and $t_m = t$. By cyclic monotonicity, $U(t_0, t_0) = 0$. Moreover, $U(t_0, t_0) \geq U(t_0, t) + v(t_0, \mathbf{x}(t)) - v(t, \mathbf{x}(t))$ for all t, so $U(t_0, t)$ is finite. Hence, $U(t_0, t) \geq U(t_0, s) + v(t, \mathbf{x}(s)) - v(s, \mathbf{x}(s))$ for all (t,s). Finally, if $\xi(t) = v(t, \mathbf{x}(t)) - U(t_0, t)$ then $v(t, \mathbf{x}(s)) - v(t, \mathbf{x}(t)) \leq \xi(s) - \xi(t)$. □

To relate cyclic monotonicity with Theorem 1, we will show that a cycle can be interpreted as an undetectable reporting strategy with rational probabilities. In so doing, we provide another characterization of implementability in terms of permutations, which may be thought of as undetectable "pure" reporting strategies.

A *finite permutation* (or simply *permutation*) is any map $\sigma : S \to S$ defined on some finite subset S of T such that σ is both one-to-one and onto. A permutation σ can be written as a pure reporting strategy, π_σ defined pointwise by $\pi_\sigma(s|t) = 1$ if $s = \sigma(t)$ and 0 otherwise. By virtue of σ being a permutation, it follows that for every $t \in S$, (i) there exists a unique $s \in S$ such that $\pi_\sigma(s|t) = 1$, and (ii) there exists a unique $s \in S$ such that $\pi_\sigma(t|s) = 1$. Therefore, π_σ is undetectable.

Corollary 1. *The following statements are equivalent for a given allocation* $\mathbf{x}$*:*
(i) *Every undetectable deviation is* $\mathbf{x}$*-unprofitable.*
(ii) $\mathbf{x}$ *is cyclically monotone.*
(iii) *Every permutation is* $\mathbf{x}$*-unprofitable.*

Proof. By Rochet's Theorem and Theorem 1, (i) is equivalent to (ii), and (i) is equivalent to (iii) by linearity of $g(w)$ with respect to $g \in \mathbb{R}^{(T\times T)}$ together with the Birkhoff-von Neumann Theorem, which states that the set of doubly stochastic matrices is the convex hull of the set of permutation matrices. □

It is instructive to consider a more direct argument for Corollary 1: (iii) implies (i) by the Birkhoff-von Neumann Theorem. (ii) implies (iii) because a permutation is a finite collection of cycles, each without repetitions, and cyclic monotonicity applied to a permutation implies that it is $\mathbf{x}$-unprofitable. Finally, (i) implies (ii) by representing a cycle as an undetectable reporting strategy with rational probabilities as follows.

Indeed, let $(t_1, \ldots, t_{m+1})$ be a cycle, so $t_1 = t_{m+1}$. Let $S = \{s_1, \ldots, s_\ell\}$, with $\ell \le m$, be the set of distinct elements in the cycle, and write $[s_j]$ for the number of times that s_j appears in $(t_1, \ldots, t_m)$. Also write $[s_i, s_j]$ for the number of times that s_i appears immediately before s_j in $(t_1, \ldots, t_{m+1})$. Let s_0 be any type that solves $[s_0] = \max_j [s_j]$, and define $\pi(s|t) = [t, s]/[s_0]$ if $s \neq t$ and $1 - \sum_{s\neq t}[t, s]/[s_0]$ otherwise. Clearly, π is a reporting strategy. To see that π is undetectable, notice that since $(t_1, \ldots, t_{m+1})$ is a cycle, $\sum_{s\neq t}[t, s] = \sum_{s\neq t}[s, t]$ for every t: the outflow from t equals the inflow to t. Finally, it is clear that by construction every element of π is a rational number, since every element is obtained as the difference between a natural number and the ratio of two natural numbers.

By Corollary 1, (iii) also characterizes implementability. Since the set of permutations are the extreme points of the set of doubly stochastic matrices, (iii) exploits linearity to provide this alternative characterization by just checking for unprofitability at the extreme points of the set of undetectable deviations. In this respect, (iii) is arguably more "efficient" than (ii) or (i), on the grounds that (iii) requires checking for unprofitability of a strict subset of the reporting strategies in (ii) or (i). In fact, (iii) checks for unprofitability of the smallest such subset of reporting strategies.

4 Implementing Allocations

It is not clear from the proof of Theorem 1 how to find a scheme that implements a given allocation. We fill this gap below. Using the notation developed in Section 2, consider the following linear programming problem given a pair of types (t_0, t_1):

$$V(t_0,t_1) \;=\; \inf_{\xi\in\mathbb{R}^T}\{\xi(t_1)-\xi(t_0) : v(t,\mathbf{x}(s))-v(t,\mathbf{x}(t)) \leq \xi(s)-\xi(t)\ \forall (t,s)\}. \qquad \text{(P)}$$

By incentive compatibility, $v(t_0,\mathbf{x}(t_1)) - v(t_0,\mathbf{x}(t_0)) \leq \xi(t_1)-\xi(t_0)$ for every ξ that implements $\mathbf{x}$. $V(t_0,t_1)$ denotes the value of (P). By definition, $V(t_0,t_1)$ is the greatest lower bound on $\xi(t_1)-\xi(t_0)$ subject to ξ implementing $\mathbf{x}$. This implies that $V(t_0,t_1) \geq v(t_0,\mathbf{x}(t)) - v(t_0,\mathbf{x}(t_0)) > -\infty$ and $V(t_0,t_0)=0$ if $\mathbf{x}$ is implementable.

Given $t_0 \in T$, let $\mathscr{F}(t_0) = \{\xi : \xi(t_0)=0 \text{ and } v(t,\mathbf{x}(s))-v(t,\mathbf{x}(t)) \leq \xi(s)-\xi(t)\ \forall(t,s)\}$ be the set of schemes ξ that implement $\mathbf{x}$ normalized so that $\xi(t_0)=0$. Endow $\mathscr{F}(t_0)$ with the pointwise order, i.e., $\xi \geq \zeta$ if and only if $\xi(t)\geq\zeta(t)$ for all $t\in T$.

Lemma 4. *The partially ordered set $(\mathscr{F}(t_0),\geq)$ is a complete lattice for every $t_0\in T$.*

Proof. It $\mathscr{F}(t_0)$ is empty then there is nothing to prove. Otherwise, $\mathbf{x}$ is clearly implementable. Let $\mathscr{G}$ be any subset of $\mathscr{F}(t_0)$. We must show that both $\bigwedge\mathscr{G}\in\mathscr{F}(t_0)$ and $\bigvee\mathscr{G}\in\mathscr{F}(t_0)$. By incentive compatibility, $v(t_0,\mathbf{x}(t))-v(t_0,\mathbf{x}(t_0))\leq\xi(t)$ for every $\xi\in\mathscr{G}$ and $t\in T$. Let $\xi_0(t) := \inf_\xi\{\xi(t):\xi\in\mathscr{G}\}$. By virtue of $\xi_0(t)$ being the greatest lower bound on $\{\xi(t):\xi\in\mathscr{G}\}$, it follows that $\xi_0(t)\geq v(t_0,\mathbf{x}(t))-v(t_0,\mathbf{x}(t_0))>-\infty$. Now, for any pair (t,s) and any $\xi\in\mathscr{G}$, $v(t,\mathbf{x}(s))-v(t,\mathbf{x}(t))\leq\xi(s)-\xi(t)$, therefore $v(t,\mathbf{x}(s))-v(t,\mathbf{x}(t))\leq\xi(s)-\xi_0(t)$, which implies $v(t,\mathbf{x}(s))-v(t,\mathbf{x}(t))\leq\xi_0(s)-\xi_0(t)$, since $\xi_0(t)$ is the greatest lower bound on $\{\xi:\xi\in\mathscr{G}\}$ and $v(t,\mathbf{x}(s))-v(t,\mathbf{x}(t))+\xi_0(t)$ is a lower bound on this set. Therefore, $\xi_0=\bigwedge\mathscr{G}\in\mathscr{F}(t_0)$. A proof that $\bigvee\mathscr{G}\in\mathscr{F}(t_0)$ uses a symmetric argument to the one above, and it is therefore omitted. $\square$

Proposition 1. *For every type t_0, the scheme $V(t_0, \cdot)$ implements $\mathbf{x}$, solves* (P) *above for all t_1, and satisfies $V(t_0, \cdot) = \bigwedge \mathscr{F}(t_0)$.*

Proof. This result follows immediately from Lemma 4. ☐

Using the notation of Section 2, it is easy to see that the dual of (P) is

$$W(t_0, t_1) \;=\; \sup_{g \geq 0} \{g(w) : Dg = \mathbf{1}_{t_1} - \mathbf{1}_{t_0}\}, \tag{D}$$

where $\mathbf{1}_s(t) = 1$ if $t = s$ and zero otherwise. Let $W(t_0, t_1)$ be the value of this problem. There clearly exists a feasible dual solution, hence $W(t_0, t_1) > -\infty$. By weak duality, $V(t_0, t_1) \geq W(t_0, t_1)$. Therefore, if $\mathbf{x}$ is implementable then the value of (the primal and hence also) the dual is finite for every choice of (t_0, t_1).

Theorem 2. *If $\mathbf{x}$ is implementable then $W(t_0, \cdot) \in \mathbb{R}^T$ implements $\mathbf{x}$ for every $t_0 \in T$.*

Proof. For any pair (t, s), let g be defined by $\operatorname{supp} g = \{(t, s)\}$ and $\lambda_{ts} = 1$, so $Dg = \mathbf{1}_s - \mathbf{1}_t$. By definition, $W(t, s) \geq g(w) = v(t, \mathbf{x}(s)) - v(t, \mathbf{x}(t))$. By revealed preference, $W(t, s) + W(s, r) \leq W(t, r)$, so $W(t, r) - W(t, s) \geq W(s, r)$. Therefore, $W(r, s) - W(r, t) \geq W(t, s) \geq v(t, \mathbf{x}(s)) - v(t, \mathbf{x}(t))$ for all (t, s, r). ☐

Theorem 2 derives a scheme to implement a given allocation from the value of (D). We will now show how it is related to the value of (P).

Corollary 2. *$V(t_0, t_1) = W(t_0, t_1)$ for every pair (t_0, t_1).*

Proof. If $\mathbf{x}$ is not implementable then $V(t_0, t_1) = \infty$. By Theorem 1, there is an $\mathbf{x}$-profitable, undetectable deviation h. Given a feasible dual solution g, $g + \alpha h$ is also feasible for any scalar $\alpha \geq 0$, and $(g + \alpha h)(w) \to \infty$ as $\alpha \to \infty$, so $W(t_0, t_1) = V(t_0, t_1)$. If $\mathbf{x}$ is implementable then $W(t_0, \cdot)$ implements $\mathbf{x}$ by Theorem 2. By weak duality, $W(t_0, t_1) \leq V(t_0, t_1)$, and since $\xi(\cdot|t_0)$ is a feasible primal solution with $W(t_0, t_0) = 0$, it follows that $W(t_0, t_1) \geq V(t_0, t_1)$. ☐

The results above yield an alternative, constructive proof of Theorem 1, stated below.

Corollary 3. *An allocation $\mathbf{x}$ is implementable if and only if $W(t, t) = 0$ for all t. In this case, the scheme $W(t, \cdot)$ implements $\mathbf{x}$ for every t.*

Proof. This result follows from Theorem 2 and Corollary 2. ☐

5 Revenue Equivalence

The linear programs of Section 4 suggest an immediate characterization of revenue equivalence. An (implementable) allocation exhibits *revenue equivalence* if any two incentive schemes that implement it differ by a constant, i.e., for any implementing schemes ξ and ζ there exists $c \in \mathbb{R}$ such that $\xi(t) = \zeta(t) + c$ for all $t \in T$.

By Lemma 4, an allocation exhibits revenue equivalence if and only if the set $\mathscr{F}(t)$ is a singleton for every type t. This is clearly equivalent to $\bigwedge \mathscr{F}(t) = \bigvee \mathscr{F}(t)$.

Lemma 5. *For every type t_1, the incentive scheme $-V(\cdot, t_1)$ implements $\mathbf{x}$, solves (P) above for all t_0, and satisfies $-V(\cdot, t_1) = \bigvee \mathscr{F}(t_1)$.*

Proof. Follows from Lemma 4 and Proposition 1. □

Theorem 3. *An allocation exhibits revenue equivalence if and only if for all (t_0, t_1),*

$$W(t_0, t_1) + W(t_1, t_0) \; = \; 0. \tag{5}$$

Proof. Follows immediately from Proposition 1, Lemma 5 and Corollary 2. □

Theorem 3 provides a dual characterization of revenue equivalence with the following strategic interpretation. $W(t_0, t_1)$ may be interpreted as the maximum profit from deviations that shift the same fixed probability mass from t_0 to t_1. Therefore, revenue equivalence holds if and only if for every (t_0, t_1), this profit is equal to the maximum profit from deviations that shift the same probability mass back from t_1 to t_0.

Mathematically, this result is equivalent to Theorem 1 of Heydenreich et al. (2009). However, the interpretation provided above for Theorem 3 is substantially different from theirs. Indeed, the interpretation here is strategic, whereas the interpretation offered by Heydenreich et al. (2009) is graph-theoretic. Specifically, Heydenreich et al. (2009) show that revenue equivalence is characterized by $U(t_0, t_1) = -U(t_1, t_0)$ for all (t_0, t_1), where U is defined in the proof of Rochet's Theorem (Section 3). It is not immediately clear how to interpret the optimization problem that defines U (although the discussion at the end of Section 3 suggests one). On the other hand, the dual problem that defines W carries the readily-available interpretation offered in the paragraph above. Finally, Theorem 1 of Heydenreich et al. (2009) does not apply to Bayesian implementation with many agents. This issue is discussed in detail in Section **??** below, which provides yet another characterization of revenue equivalence.

6 Ex Post Implementation

Let us extend the model to include several agents. Let I be a set of agents, and for every agent $i \in I$, let T_i be the set of i's possible types. Let $T = \prod_i T_i$ be the space of *type profiles*. Let $v_i(t, x)$ be the utility of agent i from social choice x if the type profile in society is t. An allocation $\mathbf{x} : T \to X$ is defined as before. An incentive scheme is now a map $\xi : I \times T \to \mathbb{R}$, and a mechanism is any pair $(\mathbf{x}, \xi)$.

It is well known that Rochet's Theorem can be extended in the context of many agents to so-called ex post implementation. Below we similarly extend Theorem 1, which is a trivial exercise given the previous results. We include it for completeness and because it will help to compare with results on Bayesian implementation and budget balance in subsequent sections.

Call $(\mathbf{x}, \xi)$ *ex post incentive compatible* (EPIC) if

$$v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t)) \ \leq \ \xi_i(s_i, t_{-i}) - \xi_i(t) \qquad\qquad \forall (i, t_i, s_i, t_{-i}).$$

An allocation $\mathbf{x}$ is *ex post implementable* if there is a scheme ξ that makes it EPIC. Intuitively, a mechanism is EPIC if for every agent, it is optimal to reveal one's true type after observing others' true types. Therefore, an EPIC mechanism will be incentive compatible regardless of one's beliefs about others, since the expected payoff from any reporting strategy is implied the state-by-state payoffs, where in this case a state is any profile of other agents' types.

A *reporting strategy* for agent i is now a map $\pi_i : S \to \Delta(S_i)$ defined on some finite subsets $S_i \subset T_i$ and $S \subset T$, where $\pi_i(s_i|t)$ is the probability i reports s_i when the profile of types is t. Now π_i is *undetectable* if $\sum_{s_i} \pi_i(s_i|t) = \sum_{s_i} \pi_i(t_i|s_i, t_{-i})$ for all t. Given an allocation $\mathbf{x}$, a reporting strategy is called *ex post* $\mathbf{x}$*-profitable* if

$$\sum_{(t_i, s_i)} \pi_i(s_i|t)[v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t))] \ > \ 0 \qquad\qquad \forall t_{-i} \in T_{-i}.$$

A deviation by agent i is any reporting strategy by i that isn't truthful. The following result is an immediate consequence of Theorem 1.

Corollary 4. *The following statements are equivalent for a given allocation* $\mathbf{x}$*:*
(i) $\mathbf{x}$ *is ex post implementable.*
(ii) *Every ex post* $\mathbf{x}$*-profitable deviation is detectable.*

7 Budget Balance

A scheme is *budget balanced* if $\sum_i \xi_i(t) = 0$ for every type profile t. An allocation $\mathbf{x}$ is *ex post implementable with budget balance* if there is a budget balanced scheme ξ such that $(\mathbf{x}, \xi)$ is EPIC. Below, we extend Theorem 1 to characterize budget balanced implementation. For this purpose, assume that $|I| = n \in \mathbb{N}$, i.e., there are finitely many agents. Although not essential, this assumption is useful for the sake of clarity.

A *finite strategy profile* (or simply *strategy profile*) is any family of reporting strategies $\pi = \{\pi_i : i \in I\}$. A *deviation profile* is any strategy profile with at least one deviation. A strategy profile π is called *unattributable* if

$$\sum_{s_i \in S_i} \pi_i(t_i|s_i, t_{-i}) \; = \; \sum_{s_j \in S_j} \pi_j(t_j|s_j, t_{-j}) \qquad\qquad \forall (i, j, t).$$

Otherwise, π is called *attributable.* Intuitively, a strategy profile is unattributable if the probability distribution over reported types is the same across players. After an unattributable deviation, even though the deviation may have been detected, it is impossible to identify the identity of the deviator or any non-deviator. Finally, a strategy profile π is *ex post* $\mathbf{x}$*-profitable* if

$$\sum_{(i,t,s_i)} \pi_i(s_i|t)[v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t))] \; > \; 0,$$

where the sum above is also taken with respect to the set of agents.

Theorem 4. *The following statements are equivalent for a given allocation* $\mathbf{x}$*:*
(i) $\mathbf{x}$ *is ex post implementable with budget balance.*
(ii) *Every ex post* $\mathbf{x}$*-profitable deviation profile is attributable.*

Proof. If T is finite then the result follows by a similar argument to the one used to prove Lemma 1. Let $R = \{(i, s_i, t) : i \in I,\; s_i \in T_i \text{ and } t \in T\}$. By a similar argument to that of Lemma 2, there exists a net of incentive schemes $\{\xi^\delta\}$ such that both $v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t)) \leq \liminf_\delta \xi_i^\delta(s) - \xi_i^\delta(t)$ for every (i, t_i, s_i, t_{-i}) and $\lim_\delta \sum_i \xi_i^\delta(t) = 0$ for all t (call this condition $(*)$) if and only if for every $\lambda \in \mathbb{R}_+^{(R)}$ and $\eta \in \mathbb{R}^{(T)}$, the system of equations given by $\sum_{s_i} [\lambda_i(s_i, t) - \lambda_i(t_i, s_i, t_{-i})] = \eta(t)$ for every (i, t) implies $\sum_{(i,s_i,t)} \lambda_i(s_i, t)[v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t))] \leq 0$ (call this condition $(**)$). Clearly, $(**)$ is equivalent to (ii). To see this, just divide every $\lambda_i(s_i, t)$ by $\Lambda = \max_{(i,t)} \sum_{s_i} \lambda_i(s_i, t)$ (if this equals zero then there's nothing to prove), as well as

$\eta(t)$, and replace $\lambda_i(t_i,t)$ with $\Lambda - \sum_{s_i} \lambda_i(s_i,t)$ for every (i,t). Now λ is proportional (with weight Λ) to an unattributable deviation profile that is also unprofitable. That (ii) implies $(**)$ is obvious. It remains to prove that $(*)$ is equivalent to (i). Again, that (i) implies $(*)$ is obvious. Conversely, let $\{\xi^\delta\}$ be a net that satisfies $(*)$. Fix any $t^0 \in T$. Given (i,t,δ), define the net $\{\zeta^\delta\}$ by $\zeta_i^\delta(t) = \xi_i^\delta(t) + \sum_{j \neq i} \xi_j^\delta(t_i^0, t_{-i})$. By $(*)$, $\lim_\delta \zeta_i^\delta(t_i^0, t_{-i}) = 0$ for all t_{-i}, and $v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t)) \leq \liminf_\delta \zeta_i^\delta(s) - \zeta_i^\delta(t)$ for every (i, t_i, s_i, t_{-i}). Hence, following the proof of Lemma 3, the scheme ζ defined by $\zeta_i(t) = \liminf_\delta \zeta_i^\delta(t) \in \mathbb{R}$ for every (i,t) ex post implements $\mathbf{x}$. Let $\{\zeta^\gamma\}$ be a subnet of $\{\zeta^\delta\}$ such that $\lim_\gamma \zeta_i^\gamma(t) = \zeta_i(t)$ for all (i,t). One such subnet exists by definition of lim inf and Lemma 4. Finally, for every $i_1 \in I$ and $t \in T$ let

$$\zeta_{i_1}^0(t) \;=\; \zeta_{i_1}(t) - \sum_{i_2 \neq i_1} \zeta_{i_2}(t_{i_1}^0, t_{-i_1}) + \sum_{i_3 \neq i_2} \zeta_{i_3}(t_{i_1}^0, t_{i_2}^0, t_{-i_1 i_2}) - \cdots + \sum_{i_n \neq i_{n-1}} \zeta_{i_n}(t_{-i_n}^0, t_{i_n}).$$

Clearly, ζ^0 ex post implements $\mathbf{x}$ because ζ does, too, since for all (i,t), $\zeta_i^0(t)$ equals $\zeta_i(t)$ plus something that does not depend on t_i. By construction, it is easy to see that the scheme ζ^0 also satisfies budget balance, since

$$\begin{aligned}
\sum_{i_1 \in I} \zeta_{i_1}^0(t) &= \sum_{i_1 \in I} \zeta_{i_1}(t) - \sum_{i_2 \neq i_1} \zeta_{i_2}(t_{i_1}^0, t_{-i_1}) + \sum_{i_3 \neq i_2} \zeta_{i_3}(t_{i_1}^0, t_{i_2}^0, t_{-i_1 i_2}) - \cdots + \sum_{i_n \neq i_{n-1}} \zeta_{i_n}(t_{-i_n}^0, t_{i_n}) \\
&= \lim \sum_{i_1 \in I} \zeta_{i_1}^\gamma(t) - \sum_{i_2 \neq i_1} \zeta_{i_2}^\gamma(t_{i_1}^0, t_{-i_1}) + \sum_{i_3 \neq i_2} \zeta_{i_3}^\gamma(t_{i_1}^0, t_{i_2}^0, t_{-i_1 i_2}) - \cdots + \sum_{i_n \neq i_{n-1}} \zeta_{i_n}^\gamma(t_{-i_n}^0, t_{i_n}) \\
&= \lim \sum_{i_1 \in I} \xi_{i_1}^\gamma(t) + \sum_{i_2 \neq i_1} \zeta_{i_2}^\gamma(t_{i_1}^0, t_{-i_1}) - \sum_{i_2 \neq i_1} \zeta_{i_2}^\gamma(t_{i_1}^0, t_{-i_1}) + \cdots \\
&\quad - \sum_{i_n \neq i_{n-1}} \zeta_{i_n}^\gamma(t_{-i_n}^0, t_{i_n}) + \sum_{i_n \neq i_{n-1}} \zeta_{i_n}^\gamma(t_{-i_n}^0, t_{i_n}) + \sum_{j \neq i_n} \xi_j^\gamma(t^0) \;=\; 0.
\end{aligned}$$

Therefore, ζ^0 ex post implements $\mathbf{x}$ with budget balance. □

The results of Sections 4 and 5 extend easily to the budget balanced setting after suitable modifications. Specifically, define $V_i(t_i^0, t_i^1; t_{-i}) = \inf \xi_i(t_i^1, t_{-i}) - \xi_i(t_i^0, t_{-i})$, where the infimum is taken with respect to incentive schemes that ex post implement $\mathbf{x}$ with budget balance. By the same argument as in Lemma 4, for any $t^0 \in T$, the set of all schemes ξ that implement $\mathbf{x}$ with budget balance and satisfy $\xi_i(t_i^0, t_{-i}) = 0$ for all t_{-i} is a complete lattice. Therefore, given $t^0 \in T$, the scheme $\xi_i(t) = V_i(t_i^0, t_i; t_{-i})$ ex post implements $\mathbf{x}$ with budget balance. This leads to the dual problem below.

$$W_i(t_i^0, t_i^1; t_{-i}) \;=\; \sup_{\lambda \geq 0, \eta} \{\lambda(w) : D\lambda = \eta + \mathbf{1}_{t_i^1}^i - \mathbf{1}_{t_i^0}^i\},$$

where $\mathbf{1}_{t_i}^i \in \mathbb{R}^{I \times T}$ is defined by $\mathbf{1}_{t_i^0}^i(j,t) = 1$ if $j = i$ and $t_i = t_i^0$, and zero otherwise.

Now Proposition 1, Theorem 2, Corollaries 2 and 3, as well as Lemma 5 easily extend with budget balance. Hence, Theorem 4 generalizes cyclic monotonicity to characterize budget-balanced implementation as follows: $W_i(t_i, t_i; t_{-i}) = 0$ for all (i, t).

Finally, Theorem 3 also extends to the case of budget balance with the following generalization. An allocation $\mathbf{x}$ has *budget-balanced revenue equivalence* if for any two schemes ξ and ζ that ex post implement $\mathbf{x}$ with budget balance, $\xi_i(t) = \zeta_i(t) + c_i(t_{-i})$ for all (i, t) and some $c_i(t_{-i}) \in \mathbb{R}$. (Hence, $\sum_i c_i(t_{-i}) = 0$ for all t.)

Corollary 5. *An allocation has budget-balanced revenue equivalence if and only if*

$$W_i(t_i^0, t_i^1; t_{-i}) + W_i(t_i^1, t_i^0; t_{-i}) \;=\; 0 \qquad \forall (i, t_i^0, t_i^1, t_{-i}).$$

The interpretation behind Corollary 5 is almost identical to that of Theorem 3. The only difference is that "attributable" now replaces "detectable." The amount $W_i(t_0, t_1; t_{-i})$ corresponds to the maximum profit from a strategy profile that is unattributable except for a given agent i, whose strategy changes the probability distribution over reports by taking probability mass from t_i^0 to t_i^1. Revenue equivalence with budget balance is equivalent to this profit plus the maximum profit from taking the probability mass back from t_i^1 to t_i^0 being always equal to zero.

Corollary 5 states that an allocation satisfies budget-balanced revenue equivalence if and only if the profit from any strategy profile that is unattributable except for a given agent i, whose strategy changes the probability distribution over reports

– To be completed. –

A Appendix

This appendix presents ancillary results that are used in the main body of the paper. Let us begin with Clark's (2006) extension of The Theorem of the Alternative.

Let X and Y be ordered, locally convex real vector spaces, with positive cones X_+ and Y_+ and topological dual spaces X^* and Y^* such that $X^{**} = X$ and $Y^{**} = Y$. Let $A : X \to Y$ be a continuous linear operator with adjoint operator $A^* : Y^* \to X^*$ and fix any $b \in Y$. Finally, for any set S let $\overline{S}$ denote its closure.

Lemma A.1 (Clark, 2006, page 479)**.** *For any $b \in Y$, there exists $x \in X_+$ such that $A(x) = b$ if and only if $A^*(y_0^*) \in \overline{X_+^* - \{A^*(y^*) : y^*(b) = 0\}}$ implies that $y_0^*(b) \geq 0$.*

Now consider the characterization of strong duality by Gretsky et al. (2002). With the same notation as in the previous result, a *linear program* is any triple (A, b, c^*) such that A is as above, $b \in Y$ and $c^* \in X^*$. The *primal* is given by the linear optimization problem $\sup\{c^*(x) : A(x) \leq b, x \geq 0\}$, and the *dual* by $\inf\{y^*(b) : A^*(y^*) \leq c^*, y^* \geq 0\}$. Say that *there is no duality gap* if the value of the primal equals the value of the dual. Denote by $V(b)$ the value of the primal as a function of b. The *subdifferential* of V at b is the set $\partial V(b) = \{y^* : V(y) - V(b) \leq y^*(y - b)\ \forall y \in Y\}$. V is *subdifferentiable* at b if $\partial V(b) \neq \emptyset$.

Lemma A.2 (Gretsky et al., 2002, page 265)**.** *Both the dual has a solution and there is no duality gap if and only if V is subdifferentiable at b.*

References

Clark, S. A. (2006): "Necessary and Sufficient Conditions for Infinite-Dimensional Linear Inequalities," *Positivity*, 10, 475–489. 13

Gretsky, N., J. Ostroy, and W. Zame (2002): "Subdifferentiability and the Duality Gap," *Positivity*, 6, 1–16. 13

Heydenreich, B., R. Müller, M. Uetz, and R. Vohra (2009): "Characterization of Revenue Equivalence," *Econometrica*, 77, 307–316, mimeo. 8

Holmes, R. B. (1975): *Geometric Functional Analysis and its Applications*, New York/Heidelberg/Berlin: Springer-Verlag. 2

Rochet, J. C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191–200. 4, 8, 9

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton, New Jersey: Princeton University Press. 2, 4