Detecting Profitable Deviations

David Rahman*
University of Minnesota

First draft: April 16, 2009
This version: February 4, 2011

Abstract

In this paper I offer necessary and sufficient conditions for implementability in a quasi-linear principal-agent model with arbitrary type spaces. I extend Rochet's Theorem by allowing the principal to observe information that may be correlated with the agent's type. By viewing the agent's gains from misreporting as payments in a hypothetical zero-sum game, I show that an allocation is implementable if and only if every infinitesimally detectable deviation is at most infinitesimally profitable. This leads to generalizations of existing results, such as revenue equivalence and implementation with moral hazard, as well as new results altogether, such as budget balanced implementation (interim and ex post), existence of bargaining solutions and revealed stochastic preference.

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

*I owe many thanks to Sushil Bikhchandani and Christoph Müller for numerous conversations and encouragement. Please send any comments to dmr@umn.edu.

1 Introduction

Understanding implementable allocations is an essential topic of mechanism design. This issue has been addressed in general (e.g., Rochet, 1987; Heydenreich et al., 2009) as well as specific contexts (e.g., Myerson, 1981; Bikhchandani et al., 2006). In this paper I restrict attention to the quasi-linear setting but otherwise keep the model as general as possible. I consider a principal-agent problem where the agent has some private information—a type—which the principal solicits. To induce honest reporting, the principal makes payments to the agent contingent on the agent's reported type and a signal that may be correlated with the agent's actual type. The main result of the paper is a characterization of implementable allocations on arbitrary type and signal spaces without placing any restrictions on utilities (apart from integrability). I also explore interesting extensions.

Such an extended principal-agent problem with additional signals for the principal is interesting for at least two reasons. Firstly, it is important to understand when the principal can exploit relevant information to induce honest reporting. For instance, the signal may be output, other agents' reports, even future reports in a dynamic environment.[1] It might be argued that correlation renders the problem trivial because it leads to full surplus extraction, but this argument is flawed: Example 5 provides a basic environment with correlation but where full surplus extraction is impossible, rendering the implementability problem still relevant.[2] Secondly, the elementary network approach emphasized by Rochet (1987) and Heydenreich et al. (2009) does not generalize easily. For instance, their approach cannot be used to characterize *interim* implementation (as opposed to ex post) relative to the additional signal—the subject of this paper—because the network structure on which they rely is quite simply absent in the present context. On the other hand, the duality-based approach that I develop in this paper to characterize implementability generalizes easily to richer environments, including budget constraints and bargaining.[3] See Section 5.

I characterize implementability with the following intuition, which I also deem to be an important contribution of the paper because it helps to understand the economics of incentive compatibility—just like the planner's problem helps to understand its decentralization through prices. Here it is:

[1] I develop the application to dynamic implementation in Rahman (2009).

[2] I provide a detailed study of full surplus extraction in Rahman (2010).

[3] Other extensions, such as limited liability, can be easily included in the present framework, too.

Consider the following hypothetical zero-sum game between a principal and an agent. The principal chooses a scheme of report- and signal-contingent money payments and the agent chooses a reporting strategy, i.e., a probability distribution over type-report pairs. The principal pays the agent the deviation gains from reporting according to the chosen strategy rather than honestly. These gains arise from both changes in the allocation (call these "gross" gains) and changes in money payments as a result of misreporting. By definition, an allocation is implementable if for some payment scheme the agent cannot make positive deviation gains. Since the agent can guarantee non-negative gains by reporting truthfully, implementability is equivalent to both the principal and agent receiving a payoff of zero, in other words, the value of this hypothetical game is zero. If this game is finite then by the Minimax Theorem it doesn't matter who goes first. Hence, an allocation is implementable if and only if for any reporting strategy for the agent there is a payment scheme that makes it unprofitable. However, the crucial insight implied by the Minimax Theorem is that different payment schemes may be used to discourage different reporting strategies.[4]

If a given strategy is detectable, i.e., the probability distribution over signals given reported types differs from that given actual types, it is easy to find a scheme that discourages it. A reported type-signal pair whose probability is greater than that with the actual type-signal pair is dubbed "bad news," and otherwise "good news." Now pay the agent for good news, charge him for bad news, and increase the wedge between good and bad news until any utility gains are outweighed by associated monetary losses. If the strategy is undetectable then the agent receives the same expected payment as if he reported truthfully, regardless of the payment scheme. Hence, deviation gains are non-positive if and only if gross gains are non-positive. Call such a strategy "unprofitable." This yields the finite version of Theorem 1: an allocation is implementable if and only if every undetectable deviation is unprofitable.

When the set of types is infinite, Theorem 1 requires a slightly stronger condition. Mathematically, to apply the Minimax argument above I require a version of bounded steepness (Gale, 1967). Strategically, an allocation is interim implementable if and only if every infinitesimally detectable deviation is at most infinitesimally profitable. Intuitively, the problem is this. Consider a sequence of asymptotically undetectable strategies. If their gross gains are positively bounded below then there exists an "infinitesimal" deviation with positive profit, and this cannot be discouraged with a payment scheme that also discourages the non-infinitesimal deviations.

[4] This argument is reminiscent of Hart and Schmeidler (1989). I discuss the connections below.

A "windfall" advantage of this result is that it generalizes easily. This contrasts, for instance, Rochet's (1987) Theorem, which is closely related (see Proposition 4). To illustrate this advantage, I extend Theorem 1 in a number of interesting directions. In Section 5, I characterize budget-balanced implementation (interim and ex post), existence of bargaining solutions even with correlated types,[5] implementability with moral hazard, revenue equivalence even if types are correlated,[6] and revealed stochastic preference.[7] As a corollary to the duality approach explored in this paper, I also obtain a subdifferential characterization of the set of implementing payment schemes for a given allocation, and offer some intuition behind it.[8]

The paper is organized as follows. In the next section, I present the principal-agent model and formally state Theorem 1, which characterizes implementable allocations. I then provide four examples to illustrate the significance of every infinitesimally detectable deviation being at most infinitesimally profitable. For instance, Example 4 shows that implementability does not require payoffs to be Lipschitz, unlike Carbajal and Ely (2010).[9] In Section 3, I prove Theorem 1 following the basic intuition above, using duality in the form of an infinite-dimensional version of the Minimax Theorem. In Section 4, I mostly place my results in relation to the literature. Specifically, I discuss Rochet's Theorem, Cremer and McLean's characterization of full surplus extraction, and Hart and Schmeidler's proof of existence of correlated equilibrium using the Minimax Theorem. I also establish revenue equivalence in this paper's richer setting, compare it to previous recent characterizations, and characterize the set of implementing payments for an implementable allocation. Section 5 derives the extensions described in the previous paragraph and Section 6 concludes. Omitted proofs and ancillary results can be found in the appendices at the end of the paper.

---

[5]This result contrasts the work of Segal and Whinston (2009), who assume independent types.

[6]I characterize when there exists a unique contingent payment scheme modulo a constant. By contrast, Heydenreich et al. (2009) only characterize unique interim expected payments assuming that types are independent, in a similar spirit to Müller et al. (2007).

[7]Although in this paper I restrict attention to revealed stochastic preference under the assumption of quasi-linearity (this assumption can be relaxed, see Afriat, 1967, p. 72), I otherwise generalize the work of McFadden and Richter (1990) and McFadden (2005) in several ways. First, I allow for prices, or "budgets" to vary randomly in a way that may be correlated with stochastic preferences. This allows for a richer interpretation of stochastic preferences as those of a population. Secondly, I do not restrict attention to compact metric type spaces.

[8]For a related result in the case of ex post implementation, see Kos and Messner (2009).

[9]Although there are similarities between Kos and Messner (2009), Carbajal and Ely (2010) and this paper, there are also important differences. Firstly, the other papers do not allow correlated signals. Secondly, even in their restricted setting they do not emphasize a strategic interpretation.

2 Model

Consider the following relatively standard mechanism design environment. There is an agent with private information and a principal who solicits this information from the agent. The agent sends a message to the principal which may or may not be truthful. The principal subsequently observes a verifiable signal possibly correlated with the agent's information, such as output or other agents' types.

Formally, let T be an arbitrary set with typical element t, interpreted as the collection of all *types* for the agent that the principal deems possible. Let X be a nonempty set of outcomes and $(Y, \mathscr{Y})$ a measurable space of possible signals that the principal may observe. For each type t, let $p(t)$ be a finitely additive probability measure on $\mathscr{Y}$ describing the likelihood of signals given t. Let $u(t, x, y) \in \mathbb{R}$ be the agent's utility from choice x when his type is t and the realized signal is y. An *allocation* is a map $\mathbf{x} : T \times Y \to X$, where $\mathbf{x}(t, y)$ represents the choice made by the principal when the agent's report is t and the realized signal is y. An *incentive scheme* (or simply *scheme*) is a map $\xi : T \times Y \to \mathbb{R}$, where $\xi(t, y)$ represents the payment from the agent to the principal when his report is t and the realized signal is y. An incentive scheme is denominated in money, which enters the agent's utility linearly with unit marginal utility, as usual. A *mechanism* is any pair $(\mathbf{x}, \xi)$ as above.

The expected utility to the agent from an allocation $\mathbf{x}$ when his type is t—assuming that he tells the truth—is given by

$$v(t) \;=\; \int_Y u(t, \mathbf{x}(t, y), y) p(dy|t),$$

and the expected utility gain from reporting s when his type is actually t is given by

$$\Delta v(t, s) \;=\; \int_Y [u(t, \mathbf{x}(s, y), y) - u(t, \mathbf{x}(t, y), y)] p(dy|t).$$

For the functions above to be well-defined, we must impose some restrictions on v and $\mathbf{x}$. Otherwise, the integrals above may not exist.

Assumption 1. Both $v(t)$ and $\Delta v(t, s)$ are well-defined and real-valued for all (t, s).

I shall maintain this assumption throughout. One way to guarantee that it holds is to assume that $v(t, \mathbf{x}(s, y), y)$ is a bounded, $\mathscr{Y}$-measurable function of y for all (t, s). But this is certainly not the only way.

Definition 1. The mechanism $(\mathbf{x}, \xi)$ is called *incentive compatible* if

$$\Delta v(t,s) \ \leq \ \int_Y [\xi(s,y) - \xi(t,y)] p(dy|t) \qquad \forall (t,s). \tag{1}$$

An allocation $\mathbf{x}$ is called *implementable* if there exists a scheme ξ such that $(\mathbf{x}, \xi)$ is incentive compatible. In this case, say ξ *implements* $\mathbf{x}$.[10]

Intuitively, incentive compatibility guarantees that any utility gain with respect to the allocation from reporting type s when the agent's true type equals t is outweighed by the associated loss from monetary transfers. Since the inequalities above apply after the agent learns his type, any prior beliefs are irrelevant for implementability. Put differently, implementability holds or fails regardless of prior beliefs.

Just as before, for the inequalities above to be well-defined, we must impose some restrictions on ξ. Otherwise, the integrals above may not exist.

Assumption 2. Every scheme ξ satisfies the following property: $\xi(t,y)$ is a bounded $\mathscr{Y}$-measurable function of y for all t, i.e., $\xi \in B(Y)^T$.

I shall also maintain this assumption throughout. Whereas Assumption 1 is relatively uncontroversial, Assumption 2 has a little more content. It reflects a trade-off between restrictions on p versus ξ. For the inequalities defining incentive compatibility above to be well-defined, $\xi(s)$ must be $p(t)$-integrable for all (t,s). Bounded measurability guarantees this without imposing restrictions on p. (The literature typically assumes this and much more, see McAfee and Reny, 1992, for instance.) On the other hand, if we assumed that the set of all $p(t)$-integrable functions $L(t)$ was the same set L for all t then we could relax Assumption 2 by requiring only that $\xi \in L^T$. Alternatively, we might also assume that $\xi \geq 0$ (or more generally that ξ is bounded below by a given integrable function) and $\xi(t) \in L(t)$ for all t, but this would require imposing restrictions on the function u for implementability to be feasible.

My main goal is to characterize implementability, i.e., find a necessary and sufficient condition in terms of v and p. Unfortunately, the elementary approach pioneered by Rochet (1987) does not apply here because his characterization relies on the network structure inherent in his environment, where Y is a singleton set. I develop below an alternative approach based on linear duality, expressed in terms of the Minimax Theorem to facilitate a strategic interpretation.

[10] This definition strictly generalizes Rochet's, and coincides with it when Y is a singleton set.

To this end, I will describe a zero-sum two-person game between the principal and the agent, where the principal chooses an incentive scheme, the agent chooses a "reporting strategy," and the agent's payoff equals the expected utility difference between the chosen strategy and reporting truthfully. I will show that implementability is equivalent to truth-telling being a Nash equilibrium of this alternative, hypothetical game. In this game, the agent's pure strategies are the set of possible type-report pairs. Since the set of types has no given structure, the only meaning mixed strategies are convex combinations of pure strategies. Formally, the agent chooses a so-called *reporting strategy* (or simply *strategy*) π in $\Delta(T \times T)$, the set of non-negative $T \times T$ matrices with finite support whose entries add up to one. The amount $\pi(t, s)$ may be interpreted as the probability that the agent's report equals s when his true type equals t. A *deviation* is any reporting strategy that isn't truthful, i.e., it lies with positive probability conditional on some type.

Definition 2. A strategy $\pi \in \Delta(T \times T)$ is called *undetectable* if

$$\sum_{s \in T} \pi(t, s) p(t) \;=\; \sum_{s \in T} \pi(s, t) p(s) \qquad \forall t \in T. \tag{2}$$

Otherwise, π is *detectable*. It is called *profitable* if $\Delta v(\pi) = \sum_{(t,s)} \Delta v(t, s) \pi(t, s) > 0$.

Intuitively, π is undetectable if the joint probability distribution over reports and the signal coincides with that of actual types and the signal.

It is relatively easy to see that a necessary condition for implementability is that every profitable deviation be detectable. To see this, apply any undetectable deviation to the inequalities in (1) above. The difference in payments on the right-hand side of (1) disappear, implying that the deviation is unprofitable.

It might be conjectured that detecting profitable deviations is not only necessary, but also sufficient for implementability. As I argue later, this is in fact the case when the set of types T is finite, when $p(t)$ does not depend on t, and also when Y is a singleton set, the latter of course being the context for Rochet's Theorem. However, this conjecture is generally false, hence detecting profitable deviations is necessary but not sufficient for implementation. Examples 1–4 below illustrate this lack of sufficiency. Intuitively, characterizing implementability requires in addition that infinitesimally detectable deviations be at most infinitesimally profitable.

Formally, let $\Delta p(\pi)$ be defined pointwise by

$$\Delta p(\pi)(t) \;=\; \sum_{s \in T} \pi(t, s) p(t) - \pi(s, t) p(s) \qquad \forall t \in T.$$

Definition 3. *Every infinitesimally detectable deviation is at most infinitesimally profitable* if "every profitable deviation is uniformly detectable," i.e.,

$$\mathscr{D} \ := \ \sup_{\pi} \frac{\Delta v(\pi)}{|\Delta p(\pi)|} \ < \ +\infty,$$ [11]

where π is a reporting strategy, $|\Delta p(\pi)| = \sum_t \|\Delta p(\pi)(t)\|$, and for each t, $\|\Delta p(\pi)(t)\|$ is the total variation norm on the space of (bounded additive) measures on Y.

Intuitively, Definition 3 requires not only that every profitable deviation be detectable, but also that every deviation's profitability be uniformly bounded by its detectability. Mathematically, it is a bounded steepness condition, ensuring that linear duality applies to an agent's deviation gains.

Theorem 1. *A given allocation is implementable if and only if every infinitesimally detectable deviation is at most infinitesimally profitable.*

Before proving Theorem 1, let me discuss the differences between Definition 3 and detecting profitable deviations in the context of some illustrative examples. Notice that all the examples below (Examples 1–4) exhibit the fact that every profitable deviation is detectable, yet implementability is impossible.

Example 1. Let $T = [0, 1]$ and $Y = \{0, 1\}$. Define $p(0) = [0]$, $p(1) = [1]$ and $p(t) = \frac{1}{2}[0] + \frac{1}{2}[1]$ for all $t \in (0, 1)$, where $[z]$ stands for Dirac measure.[12] For every $t \in (0, 1)$, let π_t be the strategy defined pointwise by $\pi_t(t, 0) = \frac{t}{2} = \pi_t(t, 1)$, $\pi_t(0, t) = \frac{t}{2} = \pi_t(1, t)$ and $\pi_t(t, t) = 1 - t$. It is not difficult to see that $|\Delta p(\pi_t)| = \frac{t}{2} \|[0] - [1]\| = t$, from which it follows that $|\Delta p(\pi_t)| \to 0$ as $t \to 0$. Define $\Delta v(t, 0) = 1/t$ for every $t \in (0, 1)$ and $\Delta v(r, s) = 0$ for all other (r, s). Clearly, every profitable deviation is detectable, since making any profit requires type 0 misreporting to some type $t \in (0, 1)$ with positive probability, and this is detectable. Now the profit from π_t is given by $\Delta v(\pi_t) = \frac{t}{2}[\Delta v(t, 0) + \Delta v(t, 1) + \Delta v(0, t) + \Delta v(1, t)] = \frac{1}{2}$ for every $t \in (0, 1)$. As a result, $\mathscr{D} = +\infty$, so implementability fails by Theorem 1. Notice that this argument fails if and only if $\Delta v(t, 0)$ is uniformly bounded above for all t.

Example 1 shows that Definition 3 and detecting profitable deviations are different in an important, non-pathological way. Although the example above relies on Δv becoming unbounded, this is by no means a prerequisite for the kind of failure of implementability that it portrays, as the next examples show.

[11] I adopt the following conventions: (i) zero divided by zero equals zero and (ii) any non-zero real number divided by zero equals $\pm\infty$ depending on the sign of the numerator.

[12] In other words, $[z](Z) = 1$ if $z \in Z$ and 0 otherwise.

Example 2. Let $T = [0,1]$, $Y = \{0,1\}$ and $p(t) = (1-t)[0] + t[1]$ for all t. Given y and any finite subset of types there is only one type with largest probability over y, so every deviation is detectable. For each $k \in \mathbb{N}$, let $t_k = 1/k$ and define π_k by $\pi_k(t_k, 0) = \frac{1}{2}(1-t_k) = \pi_k(0, t_k)$, $\pi_k(t_k, 1) = \frac{1}{2}t_k = \pi_k(1, t_k)$. (Let π_k be 0 elsewhere.) By routine calculations, $|\Delta p(\pi_k)| = \frac{1}{2k}(1-\frac{1}{k}) \,\|[0] + [1]\| = \frac{1}{k}(1-\frac{1}{k})$. Define Δv by $\Delta v(t, 0) = 1$ for all t and $\Delta v(t, s) = 0$ for all other (t, s). Clearly, $\Delta v(\pi_k) = (1-\frac{1}{k})\Delta v(t_k, 0) = (1-\frac{1}{k})$. Finally, $\lim \Delta v(\pi_k)/|\Delta p(\pi_k)| = \lim(1-\frac{1}{k})/(\frac{1}{k}(1-\frac{1}{k})) = \lim k = +\infty$. Therefore, $\mathscr{D} = +\infty$ and implementability fails by Theorem 1.

Example 2 above shows that a suitable discontinuity in Δv is sufficient to prevent implementability. Indeed, notice that in the example $\Delta v(t, 0)$ does not tend to 0 as $t \to 0$, even though $\Delta v(0, 0) = 0$. However, discontinuity is not necessary for implementation to fail, as the next example shows.

Example 3. Consider exactly the same setting and sequence $\{\pi_k\}$ as in Example 2. The only difference here is that now Δv is defined by $\Delta v(0, t) = \sqrt{t}$. It is easy to see that now $\Delta v(\pi_k) = (1-\frac{1}{k})\Delta v(t_k, 0) = (1-\frac{1}{k})\frac{1}{\sqrt{k}}$. Simple calculations show that $\lim \Delta v(\pi_k)/|\Delta p(\pi_k)| = \lim((1-\frac{1}{k})\frac{1}{\sqrt{k}})/(\frac{1}{k}(1-\frac{1}{k})) = \lim \sqrt{k} = +\infty$. Therefore, $\mathscr{D} = +\infty$ and again implementability fails by Theorem 1.

Example 3 shows that a failure of Lipschitz continuity in Δv is enough for implementation to fail. However, yet again this is not necessary. The next example highlights that what drives all these failures is not failure of Lipschitz continuity, but rather a lack of bounded steepness between the change in probabilities and the change in payoffs from misreporting, as Theorem 1 shows.

Example 4. Let $T = [0,1]$, $Y = \{0,1\}$ and $p(t) = (1-t^2)[0] + t^2[1]$ for all t. As in Example 2, every deviation is detectable. Given $k \in \mathbb{N}$, let $t_k = 1/k$ and π_k be the strategy defined by $\pi_k(t_k, 0) = \frac{1}{2}(1-t_k^2) = \pi_k(0, t_k)$ and $\pi_k(t_k, 1) = \frac{1}{2}t_k^2 = \pi_k(1, t_k)$. By routine calculations, $|\Delta p(\pi_k)| = \frac{1}{k^2}(1-\frac{1}{k^2})$. Define Δv by $\Delta v(t, 0) = t$ for all t and 0 elsewhere. Clearly, $\Delta v(\pi_k) = (1-\frac{1}{k^2})\Delta v(t_k, 0) = (1-\frac{1}{k^2})\frac{1}{k}$. After simple calculations, $\lim \Delta v(\pi_k)/|\Delta p(\pi_k)| = \lim((1-\frac{1}{k^2})\frac{1}{k})/(\frac{1}{k^2}(1-\frac{1}{k^2})) = \lim k = +\infty$. Therefore, $\mathscr{D} = +\infty$ and once again implementability fails by Theorem 1.

Example 4 exhibits a Lipschitz continuous function Δv yet implementation fails, even though every deviation is detectable. Intuitively, this happens here because the "steepness" ratio of changes in payoffs (linear) to changes in probabilities (quadratic) explodes as the deviation becomes infinitesimal. By Theorem 1, implementation is equivalent to this steepness being uniformly bounded.

3 Proof of Theorem 1

Below, I prove Theorem 1 in three steps. I begin by describing implementability as the truthful equilibrium of a zero-sum two-person game. I then characterize existence of such an equilibrium in terms of subdifferentiability of a function induced by the game's payoffs. Finally, I show that such subdifferentiability is equivalent to every infinitesimally detectable deviation being at most infinitesimally profitable.

In this zero-sum game, the principal chooses a scheme $\xi \in B(Y)^T$ and the agent a strategy $\pi \in \Delta(T \times T)$. I will explain later how to incorporate prior beliefs. The principal pays the agent the following amount:

$$F(\xi, \pi) \;=\; \sum_{(t,s)} \pi(t,s)[\Delta v(t,s) - \int_Y (\xi(s,y) - \xi(t,y))p(dy|t)].$$

Clearly, the integral above is well defined, and $F(\xi, \pi) \in \mathbb{R}$ for every (ξ, π). Intuitively, F is (proportional to) the agent's expected deviation gain from choosing π when the principal chooses ξ. This defines a zero-sum two-person game.

In this hypothetical game, the principal pays the agent the deviation gain defined by F. A pair (ξ, π) is called an *equilibrium* of F if it is a Nash equilibrium, i.e.,

$$F(\xi, \pi') \;\leq\; F(\xi, \pi) \;\leq\; F(\xi', \pi) \qquad \forall (\xi', \pi').$$

An equilibrium (ξ, π) is *honest* if π is honest, i.e., $\pi(t,s) > 0$ if and only if $t = s$. For any such honest equilibrium, clearly $F(\xi, \pi) = 0$.

Proposition 1. $\mathbf{x}$ *is implementable if and only if* F *has an honest equilibrium.*

Proof. The proof is rather simple. Suppose that $\mathbf{x}$ is implementable with scheme ξ. By definition, for any honest strategy π, the pair (ξ, π) is an honest equilibrium of F that pays 0 to the agent in the hypothetical zero-sum game. Conversely, if $\mathbf{x}$ is not implementable then for any ξ there is a strategy π for the agent that yields positive profit, given by the matrix $\delta_{(t,s)}$ that equals one at a given violated incentive constraint (t, s) and zero elsewhere. Hence, an honest equilibrium cannot exist. □

The simple observation in Proposition 1 motivates characterizing implementability with conditions for existence of a saddle point with zero saddle value. This approach yields the proof of Theorem 1 that I now develop. Afterwards, I will discuss this argument at some length.

To this end, define two quantities:

$$F_* = \sup_\pi \inf_\xi F(\xi,\pi), \qquad \text{and} \qquad F^* = \inf_\xi \sup_\pi F(\xi,\pi).$$

It is easy to see that $0 \leq F_* \leq F^*$. Indeed, the first inequality follows because given an honest strategy θ, clearly $F(\xi,\theta) = 0$ for every ξ, and the second inequality always holds. Intuitively, we may think of each of the quantities above as resulting from different order of play. Thus, F_* represents the agent's value if he moves first, whereas F^* reflects the agent's value when the principal moves first. The next three results should help to understand how these quantities are related to Lemma 1.

Proposition 2. *$F_* = 0$ if and only if every profitable deviation is detectable.*

Proof. If $F_* = 0$ then, given any strategy π, it follows that $\inf_\xi F(\xi,\pi) \leq 0$. If π is undetectable then $\sum_{(t,s)} \pi(t,s) \int (\xi(s,y) - \xi(t,y)) p(dy|t) = 0$ regardless of ξ, hence $\sum_{(t,s)} \pi(t,s) \Delta v(t,s) \leq 0$, i.e., every undetectable deviation is unprofitable. Conversely, if $F_* > 0$ there is a deviation π with $\inf_\xi F(\xi,\pi) > 0$. If π is detectable then there is a scheme ξ with $\sum_{(t,s)} \pi(t,s) \int (\xi(s,y) - \xi(t,y)) p(dy|t) = \alpha > 0$. Letting $\beta = \sum_{(t,s)} \pi(t,s) \Delta v(t,s)$, it easily follows that $F(\beta\xi/\alpha,\pi) \leq 0$, a contradiction. Therefore, π must be undetectable. □

Proposition 2 immediately delivers the finite-dimensional version of Theorem 1, i.e., if T is finite then an allocation is implementable if and only if every profitable deviation is detectable. This is because in the finite case the Minimax Theorem implies that (i) $F_* = F^*$ and (ii) both sup and inf are attained. The challenge in the infinite case is to extend this argument. Lemma 1 shows that the existence of a saddle point for F with saddle value equal to zero characterizes implementability. In general, the attainment of a saddle value is not enough to guarantee existence of a saddle point. In other words, $F_* = F^* = 0$ is not enough for implementability. What is enough for implementability is both $F_* = F^* = 0$ and attainment of the sup and inf, just as in the finite case. I now show that the condition of Theorem 1 completely characterizes these two requirements. To this end, I first introduce an important function in terms of which implementability may be characterized.

Let $ba(Y)_0^{(T)}$ be the set of functions z from T to $ba(Y)$—the space of bounded, additive measures on Y—with finite support such that $\sum_t z(Y|t) = 0$. For any $z \in ba(Y)_0^{(T)}$, let $h(z) = \sup_\pi \inf_\xi \{F(\xi,\pi) - \xi \cdot z\}$. The definition of F gives h the following intuitive equivalent representation. For any strategy π, recall that $\Delta p(\pi)$ is defined pointwise by $\Delta p(\pi)(t) = \sum_s \pi(t,s)p(t) - \pi(s,t)p(s)$, therefore $\Delta p(\pi) \in ba(Y)_0^{(T)}$.

Lemma 1. $h(z) = \sup_\pi \{\Delta v(\pi) : \Delta p(\pi) = z\}$.

Proof. By definition of F, clearly $F(\xi, \pi) - \xi \cdot z = \Delta v(\pi) - \xi \cdot [z - \Delta p(\pi)]$. Given z and π, if $\Delta p(\pi) = z$ then $\inf_\xi \{F(\xi, \pi) - \xi \cdot z\} = \Delta v(\pi)$. Otherwise, $\Delta p(\pi) \neq z$ then there exists ξ such that $\xi \cdot [z - \Delta p(\pi)] > 0$. Scaling this ξ by an arbitrarily large amount, it follows that $\inf_\xi \{F(\xi, \pi) - \xi \cdot z\} = -\infty$. □

Intuitively, Lemma 1 says that $h(z)$ may be equivalently described as the supremum of utility gains from deviations whose statistical consequences amount precisely to z. In other words, the function h collapses the principal's strategy to either punishing the agent without bound for playing a strategy that differs from z or remaining unable to prevent the agent from choosing any strategy that is statistically identical to z. This observation quickly leads to the following result.

Lemma 2. *Every profitable deviation is detectable and h is subdifferentiable at 0 if and only if every infinitesimally detectable deviation is at most infinitesimally profitable.*

Proof. Follows immediately from Lemma A.4 and observing that every infinitesimally detectable deviation being at most infinitesimally profitable implies every profitable deviation is detectable, which in turn is equivalent to $h(0) = 0$. □

Lemma 2 leads us to the proof of our main result, Theorem 1, via the Minimax Theorem as follows. By Lemma A.3, F has an equilibrium if and only if both h and k are subdifferentiable at zero, where $k(w) = \sup_\pi \inf_\xi \{F(\xi, \pi) - w \cdot \pi\}$ is defined for all $w \in \mathscr{U} = \{u \in \mathbb{R}^{T \times T} : u(t, t) = 0\}$, and these subdifferentials characterize the set of equilibria. Specifically, the subdifferential of h at 0 corresponds to the principal's equilibrium strategies, i.e., all implementing schemes (see Proposition 5 and its proof for details). Existence of an honest equilibrium for F is now characterized as follows, finally establishing Theorem 1.

Proposition 3. *F has an honest equilibrium if and only if every profitable deviation is detectable and h is subdifferentiable at 0.*

Proof. By the observations of the previous paragraph and the fact that every profitable deviation being detectable is equivalent to $h(0) = 0$, it remains to argue that subdifferentiability of k at zero is implied by the structure of F. Indeed, we require that $k(w) \geq 0$ for all w, where, since we are looking for honest equilibria, an honest strategy θ should be a subgradient of k at 0, and $w \cdot \theta = 0$ for all w. But this follows from the definition of k, because $F(\xi, \theta) - w \cdot \theta = 0$, i.e., the agent can always secure at least zero by choosing an honest strategy. □

4 Discussion

In this section I make several observations. I begin by relating Theorem 1 to Rochet's Theorem. Secondly, I compare detectability with what I'll call convex independence, introduced by Cremer and McLean (1988). Thirdly, I distinguish implementability from existence of equilibrium in the spirit of Hart and Schmeidler (1989). Finally, I revisit revenue equivalence in the richer context of this paper.

4.1 Rochet's Theorem

On the one hand, the environment of Rochet's Theorem is a special case of the one studied here because Rochet restricts attention to environments where Y is a singleton. On the other hand, both environments have in common the use of convex duality to characterize implementability. Rochet makes use of Rockafellar's (1966) characterization of subdifferentials of convex functions in terms of cyclically monotone mappings. An important insight in their argument is making essential use of the network structure behind incentive constraints when Y is a singleton. Indeed, this insight leads to an elementary proof of their characterization.

However, their approach cannot be applied here because this network structure is absent. When Y is a singleton, implementability amounts to not just every infinitesimally detectable deviation being at most infinitesimally profitable, but moreover that every profitable deviation be detectable. This interpretation of Rochet's cyclic monotonicity condition in terms of detecting profitable deviations is easy to see. To see it, recall Rochet's Theorem: an 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 $t_1 = t_{m+1}$,

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

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. To relate this cyclic monotonicity condition with Theorem 1, I now show that a cycle can be interpreted as an undetectable reporting strategy with rational probabilities. In so doing, I 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 strategy π_σ defined by $\pi_\sigma(t,s) = 1/|S|$ if $s = \sigma(t)$ and 0 otherwise. By virtue of σ being a permutation, for every $t \in S$ (i) there exists a unique $s \in S$ such that $\pi_\sigma(t,s) > 0$, and (ii) there exists a unique $s \in S$ such that $\pi_\sigma(s,t) > 0$. Therefore, π_σ is undetectable.

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

Proof. By Rochet's Theorem and Proposition 4 below, (i) is equivalent to (ii), and (i) is equivalent to (iii) by linearity of $\Delta v(\pi)$ with respect to $\pi \in \Delta(T \times T)$ 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. $\square$

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 unprofitable. Finally, (i) implies (ii) by representing a cycle as an undetectable reporting strategy with rational probabilities as follows. To see this, let $(t_1, \ldots, t_{m+1})$ be a cycle, so $t_1 = t_{m+1}$. Let $S = \{s_1, \ldots, s_\ell\}$, with $\ell \leq m$, 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 $\pi(t,s) = [t,s]/(\ell[s_0])$ if $s \neq t$ and $1/\ell - \sum_{s \neq t}[t,s]/(\ell[s_0])$ 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 also clear that π consists of rational probabilities, since every element of π is the difference between two rational 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 "conditional" deviations, i.e., the conditional probabilities $\mu(s|t)$ such that $\sum_s \mu(s|t) = 1$ for all t. Such conditional deviations μ can be obtained easily from deviations π by writing $\mu(s|t) = \pi(t,s)/\sum_r \pi(t,r)$.

The next result extends the techniques developed in this paper to derive a version of Theorem 1 when the principal's signal y is independent of the agent's type t. This, of course, includes the case of Y being a singleton. A proof appears in the appendix. I also include in the appendix a new, direct proof for when Y is a singleton that relies on linear duality. Hence, it constitutes an alternative proof of Rochet's Theorem.

Proposition 4. *Suppose $p(t)$ does not depend on t, Y is a singleton, or T is finite. An allocation is implementable if and only if every profitable deviation is detectable.*

Given Corollary 1, Proposition 4 directly verifies Rochet's Theorem without resorting to any network structure. This suggests the following important mathematical observation: when Y is a singleton, the set Π of unprofitable, undetectable deviations is closed, but it is not necessarily closed otherwise, as Examples 1–4 showed. In the language of duality, the closedness of Π is crucial for the absence of a duality gap. To see the role played by the network structure, notice that integrality constraints on network flows do not bind when flows are consistent with permutations, as in assignment problems, which is implicitly the case studied by Rochet. A duality gap may be viewed as a binding infinite-dimensional "extremal" constraint (e.g., the no-gap example in Kretschmer, 1961), so if the integrality constraint does not bind then neither does the "extremal" constraint. On the other hand, when the integrality constraint may bind, say because Y is not a singleton, then the "extremal" constraint must be incorporated explicitly. This describes the mathematical content of Theorem 1.

4.2 Cremer and McLean's Theorem

Although it might be thought that detectability of deviations is close to the notion of convex independence by Cremer and McLean (1988), there are important distinctions. This is a potentially important concern on the following grounds. Cremer and McLean's result is often paraphrased as "if types are correlated then one can extract the surplus." Hence, as soon as the conditions of Proposition 4 fail to be met, all the surplus can be (perhaps virtually) extracted, hence the mechanism design problem becomes uninteresting. Example 5 below shows that this is false: not only are types correlated, but the wedge between convex independence and detectability is basic.

Example 5. Let $T = \{0, \frac{1}{2}, 1\}$, $Y = \{a, b\}$, and $p(t) = t\delta_a + (1-t)\delta_b$, where δ stands for Dirac measure. Convex independence clearly fails, since $p(\frac{1}{2}) = \frac{1}{2}p(0) + \frac{1}{2}p(1)$, and hence $p(\frac{1}{2})$ lies in the convex hull of $\{p(0), p(1)\}$, yet every deviation is detectable.

I will discuss these two conditions next. Firstly, convex independence characterizes surplus extraction, whereas detectability characterizes only implementability. Hence, the latter is necessarily a much weaker condition. However, to see just how these two conditions differ in interpretation, let me rewrite equivalent versions of each below. As will become apparent, the essential difference between the two conditions applies equally to the case of finitely many and infinitely many types. I will therefore assume for expositional simplicity that the set T of types is finite in this subsection.[13]

Convex independence amounts to the following: $p(t) \notin \text{conv}\{p(s) : s \neq t\}$, where, of course, "conv" means convex hull. It is easy to see that this is equivalent to

$$\sum_{s \in T} \mu(s|t)p(s) \;=\; p(t) \quad \forall t \quad \Rightarrow \quad \mu(s|t) = 0 \text{ if } s \neq t,$$

where $\mu(s|t) \geq 0$ for all (t, s) and $\sum_s \mu(s|t) = 1$ for all t. Convex independence might be interpreted as follows. If type t could become type s with probability $\mu(s|t)$ then convex independence would mean that there is no way of becoming other types that is indistinguishable from remaining the original type.

On the other hand, that every deviation is detectable may be written—using the "conditional" description with μ from the previous subsection—as

$$\sum_{s \in T} \mu(t|s)p(s) \;=\; p(t) \quad \forall t \quad \Rightarrow \quad \mu(s|t) = 0 \text{ if } s \neq t.$$

Intuitively, the left-hand side above means that the probability distribution induced by truth-telling coincides with that arising from the "conditional" strategy μ, where the prior on types is given by the uniform distribution.[14]

Notationally, the difference between the two conditions is subtle. For convex independence, the antecedent says that, for every type t, the signal distribution averaged across types with weights $\mu(\cdot|t)$ equals $p(t)$, whereas for detectability it says that the signal distribution averaged across reports with weights $\mu(t|\cdot)$ equals $p(t)$.

However, there are important differences of interpretation. Whereas convex independence may be interpreted as there being no statistically indistinguishable way of *becoming* other types apart from always remaining the original type, detectability may be interpreted as the requirement that there is no way of *pretending* to be any other type that is statistically indistinguishable from always reporting honestly.

[13] For a detailed discussion of convex independence in the infinite case, see Rahman (2010).

[14] The uniform distribution does not matter—if it were any other, the μ's could be adjusted to satisfy the same system of equations as with the uniform. Details are available on request.

4.3 Hart and Schmeidler's Theorem

The essential use of duality in this paper—embodied by the Minimax Theorem—is reminiscent of the work by Hart and Schmeidler (1989). Moreover, the condition that describes whether a deviation is undetectable is remarkably close to Equation 2 of Nau and McCardle (1990, p. 433) and Myerson (1997, p. 190). Below, I discuss some similarities and differences between their work and the present paper.

The authors mentioned above showed that correlated equilibrium exists in a finite game using duality.[15] To paraphrase, consider a zero-sum game between a mediator and a surrogate for the players. The mediator chooses a correlated strategy, whereas the surrogate randomly selects a player, followed by a deviation by that player, i.e., a (possibly mixed) recommendation-contingent plan. The mediator pays the surrogate the sum of expected unilateral deviation gains to the players in the original game. By the Minimax Theorem it doesn't matter who goes first. When the mediator goes first, the value of the game being zero means that a correlated equilibrium exists. Indeed, in this case there is a correlated strategy that discourages all unilateral deviation profiles simultaneously. When he moves second, he may discourage them one by one. A player's recommendation-contingent plan is a Markov chain with an invariant distribution. The product of these is a strategy for the mediator that secures his payoff to at least 0. Since the surrogate's security payoff is also 0 by having players always obey, the value is 0 and correlated equilibrium exists.

The previous argument had the mediator choose an allocation (or correlated strategy) by keeping fixed utilities of the underlying game. In this paper, the principal chooses (rather, influences) utilities via a payment scheme but keeps fixed the allocation. The key step that every Markov chain has an invariant distribution does not apply to this paper but its equation does. Indeed, given a deviation π_i, Hart and Schmeidler and Nau and McCardle find a probability vector p_i that makes the deviation undetectable relative to p_i. Since such a p_i always exists, the product measure $p = \prod_i p_i$ is optimal for the mediator when he moves after the surrogate. On the other hand, in this paper the vector p_i is exogenously given as i's beliefs, and only the deviations π_i that satisfy undetectability are considered feasible. Therefore, even though apparently the same equation contributes to determining both implementability and existence of correlated equilibrium, its interpretation and the way it is used differ in important ways with each application.

[15] I extend their approach to infinite games with bounded measurable utilities in Rahman (2008).

4.4 Revenue Equivalence Revisited

In this subsection, I recast the problem of revenue equivalence in the current context, and characterize implementing incentive schemes. This extends Rochet (1987)'s explicit derivation of an incentive scheme that implements a given allocation under cyclic monotonicity, but does so differently, since, again, his approach relies on a network structure that is absent here.

The results below differ from other characterizations of revenue equivalence and implementing schemes in the literature, such as Heydenreich et al. (2009) and Kos and Messner (2009). Indeed, these papers not only focus on the case of independent types (in this case, t and y are independent), but moreover just characterize when expected payments only differ by a constant, rather than when the entire payment schedule is unique up to a constant. Of course, when types are independent, the most we can hope for in terms of revenue equivalence is that expected payments differ by a constant. On the other hand, when types are not independent or there are other additional constraints imposed on payments, it becomes meaningful to consider revenue equivalence in terms of the entire schedule.

Formally, say that there is *revenue equivalence* if for any two incentive schemes ξ and ζ that implement a given allocation, there exists a constant $c \in \mathbb{R}$ such that

$$\zeta(t, y) = \xi(t, y) + c.$$

This definition differs somewhat from others. When types are independent, the above cited authors describe revenue equivalence as $\zeta(t, y) = \xi(t, y) + c(y)$, where y stands for others' types. Of course, my version is much more restrictive. Correlation amongst types diminishes the set of functions $c(y)$ with respect to which implementing schemes may differ. I take this idea to the extreme and require that schemes differ only by a constant. This may be useful, for instance, when agents exhibit risk aversion over payments. In this case, the stronger version of revenue equivalence is required for implementation to yield a unique payoff profile modulo a constant.

Mathematically, Theorem 1 shows that an allocation is implementable if and only if the function $h(z) = \sup_\pi\{\Delta v(\pi) : \Delta p(\pi) = z\}$ is subdifferentiable at 0 (Section 3). Let $\partial h(0)$ denote the subdifferential of h at 0 and $\operatorname{dom} h = \{z : h(z) > -\infty\}$ be the effective domain of h. Next, I argue that revenue equivalence is characterized in terms of differentiability of h. This brings together several results in the literature.

Proposition 5. *For any implementable allocation, $\partial h(0)$ completely characterizes all implementing payment schemes modulo a constant. Hence, revenue equivalence holds if and only if $h : dom\ h \to \mathbb{R}$ is differentiable at 0.*

Proof. Let $C = \{\xi \in B(Y)^T : \exists c \in \mathbb{R} \text{ s.t. } \xi(t, y) = c\ \forall (t, y)\}$ be the vector space of constant schemes, and write $B(Y)^T/C$ for the quotient space of (equivalence classes of) incentive schemes modulo constant payments. This is the dual space of $ba(Y)_0^{(T)}$, defined in the paragraph just before Lemma 1. With dom h being a subset of $ba(Y)_0^{(T)}$, we obtain $B(Y)^T/C$ as the space of linear functionals acting on the domain of h, and from which derivatives are defined. Now, the function h is differentiable at 0 if and only if its subdifferential at 0 is a singleton. By Lemma A.3, the subdifferential characterizes the set of solutions to the dual of the linear program that describes h in Lemma 1, i.e., the set of payment schemes, modulo a constant, that implement a given allocation, proving the first claim. The second claim now follows. $\square$

Proposition 5 may be interpreted as follows. Revenue equivalence is the statement that the function h is differentiable at 0, i.e., that its directional derivatives in any direction collapse into a single linear functional. Thus, the maximal gain associated with shifting probability mass from one type to another must exactly coincide with the minimal loss that shifts the probability mass back. Revenue equivalence is another way of saying that such equality holds for every change in probability mass.

This observation sheds light on previous results. For the special case of ex post implementation, Heydenreich et al. (2009) and Kos and Messner (2009) provide a similar characterization, but again they rely on the network structure inherent in their environment. Their results may be thought of as coming from a single agent model with Y a singleton, deriving revenue equivalence as in the previous paragraph. Indeed, they define revenue equivalence as $\zeta(t, y) = \xi(t, y) + c(y)$, i.e., conditional on y, and impose ex post implementation.[16]

The characterization of revenue equivalence for arbitrary Y, with possible correlation amongst types, is new, but given the duality approach followed above, shows clearly the underlying structure behind revenue equivalence results in simpler settings that relied on a network structure that is no longer present here.

[16] Heydenreich et al. (2009) briefly discuss Bayesian implementation with independent types, using the observation from Müller et al. (2007) that with respect to interim expected payments, Bayes-Nash implementation boils down to the same mathematical structure as ex post implementation with a single agent. This can also be seen from Proposition 4, which characterizes implementability in exactly the same way when Y is a singleton and when y is independent of t.

5 Extensions

An advantage of using duality for Theorem 1 is that it extends, as I argue next. In this section I discuss the following extensions of the model: moral hazard, revealed stochastic preference, budget balanced implementation, bargaining with interdependent values and finally a "subdifferential" characterization of implementing incentive schemes together with revenue equivalence.

5.1 Moral Hazard

The moral hazard problem fits easily into the framework developed above. To see this, consider a prototypical such problem. An agent's possible actions are described by an arbitrary set A. The principal wants the agent to choose some fixed action $a \in A$, but the agent may choose any action $b \in A$. Let Y be another measurable space of verifiable output, and $p(a) \in \Delta(Y)$ the conditional probability of such output. Finally, let $v(a) \in \mathbb{R}$ be the agent's utility from each action a.

An action a is *enforceable* if there is a payment scheme $\xi \in B(Y)$ such that

$$v(b) - v(a) \ \leq \ \int_Y \xi(y)[p(dy|b) - p(dy|a)] \qquad \forall b \in A.$$

A *deviation* in this setting is any $\pi \in \mathbb{R}^{(A)}$ such that $\pi \geq 0$ and $\sum_a \pi(a) = 1$. Such a π is called *undetectable* if

$$p(a) \ = \ \sum_{b \in A} \pi(b) p(b).$$

A deviation π is called *a-profitable* if $\sum_b \pi(b)[v(b) - v(a)] > 0$. Finally, say that *every infinitesimally detectable deviation is at most infinitesimally a-profitable* if

$$\sup_\pi \frac{\Delta v(\pi)}{|\Delta p(\pi)|} < +\infty,$$

where $\Delta p(\pi) = \|p(a) - \sum_b \pi(b)p(b)\|$ and $\Delta v \in \mathbb{R}^A$ satisfies $\Delta v(b) = v(b) - v(a)$ for all $b \in A$. The next result follows easily from previous ones, so its proof is omitted.

Theorem 2. *An action a is enforceable if and only if every infinitesimally detectable deviation is at most infinitesimally a-profitable.*

Theorem 2 shows how implementability of an allocation is characterized in the same manner under adverse selection as moral hazard. In each context, implementability boils down to detecting profitable deviations from either honesty or obedience.

5.2 Revealed Stochastic Preference

It is well-known that cyclic monotonicity characterizes rationalizable economic behavior in the spirit of Afriat (1967) and others, so that Rochet's Theorem is comparable to Afriat's Theorem of revealed preference. This comparison is formalized by Rochet himself (Rochet, 1987, pp. 195–196) in a quasi-linear context.[17]

Briefly, recall that an allocation $\mathbf{x} : T \to X$ is implementable if

$$v(t, \mathbf{x}(t)) - \xi(t) \ \geq \ v(t, \mathbf{x}(s)) - \xi(s) \qquad \forall (t, s).$$

By the taxation principle, any two reports that lead to the same choice must cost the same amount of money, i.e., $\mathbf{x}(t) = \mathbf{x}(s)$ implies that $\xi(t) = \xi(s)$ whenever $\mathbf{x}$ is implementable. Hence we may rewrite the previous inequalities as

$$v(t, \mathbf{x}(t)) - \xi(\mathbf{x}(t)) \ \geq \ v(t, \mathbf{x}(s)) - \xi(\mathbf{x}(s)) \qquad \forall (t, s).$$

Reinterpret $v'(t) = -v(t) \in \mathbb{R}^X$ as a vector of "nonlinear prices," $\mathbf{x}(t)$ as a "choice" and t as a parameter indexing price/choice outcomes. Finally, interpret $\xi'(x) = -\xi(x)$ as a utility function over the range of $\mathbf{x}$. By definition, there exists a quasi-linear utility function ξ' that rationalizes every choice $\mathbf{x}(t)$ given prices $v'(t)$ if

$$\xi'(\mathbf{x}(t)) - v'(t, \mathbf{x}(t)) \ \geq \ \xi'(\mathbf{x}(s)) - v'(t, \mathbf{x}(s)) \qquad \forall (t, s).$$

Now it is clear how Rochet's Theorem and Afriat's Theorem follow from each other.

Similarly, Theorem 1 is comparable to the work of McFadden (2005) on revealed stochastic preference. Let us follow the previous logic in the stochastic setting. We will think of "output" Y as a summary of uncertainty subsequent to the determination of an agent's type t. Thus, we now think of a random allocation $\mathbf{x} : T \times Y \to X$, where the randomness comes from Y. We appeal once more to the interpretation of $v(t, y) \in \mathbb{R}^X$ as a vector of "nonlinear prices," although we now allow it to be random. This is captured by its dependence on y. Similarly, we may think of $\mathbf{x}(t, y)$ as "random choices." Hence, y determines both prices v and choices $\mathbf{x}$. In particular, these two variables could be correlated given t. Intuitively, we are given a collection of price-choice distributions, and ask whether or not such a distribution may be generated by a *population* of quasi-linear utility maximizers, indexed by y, whose members make choices given personalized nonlinear prices.

[17] However, results in the quasi-linear context can be used to derive general rationalizability results, as Afriat (1967, p. 72) does in a neoclassical setting. See also Afriat (1963), Richter and Wong (2005).

Following the argument for revealed preference, we seek to interpret $\xi'(t, y) = -\xi(t, y)$ as a (random, expected) utility function over choices by appealing to a suitable version of the taxation principle. Unfortunately, this principle is not available, since implementability does not require that $\xi(t, y) = \xi(s, y)$ whenever $\mathbf{x}(t, y) = \mathbf{x}(s, y)$. Therefore, Theorem 1 does not directly capture revealed stochastic preference. On the other hand, Theorem 1 may be extended by *imposing* such restrictions on ξ. The outcome of this exercise is documented in the next result.

Say that $\mathbf{x}$ is *implementable as a menu* if it is implementable with a scheme ξ such that $\xi(t, y) = \xi(s, y)$ whenever $\mathbf{x}(t, y) = \mathbf{x}(s, y)$. To characterize such version of implementability, we require further notation. Let $\mathscr{R} = \{(t, s, y) : \mathbf{x}(t, y) = \mathbf{x}(s, y)\}$ be the set that indexes restrictions on ξ, and write $\mathbf{1}_{\mathscr{R}}$ for the indicator function of $\mathscr{R}$, so $\mathbf{1}_{\mathscr{R}}(t, s, y) = 1$ if $(t, s, y) \in \mathscr{R}$ and 0 otherwise.

Theorem 3. *An allocation $\mathbf{x}$ is implementable as a menu if and only if for any net $\{(\lambda_\delta, \mu_\delta)\}$ such that $\lambda_\delta \in \mathbb{R}_+^{(T\times T)}$ and $\mu_\delta \in ba(Y)^{(T\times T)}$, if*

$$\lim \sum_{s\in T}(\lambda_\delta(t, s)p(s) - \lambda_\delta(s, t)p(t)) - \sum_{s\in T}\int_Y [y]\mathbf{1}_{\mathscr{R}}(t, s, y)(\mu_\delta(dy|t, s) - \mu_\delta(dy|s, t)) \;=\; 0$$

for every t then $\lim \Delta v(\lambda_\delta) \leq 0$, where $[y]$ stands for Dirac measure and the integral above is vector-valued in $ba(Y)$.

This result follows similarly to previous results, so a proof is omitted. Theorem 3 generalizes previous results in several directions. To help describe them, think of y as a parameter for different types of decision maker in a heterogeneous population.[18]

First, Theorem 3 characterizes revealed stochastic preference of a population of decision makers under the following weaker assumptions: (i) it allows for "personalized budgets" because v may depend on y and therefore is compatible with correlation between prices, choices and utility, (ii) it allows for different populations in the sample of observed choices because p may depend on t, and (iii) it does not impose any structure on T, so is compatible with any possibly infinite set of types. This contrasts with McFadden (2005), who in characterizing revealed stochastic preference in the infinite case confines attention to compact metric type spaces. On the other hand, Theorem 3 is restricted by the assumption of quasi-linearity, although this assumption may be dropped by applying Afriat's (1967, p. 72) trick.

[18]The interpretation of revealed stochastic preference as a population distribution of behavior rather than uncertain behavior by a single decision maker may be attributed to McFadden and Richter (1990, Footnote 25, pp. 174–5).

Adding structure to the problem reveals further insights in Theorem 3. For instance, suppose that $p(t) = p$ for all t, so the population does not change with observed behavior, and $\mathbf{x}(t) \neq \mathbf{x}(s)$ with positive p-probability for every pair (t, s). Intuitively, there are no duplicate observations. Now it is easy to see that Theorem 3 boils down to a version of Proposition 4, i.e., every profitable deviation is detectable.

Corollary 2. *Suppose that $p(t) = p$ does not depend on t and that $\mathbf{x}(t) \neq \mathbf{x}(s)$ with positive p-probability for every pair (t, s). An allocation $\mathbf{x}$ is implementable as a menu if and only if every profitable deviation is detectable.*

A slightly different version of Theorem 3 obtains by imposing $\xi(t, y) = \xi(s, y)$ for all y whenever $\mathbf{x}(t) = \mathbf{x}(s)$, instead of $\xi(t, y) = \xi(s, y)$ whenever $\mathbf{x}(t, y) = \mathbf{x}(s, y)$. This means that we may rewrite ξ as $\xi(\mathbf{x}(t), y)$ for every y. In other words, observed choices may be represented as coming from a population choosing an efficient Y-contingent allocation when individual utility functions are quasi-linear and exhibit consumption externalities, since individuals of type y now care about the entire allocation $\mathbf{x}(t)$. This is an easy exercise given the techniques developed above, hence omitted.

5.3 Budget Balanced Implementation

I will now characterize budget balanced implementation. Consider the following multi-agent setting, where $I = \{1, \ldots, n\}$ is a finite set of agents, and for each $i \in I$, T_i is a measurable space of types, $T = \prod_i T_i$ is the product space of type profiles with the product σ-algebra, and $v_i(t, \mathbf{x}(s_i, t_{-i})) \in \mathbb{R}$ is the utility to agent i under allocation $\mathbf{x} : T \to X$ when his type is t_i but he reported s_i. For each agent i and type t_i, let $p_i(t_i)$ be a finitely additive probability measure over T_{-i}—the space of other agents' type profiles—that describes i's posterior beliefs over others' types.

An incentive scheme is now a vector $\xi = (\xi_1, \ldots, \xi_n)$ of payment schedules, one per agent. It is called *budget balanced* if $\sum_i \xi_i(t) = 0$ for every t. Say $\mathbf{x}$ is *implementable with budget balance* if there is a budget balanced scheme that implements $\mathbf{x}$ for all i. For this to be well defined I maintain Assumptions 1 and 2, so $\xi \in \prod_i B(T_{-i})^{T_i}$. Notice that I do *not* require v_i to be uniformly bounded. A *strategy profile* is any profile $\pi = (\pi_1, \ldots, \pi_n)$ of strategies for each agent i. Call any such π *profitable* if $\Delta v(\pi) = \sum_{(i,t_i,s_i)} \Delta v_i(t_i, s_i)\pi_i(t_i, s_i) > 0$, where

$$\Delta v_i(t_i, s_i) \;=\; \int_{T_{-i}} [v_i(t, \mathbf{x}(s_i, t_{-i})) - v_i(t, \mathbf{x}(t))] p_i(dt_{-i}|t_i).$$

Given any vector $\eta \in \mathbb{R}^{(T)}$, let $\Delta p(\pi|\eta)$ be defined pointwise for each (i, t_i) as follows:

$$\Delta p_i(\pi|\eta)(t_i) \;=\; \sum_{s_i \in T_i} [\pi_i(t_i, s_i)p_i(t_i) - \pi_i(s_i, t_i)p_i(s_i)] - \sum_{t_{-i}} \eta(t_i, t_{-i})\delta_{t_{-i}},$$

where $\delta_{t_{-i}}$ is Dirac measure. Think of η as a vector of multipliers on the ex post budget constraints. This vector η leads to a different function from h in Lemma 1 whose subdifferentiability will characterize budget balanced implementation.

To this end, I will use a condition similar to that for Theorem 1, based on the notion of attribution introduced in the finite case by Rahman and Obara (2010). A strategy profile π is called *unattributable* if η exists such that $\Delta p_i(\pi|\eta) = 0$ for every agent i. Intuitively, for any unilateral deviation in π, the change in probabilities over reports is the same across agents, making it impossible to statistically identify an innocent agent (let alone the deviator). Following the proof of Theorem 1, it is straightforward that subdifferentiability of $k(z) = \sup_{(\pi,\eta)}\{\Delta v(\pi) : \Delta p(\pi|\eta) = z\}$ at 0 characterizes budget balanced implementation. Just as before, an equivalent description of this subdifferentiability is the following. Say *every infinitesimally attributable strategy profile is at most infinitesimally profitable* if $\sup_{\pi,\eta} \Delta v(\pi)/\left|\Delta p(\pi|\eta)\right| < +\infty$, where

$$|\Delta p(\pi|\eta)| \;=\; \sum_{i=1}^{n}\sum_{t_i \in T_i} \|\Delta p_i(\pi|\eta)(t_i)\|$$

and the norm above is the total variation norm on $B(T_{-i})$ for each (i, t_i).

Theorem 4. *An allocation is implementable with budget balance if and only if every infinitesimally attributable strategy profile is at most infinitesimally profitable.*

A proof of Theorem 4 is very close to Theorem 1, so it is omitted. To illustrate the usefulness of Theorem 4, let us now characterize exactly when budget balance is a binding constraint. Say that *budget balance is not binding* if for any "budget" $b \in \mathbb{R}^T$ there is an incentive scheme ξ such that $\sum_i \xi_i(t) = b(t)$ for all t and

$$\int_{T_{-i}} (\xi_i(s_i, t_{-i}) - \xi_i(t))p_i(dt_{-i}|t_i) \;\geq\; 0 \qquad \forall (i, t_i, s_i).$$

To understand this definition, consider a scheme ζ that implements some allocation. If budget balance is not binding then there is an additional scheme ξ that absorbs any budgetary surpluses and deficits from ζ without disrupting incentive constraints.

Say that *detection implies attribution* whenever for any strategy profile π, if π is unattributable then every π_i is undetectable. Say that *detection implies attribution asymptotically* if $\sup_{\pi,\eta} b \cdot \eta/\left|\Delta p(\pi|\eta)\right| < +\infty$ for every budget $b \in \mathbb{R}^T$.

Proposition 6. *Budget balance is not binding if and only if detection implies attribution asymptotically. With independent types, this holds if and only if detection implies attribution.*

The proof of Proposition 6 is similar to previous ones, hence omitted.

Corollary 3. *With independent types, detection implies attribution, so an allocation is implementable with budget balance if and only if it is implementable for every agent.*

Proof. By Proposition 6, it suffices to show that detection implies attribution with independent types. If not, there exists an unattributable deviation profile π, i.e., such that $\sum_{s_i}(\pi_i(s_i|t_i) - \pi_i(t_i|s_i)) = \eta(t)$ for each i, and an agent i such that π_i is detectable. Hence, $\sum_t \eta(t) = 0$. By detectability, $\eta(t) > 0$ and $\eta(s) < 0$ for some pair (t, s). But then $\eta(s) = \eta(s_i, t_{-i}) < 0$ and $\eta(t) = \eta(t_j, s_i, t_{-ij}) > 0$, a contradiction. □

To illustrate, consider the special case of *private values*, where each agent's utility is independent of others' types, i.e., $v_i(t, x) = v_i(t_i, x)$ for all x, and an *ex post efficient* allocation, i.e., $\mathbf{x}^*$ such that $\mathbf{x}^*(t) \in \arg\max_x \sum_i v_i(t_i, x)$ for all $t \in T$. An allocation is *ex post implementable* if it is implementable when $p_i(t_i) = [t_{-i}]$ for all (i, t).

Corollary 4. *With private values, $\mathbf{x}^*$ is ex post implementable. Therefore, with independent private values $\mathbf{x}^*$ is implementable with or without budget balance.*

Proof. By Proposition 4, we must show that every profitable deviation is detectable. Otherwise, suppose that π_i is a profitable, undetectable deviation and consider the welfare consequences of agent i reporting according to π_i instead of truthfully. Since π_i is undetectable and values are private, the expected utility to any agent $j \neq i$ is the same if i plays π_i or if he reports truthfully. On the other hand, agent i is strictly better off, therefore, welfare increases when agent i plays π_i instead of reporting truthfully. But this contradicts ex post efficiency. The rest follows by Corollary 3. □

I end this section by extending Rochet's Theorem to include budget balance. In a general environment with possibly interdependent values, I now characterize allocations that are ex post implementable with a scheme that also satisfies budget balance. Just as with Rochet's Theorem relative to Theorem 1, the infinitesimal qualifiers of Theorem 4 are not necessary for budget balanced ex post implementability. A proof of this result appears in the appendix.

Theorem 5. *An allocation is ex post implementable with budget balance if and only if profitable deviation profile is attributable.*

5.4 Bargaining with Interdependent Values

A *bargaining problem* is the task of finding an incentive scheme that implements a given allocation without violating budget balance or individual rationality, described below. In this subsection I characterize existence of solutions to such problems.

A mechanism $(\mathbf{x}, \xi)$ is called *individually rational* if

$$\int_{T_{-i}} [v_i(t, \mathbf{x}(t)) - \xi_i(t)] p_i(dt_{-i}|t_i) \ \geq \ \int_{T_{-i}} v_i(t, \mathbf{x}(0)) p_i(dt_{-i}|t_i) \qquad \forall (i, t_i),$$

where $\mathbf{x}(0)$ is the *disagreement outcome*, i.e., what happens when an agent decides to opt out of the mechanism. A *bargaining solution* for $\mathbf{x}$ is any incentive scheme ξ that implements $\mathbf{x}$ with budget balance and renders $(\mathbf{x}, \xi)$ individually rational.

The multipliers on an agent's individual rationality constraint may be interpreted as the probability with which the agent deviates to opting out. Therefore, in this setting we redefine a strategy to be any $\pi_i \in \Delta(T_i \times T_i \cup \{0\})$, where π_i has finite support and $\pi_i(t_i, 0)$ stands for the probability that agent i chooses type t_i to opt out. Since $0 \notin T_i$ it is clear that every deviation where opting out has positive probability is detectable. As before, a strategy profile π is *profitable* if $\Delta v(\pi) = \sum_{(i,t_i,s_i)} \Delta v_i(t_i, s_i) \pi_i(t_i, s_i) > 0$, where where the summation is indexed by $i \in I$, $t_i \in T_i$ and $s_i \in T_i \cup \{0\}$, $\Delta v_i(t_i, s_i)$ is defined as usual except that now its domain is $I \times T_i \times T_i \cup \{0\}$ and

$$\Delta v_i(t_i, 0) = \int_{T_{-i}} [v_i(t, \mathbf{x}(0)) - v_i(t, \mathbf{x}(t))] p_i(dt_{-i}|t_i).$$

(Obviously, everything goes through just the same even if disagreement outcomes depend on who opts out and others' types.) The definition of attribution and its infinitesimal counterpart is just the same as in the previous subsection, except for the caveat that s_i ranges across $T_i \cup \{0\}$.

Theorem 6. *Fix an arbitrary allocation.* (1) *A bargaining solution exists if and only if every infinitesimally attributable strategy profile is at most infinitesimally profitable.* (2) *When types are independent, a bargaining solution exists if and only if every profitable strategy profile is attributable.*

Once again, the proof of this result is similar to previous ones, and is therefore omitted. Theorem 6 may be contrasted with Segal and Whinston (2009) in that—using duality—it characterizes existence of bargaining solutions even when values are interdependent, the type space is arbitrary utility functions are not necessarily uniformly bounded, and especially when types are correlated.

6 Conclusion

In this paper I characterize implementability (Theorem 1) by making use of the Minimax Theorem, emphasizing a strategic interpretation. I also suggest some extensions of this result. I have tried to improve upon Rochet's Theorem both in supplying a strategic interpretation and also generalizing its result.

Mathematically, a notable difference with Rochet's or Rockafellar's approach is that they derived cyclic monotonicity in some sense by "integrating" a subdifferential correspondence and exploiting an inherent network structure. Using Rockafellar's "fundamental theorem of calculus" for convex functions, Rochet constructed an implementing payment scheme by integrating the utility gains function. On the other hand, I take the alternative system of inequalities from incentive compatibility and think of payment schemes as multipliers on the dual undetectability constraints, i.e., I view them as (directional) derivatives. Hence, I obtain the payment schemes by differentiating a dual value function, rather than integrating a subdifferential correspondence. A substantial obstacle in generalizing Rochet's basic approach is that the network structure inherent in his argument is lost when the principal observes information above and beyond the agent's report. However, the approach I follow in this paper still bears a resemblance in that it hinges on duality without relying on networks. As a result, the approach generalizes.

As a final comment, although the approach used in this paper may appear reminiscent of linear semi-infinite programming (LSIP, see, e.g., Goberna and López, 1998), please note that in this paper there may be both (a) infinitely many (incentive) constraints and (b) infinitely many unknowns. Therefore, this isn't strictly speaking LSIP.

A Preliminaries

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 above, 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 a function 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.*

For the next two results, we need some definitions. Let $f : X \times Y \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Let $\mathrm{dom} f = \{(x, y) : |f(x, y)| < \infty\}$. Write $\mathrm{dom}_1 f$ and $\mathrm{dom}_2 f$ for the projections of $\mathrm{dom} f$ on X and Y, respectively. Say that f is *closed* if both $\{x' : f(x', y) \geq c\}$ and $\{y' : f(x, y') \leq c\}$ are closed sets for every $c \in \mathbb{R}$, $x \in \mathrm{dom}_1 f$ and $y \in \mathrm{dom}_2 f$. The function f is *concave-convex* if it is concave with respect to x for all $y \in \mathrm{dom}_2 f$ and convex with respect to y for all $x \in \mathrm{dom}_1 f$.

Lemma A.3 (Ioffe and Tikhomirov, 1968, page 84). *Let $f : X \times Y \to \mathbb{R} \cup \{\pm\infty\}$ be a closed concave-convex function, and define the following functions on X^* and Y^*:*

$$h(z) = \inf_{y \in \mathrm{dom}_2 f} \sup_x \{x \cdot z - f(x, y)\} \quad \text{and} \quad k(w) = \sup_y \inf_{x \in \mathrm{dom}_1 f} \{f(x, y) + y \cdot w\}.$$

For f to have a saddle point it is necessary and sufficient that $\partial h(0) \neq \emptyset \neq \partial k(0)$. The set of saddle points coincides with the product $\partial h(0) \times \partial k(0)$.

Since subdifferentiability is such a prominent criterion in the results above, I document below a test for it that will be prove useful. The test is given by the notion of "bounded steepness," which may be attributed to Gale (1967).

Lemma A.4 (Gretsky et al, 2002, page 267). *Let $f : Y \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function on a normed linear space Y, and suppose that $f(b) < +\infty$. The function f is subdifferentiable at b if and only if f has* bounded steepness *at b, i.e., the quotients*

$$\frac{f(b) - f(y)}{\|y - b\|}$$

are bounded above.

B Proof of Proposition 4

I prove Proposition 4 with the next two results.

Lemma 3. *F has an honest equilibrium if and only if for every net $\{\lambda_\delta\} \subset \mathbb{R}_+^{(T\times T)}$,*

$$\lim \Delta p(\lambda_\delta)\cdot\xi \;=\; 0 \quad \forall \xi \in B(Y)^T \quad \Rightarrow \quad \limsup \Delta v(\lambda_\delta) \;\le\; 0, \tag{3}$$

where $\Delta p(\lambda_\delta)\cdot\xi = \int_Y \sum_{(t,s)} \xi(t,y)[\lambda_\delta(t,s)p(dy|s) - \lambda_\delta(s,t)p(dy|t)]$ for each scheme ξ and $\Delta v(\lambda_\delta) = \sum_{(t,s)} \Delta v(t,s)\lambda_\delta(t,s)$ for every $\Delta v \in \mathscr{U}$.

Proof. I will show that (3) is the dual of implementability. Let $X = B(Y)^T/C \times \mathbb{R}^{T\times T}$ with $[B(Y)^T/C]_+ = B(Y)^T/C$ and $Y = \mathbb{R}^{T\times T}$. Given $\xi \in B(Y)^T/C$, $\sigma \in \mathbb{R}^{T\times T}$ and $(t,s) \in T\times T$, define $A(\xi,\sigma)(t,s) = \int_Y \xi(s,y) - \xi(t,y)p(dy|t) - \sigma(t,s)$. Clearly, $A : X \to Y$ is a continuous linear operator. By Lemma A.1, there exists $(\xi,\sigma) \in X_+$ such that $A(\xi,\sigma) = \Delta v$ if and only if $\Delta v(\mu_0) \ge 0$ for any $\mu_0 \in \mathbb{R}^{(T\times T)}$ for which a net $\{(\lambda_\delta,\mu_\delta)\} \subset Y_+^* \times Y^*$ exists with $\Delta v(\mu_\delta) = 0$ for all δ, $\Delta p(\mu_0) = \lim -\Delta p(\mu_\delta)$ and $\mu_0 = \lim \lambda_\delta - \mu_\delta$, where $\Delta p(\mu)(t) = \sum_s \mu(t,s)p(t) - \mu(s,t)p(s)$ for all t. Since Δp is continuous and linear, $\Delta p(\mu_0) = \lim \Delta p(\lambda_\delta) - \Delta p(\mu_\delta) = -\Delta p(\mu_\delta)$, therefore $\lim \Delta p(\lambda_\delta) = 0$. Since Δv is also a continuous linear functional of μ and $\Delta v(\mu_\delta) = 0$, applying the same argument that derived $\Delta p(\lambda_\delta) \to 0$ yields the dual condition $\lim \Delta p(\lambda_\delta) = 0 \Rightarrow \limsup \Delta v(\lambda_\delta) \le 0$. Finally, since convergence in $ba(Y)^{(T)}$ is defined with respect to the weak$*$ topology, (3) follows. $\square$

We now establish the last lemma we need to prove Proposition 4.

Lemma 4. *Let $\lambda \in \mathbb{R}_+^{(T\times T)}$ and suppose that $p(t)$ does not depend on t. The equation $\Delta p(\lambda) = 0$ implies that $\Delta v(\lambda) \le 0$ if and only if (3) holds for every net $\{\lambda_\delta\}$.*

Proof. Necessity follows by restricting attention to constant nets. For sufficiency, consider a net $\{\lambda_\delta\}$ with $\lim \Delta p(\lambda_\delta)\cdot\xi = 0$ for all ξ but $\limsup \Delta v(\lambda_\delta) = 1$. Without loss, take a convergent subsequence, $\{\lambda_m\}$, by picking λ_m with $|\Delta v(\lambda_m) - 1| < 1/m$. Define the vector $\mu_m \in ba(Y)^{(T)}$ pointwise by $\mu_m(t) = \sum_s \lambda_m(t,s)p(s) - \lambda_m(s,t)p(t)$ for every t. Let us consider the following three cases.

– *Case 1a*: Suppose that supp λ_m does not depend on m.

Let $S_m = \{t : \lambda_m(t,s) > 0 \text{ for some } s\} \cup \{t : \lambda_m(s,t) > 0 \text{ for some } s\}$ be the set of types to which λ_m gives positive weight. By hypothesis, and $S_m = S$ is independent of m and $|S| < \infty$. Consider the cone $C = \{\lambda \in \mathbb{R}_+^{S \times S} : \Delta p(\lambda) = 0\}$. Since this cone is finitely generated, it is closed. Therefore, by the Theorem of the Alternative, there does not exist $\lambda \geq 0$ such that $\Delta p(\lambda) = 0$ in $ba(Y)^S$ and $\Delta v(\lambda) > 0$ if and only if there is a scheme $\xi \in B(Y)^S$ that implements $\mathbf{x}$ assuming that the type space is S. Applying Lemma 3, the result now follows.

– *Case 1b*: Suppose that supp μ_m does not depend on m but supp λ_m does.

Let supp $\mu_m = T_0$, which, by hypothesis, does not depend on m. Clearly, $T_0 \subset S_m$. If $T_0 = S_m$ then we are in Case 1a, and we are done. Otherwise, $S_m \setminus T_0 \neq \emptyset$. Let $\lambda_m^0(t,s) = \lambda_m(t,s)$ if either t or s (or both) belong to $S_m \setminus T_0$ and 0 otherwise, and let $\lambda_m^1 = \lambda_m - \lambda_m^0$. Consider the following optimization problem:

$$\begin{aligned} V_m &= \min_{\eta_m \geq 0} \left\| \lambda_m^1 - \eta_m \right\|_1 \qquad \text{s.t.} \\ \sum_{s \in T_0} \eta_m(t,s)p(s) - \eta_m(s,t)p(t) + \sum_{s \in S_m \setminus T_0} \lambda_m(t,s)p(s) - \lambda_m(s,t)p(t) &= 0 \qquad \forall t \in T_0. \end{aligned}$$

If types are independent, i.e., $p(t)$ does not depend on t, then this problem has a feasible solution and we may avoid reference to y without any loss of generality. Indeed, for any vector $b \in \mathbb{R}^{T_0}$, by the Theorem of the Alternative $\eta \geq 0$ exists such that $\sum_s \eta(t,s) - \eta(s,t) = b(t)$ for all t if and only if $\sum_t b(t) = 0$, and clearly $\sum_{t \in T_0} \sum_{s \in S_m \setminus T_0} \lambda_m(t,s) - \lambda_m(s,t) = \sum_{(t,s)} \lambda_m^1(t,s) - \lambda_m^1(s,t) = 0.$[19] Taking the dual of this problem, manipulating it and applying strong duality yields

$$V_m = \max_{\zeta \in \mathbb{R}^{T_0}} \sum_{t \in T_0} \zeta(t) \sum_{s \in S_m} \lambda_m(t,s) - \lambda_m(s,t) \;\; \text{s.t.} \;\; -1 \leq \zeta(t) - \zeta(s) \leq 1 \quad \forall (t,s).$$

But this dual problem is easily solved, yielding $V_m = \frac{1}{2} \sum_t \left| \sum_s \lambda_m(t,s) - \lambda_m(s,t) \right|$. Since T_0 is finite, it follows that $\mu_m \to 0$ in norm, hence $V_m \to 0$. If a subsequence exists such that for every k there is (t,s) such that $\Delta v(t,s) = -\infty$ and $\lambda_{m_k}(t,s) > 0$ then we are done, so suppose not, i.e., for m sufficiently large, $\Delta v(t,s) > -\infty$ for all (t,s) such that $\lambda_m(t,s) > 0$. Finally, if η_m is an optimal primal solution then $\Delta v(\lambda_m) = \Delta v(\lambda_m^0 + \eta_m + \lambda_m^1 - \eta_m) \leq \Delta v(\lambda_m^1 - \eta_m) \leq \|\Delta v\|_\infty \|\lambda_m^1 - \eta_m\|_1 \to 0$, where the first inequality follows because by construction $\lambda_m^0 + \eta_m$ is undetectable, and we are assuming that $\Delta v(\lambda) \leq 0$ for every undetectable λ. But this contradicts the original hypothesis that $\Delta v(\lambda_m) \to 1$, and the claim is established for this case.

[19] Notice that the conclusion that the problem above is feasible does not necessarily follow if types fail to be independent. For instance, in the setting of Example 1, suppose that $\lambda_m(t,s) = 1$ if t and s both belong to $\{0, \frac{1}{2}, 1\}$ and $s = t \pm \frac{1}{2}$, otherwise $\lambda_m(t,s) = 0$.

– *Case 2*: Suppose that supp μ_m depends on m.

If supp $\mu_m = T_m$ has a subsequence $\{T_{m_k}\}$ that does not depend on k then we are back to Case 1, so suppose not, i.e., there is a subsequence $\{T_{m_k}\}$ with $T_{m_k} \setminus \bigcup_{\ell<k} T_{m_\ell} \neq \emptyset$ for all k. Without loss, assume that this is the sequence with which we began. Construct $\{t_m\}$ with to innovations in T_m, i.e., $t_m \in T_m$ yet $t_m \notin T_k$ for all $k < m$.

By taking a subsequence if necessary, without loss $\{\lambda_m\}$ satisfies $\|\mu_m(t_k)\| < 2^{-m}$ for all $k \leq m$. Let $\xi(t_m, y) = [\mathbf{1}_{P_m}(y) - \mathbf{1}_{N_m}(y)]/\left\|\mu_m(t_m)\right\|$ for all m and all y, where P_m and N_m are the positive and negative sets in a Hahn decomposition of Y relative to μ_m (see, e.g., Folland, 1999, for a definition of Hahn decomposition), and $\mathbf{1}_X(y)$ is the indicator function of $X \subset Y$. (If $t \notin \{t_m\}$ then $\xi(t, y) = 0$.) For every $m \in \mathbb{N}$,

$$\begin{aligned} \xi \cdot \mu_m \;&=\; \sum_{k=1}^{m} \xi(t_k) \cdot \mu_m(t_k) \;=\; 1 + \sum_{k<m} \xi(t_k) \cdot \mu_m(t_k) \\ >\; 1 - \sum_{k<m} 2^{-m} \cdot 2^k \;&=\; 1 - \sum_{k<m} 2^{k-m} \;=\; 1 - (1 - 2^{-m}) \;\geq\; 1/2. \end{aligned}$$

Hence, it is not the case that $\Delta p(\lambda_m) \cdot \xi \to 0$ for all ξ, so (3) follows vacuously. $\square$

Let me make a few remarks about this proof. The proof of Proposition 4 is useful for two reasons. Firstly, Lemma 3 shows how the dual, alternative system of inequalities that are equivalent to implementability generalizes from the finite case to the arbitrary case. Secondly, Lemma 4 reconciles Theorem 1 with Rochet's Theorem by showing that when types are independent we revert back to the requirement of detecting profitable deviations to characterize implementability. Hence, the additional requirement of infinitesimally detectable deviations being at most infinitesimally profitable loses its bite in this setting.

It is interesting to note that the independence assumption is used to prove Lemma 4 only in Case 1b. In the other two cases, the assumption is not necessary, and in fact not used. This observation reveals the structure of the examples used to illustrate the differences between Theorems 1 and Rochet's Theorem. There, the sequence of deviations constructed fit into Case 1b, i.e., the support of μ_m was independent of m but the support of π_m crucially was not. As a final remark, note that the duality used to establish Case 1b can be used to provide a dual characterization of when detecting profitable deviations implies its infinitesimal counterpart. Namely, as long as the primal problem of Case 1b is feasible, or equivalently its dual is bounded, we obtain the result that detecting profitable deviations implies that every infinitesimally detectable deviation is at most infinitesimally profitable.

C An Alternative Proof of Rochet's Theorem

Let us prove directly that an allocation is implementable (with Y a singleton) if and only if every profitable deviation is detectable. First, assume that T is a finite set.

Lemma 5. *If T is a finite set then an allocation $\mathbf{x}$ is implementable if and only if every 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 profitable. $\square$

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

Any $\lambda \in \mathbb{R}^{(Z \times Z)}$ 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 $\mathrm{supp}_Z\, g = \{z_1, \ldots, z_n\}$ 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.

Write $\Delta v(t,s) = v(t,\mathbf{x}(s)) - v(t,\mathbf{x}(t))$ and $\Delta v(\lambda) = \sum_{(t,s)} \lambda(t,s) \Delta v(t,s)$ as usual. Define the linear operator $D : \mathbb{R}^{(T \times T)} \to \mathbb{R}^{(T)}$ as follows. Given $\lambda \in \mathbb{R}^{(T \times T)}$, let $D\lambda = \sum_{(k,\ell)} \lambda_{k\ell}(\mathbf{e}_{t_k} - \mathbf{e}_{t_\ell})$ and $D\lambda(f) = \sum_{(k,\ell)} \lambda_{k\ell}[f(t_k) - f(t_\ell)]$ for all $f \in \mathbb{R}^T$.

Lemma 6. *The following are equivalent:*
(i) *For every $\lambda \in \mathbb{R}_+^{(T \times T)}$, $D\lambda = \mathbf{0}$ implies that $\Delta v(\lambda) \leq 0$.*[20]
(ii) *There exists a net $\{\xi_\delta\}$ such that $\Delta v(t,s) \leq \liminf_\delta \xi_\delta(s) - \xi_\delta(t)$ 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 $A(\lambda) = (D\lambda, \Delta v(\lambda))$. Since $\mathbb{R}_+^{(T \times T)}$ is a cone, (i) fails if and only if there exists $\lambda \in \mathbb{R}_+^{(T \times T)}$ such that $D\lambda = \mathbf{0}$ and $\Delta v(\lambda) = 1$, i.e., $A(\lambda) = (\mathbf{0}, 1)$.

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

Clearly, A is linear and continuous, so by Lemma A.1, $A(\lambda) = (\mathbf{0}, 1)$ if and only if given any number ε, incentive scheme ξ and net $\{(w_\delta, \xi_\delta) \in \mathbb{R}_+^{T\times T} \times \mathbb{R}^T\}$,

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

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

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

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

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

Hence, (i) holds if and only if there exists a number $\varepsilon > 0$ and a net $\{\xi_\delta\}$ such that $\varepsilon \Delta v(t,s) \leq \liminf \xi_\delta(s) - \xi_\delta(t)$ 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. $\square$

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

The last step in our proof is to show that (ii) implies ex post implementability.

Lemma 7. *The following statements are equivalent:*
(i) *There exists a net $\{\xi_\delta\}$ such that $\Delta v(t,s) \leq \liminf_\delta \xi_\delta(s) - \xi_\delta(t)$ for all (t,s).*
(ii) *There exists an incentive scheme ξ such that $\Delta v(t,s) \leq \xi(s) - \xi(t)$ 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,

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

Hence, $\xi(t) = \liminf \xi_\delta(t)$ is bounded. Since the $\liminf$ 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), $\Delta v(t,s) \leq \liminf \xi_\delta(s) - \xi_\delta(t)$. Collecting these last two inequalities finally yields $\Delta v(t,s) \leq \xi(s) - \xi(t)$, as required. $\square$

D Proof of Theorem 5

If T is finite then the result follows by a similar argument to the one used to prove Lemma 5. 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 6, 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 7, 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 $\liminf$. 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.

References

Afriat, S. (1963): "The system of inequalities $a_{rs} > X_r - X_s$," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 59, 125. 21

——— (1967): "The construction of utility functions from expenditure data," *International Economic Review*, 67–77. 4, 21, 22

Bikhchandani, S., S. Chatterji, A. Sen, R. Lavi, A. Mu'alem, and N. Nisan (2006): "Weak Monotonicity Characterizes Deterministic Dominant-Strategy Implementation," *Econometrica*, 74, 1109–1132. 2

Carbajal, J. and J. Ely (2010): "Mechanism Design Without Revenue Equivalence," Working paper. 4

Clark, S. A. (2006): "Necessary and Sufficient Conditions for Infinite-Dimensional Linear Inequalities," *Positivity*, 10, 475–489. 27, 28

Cremer, J. and R. McLean (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica*, 56, 1247–1257. 4, 13, 15

Folland, G. (1999): *Real Analysis: Modern Techniques and their Applications*, John Wiley & Sons, second ed. 31

Gale, D. (1967): "A geometric duality theorem with economic applications," *The Review of Economic Studies*, 34, 19–24. 3, 28

Goberna, M. and M. López (1998): *Linear semi-infinite optimization*, John Wiley & Son Ltd. 27

Gretsky, N., J. Ostroy, and W. Zame (2002): "Subdifferentiability and the Duality Gap," *Positivity*, 6, 1–16. 28

Hart, S. and D. Schmeidler (1989): "Existence of Correlated Equilibria," *Mathematics of Operations Research*, 14, 18–25. 3, 4, 13, 17

Heydenreich, B., R. Müller, M. Uetz, and R. Vohra (2009): "Characterization of Revenue Equivalence," *Econometrica*, 77, 307–316. 2, 4, 18, 19

Ioffe, A. D. and V. M. Tikhomirov (1968): "Duality of Convex Functions and Extremum Problems," *Russian Mathematical Surveys*, 23, 53–124. 28

Kos, N. and M. Messner (2009): "Incentive Compatible Transfers," Mimeo. 4, 18, 19

Kretschmer, K. (1961): "Programmes in paired spaces," *Canad. J. Math*, 13, 221–238. 15

McAfee, R. and P. Reny (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395–421. 6

McFadden, D. (2005): "Revealed stochastic preference: a synthesis," *Economic Theory*, 26, 245–264. 4, 21, 22

McFadden, D. and M. Richter (1990): "Stochastic rationality and revealed stochastic preference," *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz*, 161–186. 4, 22

Müller, R., A. Perea, and S. Wolf (2007): "Weak monotonicity and Bayes-Nash incentive compatibility," *Games and Economic Behavior*, 61, 344–358. 4, 19

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73. 2

——— (1997): "Dual Reduction and Elementary Games," *Games and Economic Behavior*, 21, 183–202. 17

Nau, R. F. and K. F. McCardle (1990): "Coherent Behavior in Noncooperative Games," *Journal of Economic Theory*, 50, 424–444. 17

Rahman, D. (2008): "The Alternative to Equilibrium Existence," Working paper. 17

——— (2009): "Dynamic Implementation," Working paper. 2

——— (2010): "Correlated Information and Direct Mechanism Design," Working paper. 2, 16

Rahman, D. and I. Obara (2010): "Mediated Partnerships," *Econometrica*, 78, 285–308. 24

Richter, M. and K. Wong (2005): "Infinite inequality systems and cardinal revelations," *Economic Theory*, 26, 947–971. 21

Rochet, J. C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191–200. 1, 2, 4, 6, 7, 13, 15, 18, 21, 25, 27, 31, 32

Rockafellar, R. (1966): "Characterization of the subdifferentials of convex functions," *Pacific J. Math*, 17, 497–510. 13

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton University Press. 13, 27, 32

Segal, I. and M. D. Whinston (2009): "An Expected-Efficient Status Quo Allows Efficient Bargaining," *Theoretical Economics*, forthcoming. 4, 26